

Yanwen Wu (Ed.)

Software Engineering and Knowledge Engineering: Theory and Practice

 Springer

Advances in Intelligent and Soft Computing

Editor-in-Chief

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

Further volumes of this series can be found on our homepage: springer.com

Vol. 101. D. Dicheva, Z. Markov,
and E. Stefanova (Eds.)
*Third International Conference on Software,
Services and Semantic Technologies
S3T 2011, 2011*
ISBN 978-3-642-23162-9

Vol. 102. R.S. Choraś (Ed.)
*Image Processing and Communications
Challenges 3, 2011*
ISBN 978-3-642-23153-7

Vol. 103. T. Czachórski, S. Kozielski, and
U. Stańczyk (Eds.)
Man-Machine Interactions 2, 2011
ISBN 978-3-642-23168-1

Vol. 104. D. Jin and S. Lin (Eds.)
*Advances in Computer Science,
Intelligent System and Environment, 2011*
ISBN 978-3-642-23776-8

Vol. 105. D. Jin and S. Lin (Eds.)
*Advances in Computer Science,
Intelligent System and Environment, 2011*
ISBN 978-3-642-23755-3

Vol. 106. D. Jin and S. Lin (Eds.)
*Advances in Computer Science,
Intelligent System and Environment, 2011*
ISBN 978-3-642-23752-2

Vol. 107. P. Melo-Pinto, P. Couto, C. Serôdio,
J. Fodor, and B. De Baets (Eds.)
Eurofuse 2011, 2011
ISBN 978-3-642-24000-3

Vol. 108. Y. Wang (Ed.)
Education and Educational Technology, 2011
ISBN 978-3-642-24774-3

Vol. 109. Y. Wang (Ed.)
*Education Management, Education Theory
and Education Application, 2011*
ISBN 978-3-642-24771-2

Vol. 110. L. Jiang (Ed.)
*Proceedings of the 2011 International Conference
on Informatics, Cybernetics, and Computer
Engineering (ICCE 2011) November 19–20, 2011,
Melbourne, Australia, 2011*
ISBN 978-3-642-25184-9

Vol. 111. L. Jiang (Ed.)
*Proceedings of the 2011 International Conference
on Informatics, Cybernetics, and Computer
Engineering (ICCE 2011) November 19–20, 2011,
Melbourne, Australia, 2011*
ISBN 978-3-642-25187-0

Vol. 112. L. Jiang (Ed.)
*Proceedings of the 2011 International Conference
on Informatics, Cybernetics, and Computer
Engineering (ICCE 2011) November 19–20, 2011,
Melbourne, Australia, 2011*
ISBN 978-3-642-25193-1

Vol. 113. J. Altmann, U. Baumöl,
and B.J. Krämer (Eds.)
Advances in Collective Intelligence 2011, 2011
ISBN 978-3-642-25320-1

Vol. 114. Y. Wu (Ed.)
*Software Engineering and Knowledge
Engineering: Theory and Practice, 2012*
ISBN 978-3-642-03717-7

Vol. 115. Y. Wu (Ed.)
*Software Engineering and Knowledge
Engineering: Theory and Practice, 2012*
ISBN 978-3-642-25348-5

Yanwen Wu (Ed.)

Software Engineering and Knowledge Engineering: Theory and Practice

Volume 2



Springer

المنارة للاستشارات

Editor

Prof. Yanwen Wu
Central China Normal University
Lvting Yajing 10-3-102
Hongshan Qu
Wuhan, 430079
China
E-mail: wyw1970_cn@yahoo.com.cn

ISBN 978-3-642-25348-5

e-ISBN 978-3-642-25349-2

DOI 10.1007/978-3-642-25349-2

Advances in Intelligent and Soft Computing

ISSN 1867-5662

Library of Congress Control Number: 2011941486

© 2012 Springer-Verlag Berlin Heidelberg

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typeset by Scientific Publishing Services Pvt. Ltd., Chennai, India

Printed on acid-free paper

5 4 3 2 1 0

springer.com

المنارة للاستشارات

Preface

The 2009 Pacific-Asia Conference on Knowledge Engineering and Software Engineering (KESE 2009) will be held on December 19–20, 2009, Shenzhen, China . The two book are composed of 275 excellent papers selected from KESE 2009, which I will be delighted to introduce to you.

These high-quality papers of the Proceedings have been strictly peer-reviewed by famous experts, scholars and professors of the related fields from home and abroad. Therefore, I am sure you will gain lots of useful information from it. The purposes of the excellent Proceedings are to encourage more scholars and professors to participate in the conference, communicate with peers on the recent development of the related fields and ensure the correct research direction.

Knowledge engineering (KE) was defined in 1983 by Edward Feigenbaum, and Pamela McCorduck as follows: KE is an engineering discipline that involves integrating knowledge into computer systems in order to solve complex problems normally requiring a high level of human expertise.

At present, it refers to the building, maintaining and development of knowledge-based systems. It has a great deal in common with software engineering, and is used in many computer science domains such as artificial intelligence, including databases, data mining, expert systems, decision support systems and geographic information systems. Knowledge engineering is also related to mathematical logic, as well as strongly involved in cognitive science and socio-cognitive engineering where the knowledge is produced by socio-cognitive aggregates (mainly humans) and is structured according to our understanding of how human reasoning and logic works.

Software Engineering (SE) is a profession dedicated to designing, implementing, and modifying software so that it is of high quality, affordable, maintainable, and fast to build. It is a “systematic approach to the analysis, design, assessment, implementation, test, maintenance and reengineering of software, that is, the application of engineering to software.” The term software engineering first appeared in the 1968 NATO Software Engineering Conference, and was meant to provoke thought regarding the perceived “software crisis” at the time. The IEEE Computer Society's Software Engineering Body of Knowledge defines “software engineering” as the application of a systematic, disciplined, quantifiable approach to the development, operation, and maintenance of software, and the study of these approaches; that is, the application of engineering to software. It is the application of Engineering to software because it integrates significant mathematics, computer science and practices whose origins are in Engineering.

KESE 2009 is to provide a forum for researchers, educators, engineers, and government officials involved in the general areas of Knowledge Engineering and Software Engineering to disseminate their latest research results and exchange views on the future research directions of these fields.

Special thanks for editors and experts from home and abroad. We also thank every participant. Without you, nothing could be done.

Organization

General Chairs

Chin-Chen Chang
Biswanath Vokkarane

Feng Chia University, Taiwan
Society on Social Implications of Technology and
Engineering

Publication Chair

Yanwen Wu

Huazhong Normal University, China

Organizing Chairs

Khine Soe Thaug

Maldives College of Higher Education, Maldives

Program Chair

Junwu Zhu

Yangzhou University, China

International Committee

Junwu Zhu
He Ping
Yiyi Zhouzhou
Garry Zhu
Ying Zhang
David Wang

Yangzhou University, China
Liaoning Police Academy, China
Azerbaijan State Oil Academy, Azerbaijan
Thompson Rivers University, Canada
Wuhan University, China
IEEE Nanotechnology Council Cambodia Chapter
Past Chair, Cambodia
ACM NUS Singapore Chapter, Singapore
ACM NUS Singapore Chapter, Singapore
National University of Singapore, Singapore
Jiangsu Teachers University of Technology, China

Contents

Knowledge Engineering and Communication Technology

- Analytic Solutions of a Second-Order Functional Differential Equation** 1
Lingxia Liu
- Analytic Solutions of an Iterative Functional Differential Equation with State Dependent Delay** 9
Lingxia Liu
- The Study of Optimal Conditions of Electroporation in *Gluconacetobacter Xylinum*** 17
Huixia Zhu, Shiru Jia, Hongjiang Yang, Lin Yan, Jing Li
- Study on the Mechanism of Apoptosis in the Swine Small Intestine Epithelium Treated by Heat Stress** 25
Dan Jia, Fenghua Liu, Weili Luan, Kaijun Guo, An Lu, Jin Yu
- Effect of Spleen Deficiency Syndrome on the Expression of TLR2 and TLR4 in Rats** 33
Tao Liu, An Lu, Xiaoyu Zhu, Nuowei Zhang, Guilin Cheng, Ning Wang, Fenghua Liu, Kaijun Guo, Huichuan Wang
- Reserpine Caused Gastrointestinal Damages and Hormonal Changes in Rats** 41
Kaijun Guo, Fei Cheng, Xiaorui Wang, Xiaoping Song, Xiaoyu Zhu, Jingyi Gan, Fenghua Liu, Jianqin Xu
- Heat Stress Induced Apoptosis in Rat Intestinal Epithelial Cell Line-6 (IEC-6)** 49
Kaijun Guo, Weili Luan, Huichuan Wang, Jin Yu, Ning Wang, Guilin Cheng, Fenghua Liu

| | |
|---|-----|
| Evaluation of Chinese New Manufacturing Development Based on Factor Analysis | 57 |
| <i>Wei Sun</i> | |
| Research of Evaluation of New Computer Complex Talents | 63 |
| <i>Wei Sun</i> | |
| Optimal Iterative Learning Control for Nonlinear Discrete-Time Systems | 69 |
| <i>Hong-wei Xu</i> | |
| The Research and Design of Realizing Interconnection and Intercommunication Local Public Emergency Platform | 77 |
| <i>Gu Yan</i> | |
| The Normal Range of House Price Rising Rate of China | 85 |
| <i>Zhang Hongli</i> | |
| A Bottom-Up Incremental Algorithm of Building Concept Lattice | 91 |
| <i>LingLing Lv, Lei Zhang, PeiYan Jia, FuNa Zhou</i> | |
| Implementing an Active Database for Maintaining Asset Data | 99 |
| <i>Rozana Kasbon, Sofia Shaharom, Eliza Mazmee Mazlan, Saipunidzam Mahamad</i> | |
| The Design and Implementation of a Gesture-Driven System for Intelligent Wheelchairs Based on the Orientation Histogram Method | 109 |
| <i>Yuan Luo, Yu Xie</i> | |
| A Flash-Based Framework for Learning Thai Language as Second Language in Preschool Education | 115 |
| <i>Saipunidzam Mahamad, Yossawadee Kaoemane, Mohammad Noor Ibrahim, Rozana Kasbon</i> | |
| Emergency Personnel Evacuation Model and Its Application Based on Cellular Automaton | 123 |
| <i>Demin Wu, Wenyong Dong, Jiangpeng Wang</i> | |
| Software Protection with Encryption and Verification | 131 |
| <i>Wentao Liu</i> | |
| Improved Tri-training with Unlabeled Data | 139 |
| <i>Tao Guo, Guiyang Li</i> | |
| Improved Particle Swarm Optimization Based on Genetic Algorithm | 149 |
| <i>Chunhong Dou, Jinshan Lin</i> | |
| A Scheme of LSB Steganography Based on Concept of Finding Optimization Pixels Selection | 155 |
| <i>Jianhong Sun, Yingjiang Li, Xiaohui Zhong, Junsheng Li</i> | |

| | |
|--|-----|
| Research on Intelligent Generating Test Paper Based on Parallel Genetic Algorithm | 161 |
| <i>Jianjun Li, Meng Wang</i> | |
| Design of Fault Diagnosis System for the Parameter-Depended Control System | 169 |
| <i>Jing Zhou</i> | |
| Research of Method about Geomagnetic Sensor Data Processing by Optimal Idea | 177 |
| <i>Chanyuan Liu</i> | |
| Design and Implementation of a Communication Simulation Training System Based on HLA | 181 |
| <i>Chen Bin, Li Juan</i> | |
| Numerical Simulation on Heat Transfer from Envelope of the Underground Engineering | 189 |
| <i>Wang Li-jun, Miao Xiao-ping, Wang Rui-hai, Wang Bo</i> | |
| A Reversible Watermarking Scheme for 2D Vector Maps | 197 |
| <i>Dan Wu</i> | |
| Research of Mobile Public Transport Information System Based on J2ME and SVG | 205 |
| <i>Zhang Suqi, Gu Junhua, Li Weini, Gu Qiang</i> | |
| Based on the ExtJS Technology and SSH Framework Authority Management Research | 213 |
| <i>He Yueshun, Luo Min</i> | |
| Public-Private Partnerships and E-Government | 221 |
| <i>Faizullah Khan, Surat Khan, Bin Zhang</i> | |
| Reverse e-Logistics for SMEs in Pakistan | 229 |
| <i>Surat Khan, Faizullah Khan, Bin Zhang</i> | |
| A Teaching Method of “In One or Two Words” in Higher Mathematics | 239 |
| <i>Zhibin Li, Xiaoying Sun, Jing Chang</i> | |
| Ponder on a New Internet Community Space Pattern with Ubiquitous-Network Service | 247 |
| <i>Zhao Hongyu, Chang Sheng</i> | |
| HAZOP Analysis-Based Method on the Risk Assessment of the Main Engine of the 1000TEU Container Ship | 257 |
| <i>Yulong Zhan, Yitong Hao, Yao Yu, Di Wu</i> | |
| ‘Peer to Peer’ Network Graphics Art Design | 265 |
| <i>Lv Jiefeng, Ge Xiazhi</i> | |

| | |
|--|-----|
| Design of Musical Performance Mode Based on Cooperative Learning | 271 |
| <i>Yumeng He</i> | |
| Building Distance Education and Training System for Police by Making Use of Information Technology | 277 |
| <i>Shuxia Liu</i> | |
| Research on Mathematica Experimental Teaching Reform of High Vocational Education | 285 |
| <i>Jing Ruan</i> | |
| ERP Electronic Sand Table Based on Web Service | 291 |
| <i>Zhang Changsheng, Zou Huasheng</i> | |
| A Mathematical Modeling of the Best Design of Pipeline Construction | 299 |
| <i>Ruan Jing, Lin Bin, Xiang Haifei</i> | |
| Trust-Based Privacy Authorization Model for Web Service Composition . . . | 307 |
| <i>Jun Zheng, Zhiqiu Huang, Jun Hu, Ou Wei, Linyuan Liu</i> | |
| Talent Training of Software Testing Based on University-Enterprise Cooperation | 315 |
| <i>Zhang Shanwen, Zhou Zhengguang</i> | |
| Using Surface Electromyography Signals to Control an Exoskeleton Arm Driven by Pneumatic Artificial Muscles | 321 |
| <i>Jie Liu, Yu Wang, Heting Tong, Ray P.S. Han</i> | |
| A Study on Regression Analysis of Bodyof 4-6 Years Old Children in Henan Area | 327 |
| <i>Wang Shi-zhu, Zhou Ping</i> | |
| Recognition and Evaluation of Information Security Based on Non-optimum Analysis | 335 |
| <i>Lihua Duan</i> | |
| The Analysis of Two Dimensional Resource Allocation Procedures for IEEE802.16m of Mobile WiMAX Systems | 343 |
| <i>Young-il Kim, Otgonbayar Bataa, Bat-Enkh Oyunbileg, Khishigjargal Gonchigsumlaa</i> | |
| DAIM: A Distributed Algorithm for Isolating Malfunctioning Nodes in Wireless Sensor Networks | 351 |
| <i>Mohamed K. Watfa, Rawad Abu Assi</i> | |
| IPurse: An Innovative RFID Application | 359 |
| <i>Mohamed K. Watfa, Manprabhjot Kaur, Rashida Firoz Daruwala</i> | |

| | |
|--|-----|
| An Epistemic Logic Based on Change and the Paradox of Unknowability | 367 |
| <i>Lei Liu</i> | |
| Criminal Investigation DSS Based on Trust Intuition Analysis Model | 375 |
| <i>Lihua Duan</i> | |
| Crime Data Mining Based on Extension Classification | 383 |
| <i>Weidong Tao</i> | |
| A New Data Mining Approach with Data-Intuition Cooperative Analysis ... | 391 |
| <i>Weidong Tao</i> | |
| A Discovery for the Association Rules of Operation Sequence Based on Improved Apriori Algorithm | 399 |
| <i>Tan Yanhua</i> | |
| A Parameter-Driving Method of Redeveloping Mould Standard-Part Library Based on SolidWorks | 407 |
| <i>Tan Yanhua</i> | |
| Study on Magnesium Alloys Ignition Temperature Test System | 415 |
| <i>Li Yongyan, Zhao Weimin, Xue Haitao, Ding Jian</i> | |
| Defect Recognition of Resistance Spot Welding Based on Artificial Neural Network | 423 |
| <i>Li Yongyan, Zhao Weimin, Xue Haitao, Ding Jian</i> | |
| Theoretical Research on University Outward Bound Curriculum System ... | 431 |
| <i>Xinhua Zhu</i> | |
| The Effect of Guangzhou's Temperature Change to the Electric Power Consumption | 439 |
| <i>Feng Jing, Pan Anding</i> | |
| The Relationship and Strategy of Informatization on Farmers, Agriculture and Rural Areas and New Rural Construction | 445 |
| <i>Li Quan-guo, Kang Ling</i> | |
| Measuring the Efficiency of Rural Informatization in China 2007-2009: An Application of Data Envelopment Analysis | 451 |
| <i>Xiaochun Xu, Jingdong Luan</i> | |
| Research on Software Architecture Base on Data Snap Technology | 459 |
| <i>Dan Guo, Yu Zhang</i> | |
| Research of Agricultural Land Classification and Evaluation Based on Genetic Algorithm Optimized Neural Network Model | 465 |
| <i>Liu TingXiang, Zhang ShuWen, Wu QuanYuan, Bao WenDong</i> | |

| | |
|---|-----|
| Evaluation of Agricultural Water Use Efficiency and Regional Irrigation Strategies in China Based on GIS | 473 |
| <i>Li Yuyi, Pang Huancheng, Yu Tianyi, Li Hua, Li Yibing, Ren Tianzhi</i> | |
| The Research on Network Courses Education Resources Integration Based on P.E. | 483 |
| <i>Xu Xiaoyang</i> | |
| Design of Field Information Collection System Based on uC/OS-II | 489 |
| <i>ZhiLai Zhang, Yan Li, ZhiAn Wang</i> | |
| Analysis and Prediction of the Total Number of Harbin Ice-Snow Tourism Based on Times Series | 495 |
| <i>Yu Zhang</i> | |
| ZPD Incidence Development Strategy for Demand of Internet in Business — Teaching of New and Old Comprehensive Regional Higher Education Institutes | 503 |
| <i>Hong Liu</i> | |
| Analysis of Changes of Agricultural Tax System | 507 |
| <i>Zhang Huiqing</i> | |
| Application Research of Smart Water Systems Based on I2C Bus | 515 |
| <i>Shi Binbin, Hu Zhifen, Dai Minli</i> | |
| Evaluation of Mountain Road Influences on Driving Fatigue | 521 |
| <i>Xiao Jinjian, Huang Simin, Wang Yunsong</i> | |
| Multi-channel Feedback Data-Curtail Identification | 529 |
| <i>Bo Zhao, Qinrang Liu, Bo Yuan</i> | |
| Research on Parametric Design of Hydraulic Retarder Cascade | 535 |
| <i>Zou Bo, Wei Wei, Yan Qing-dong</i> | |
| A Hardware Design of CCM3118 Evaluation System | 543 |
| <i>Xu Feng-liang, Shi Binbin</i> | |
| RNP Flight Procedure Analysis and Simulation | 551 |
| <i>Lan Ma, Ying Lin</i> | |
| Infrastructure-Aid Information Forwarding Based on Network Coding in Vehicular Ad Hoc Networks | 557 |
| <i>Gang Wang, Yi Xu, Xia Dai</i> | |
| Research on Evaluation and Optimization of Agricultural Technology Innovation Resources Allocation in China | 567 |
| <i>Jingcui Sun, Qingran Guo</i> | |

| | |
|---|-----|
| CJSYS-A01 Piezoelectric Fluidic Angular Rate Sensor | 575 |
| <i>Xing Wang, Linhua Piao, Quangan Yu</i> | |
| Research on Evaluation of Clients' Service Performance to Bonded Logistic Enterprise Based on Entropy Weight | 581 |
| <i>Li Zhen-fu, He Jian-tong, Chen Ping</i> | |
| Necessity Analysis of Inside Shoulder Setting in Multi-lane Freeway with Traffic Simulation Method | 593 |
| <i>Zhong Liande, Li Xiuwen, Wu Keman</i> | |
| Agricultural Products Quality Cost-Allocating Based on Supply Chain Management | 599 |
| <i>Wang Xin-li, Wei Qiong</i> | |
| Research on Role of Learning Organizational Theory to the P.E. Majors in Teachers College | 611 |
| <i>Yao Bin-bin</i> | |
| Enterprise Logistics Risk Factors Analysis Based on Interpretative Structural Modeling Method | 617 |
| <i>Liu Yongsheng, Li Yanru</i> | |
| A Study and Development of Onboard LED Lighting for Train | 627 |
| <i>Deqiang He, Yixuan Bai, Xiongwen Yang, Xiaochun Su, Zengqi Yang</i> | |
| Study on Agricultural Products Logistics Mode in Henan Province of China | 635 |
| <i>Wang Liping</i> | |
| The Research on Torsional Vibration Signal Processing and Automobile Transmission Shaft | 641 |
| <i>Liu Xiaoqun, Chen Hu</i> | |
| Pattern Analysis of the CRM Application in the China's Telecommunication Industry | 649 |
| <i>Li Wen, Wu Meng-yun, Du Jian-guo</i> | |
| Service-Oriented Data Integration Using GIS: An Anchorage Management Tool for Wenzhou Government, China | 657 |
| <i>Li Yubao, Li Longchang, Liao Fei</i> | |
| An Index Assignment Algorithm over Noisy Channel | 663 |
| <i>Wang Yue</i> | |
| Application of Artificial Neural Network (ANN) for Prediction of Power Load | 673 |
| <i>Fei Xia, Li Fan</i> | |

| | |
|---|-----|
| The Research on Tracking Concept Drift Based on Genetic Algorithm | 679 |
| <i>Zhong Zhishui</i> | |
| Prototype Design of Hierarchical Cluster Multiprocessor on Platform FPGA | 687 |
| <i>RuiHui Mu, PeiZhang Zhao, Dong Li</i> | |
| Efficient Memory Processors Design of Multiple Applications for Multiprocessor Architecture | 693 |
| <i>Liang Tian, Hong Pan, Dong Li</i> | |
| Research of Discrete Element Method in Vibratory Feeder | 699 |
| <i>Yaluo Yang, Yayu Huang, Hao Wu</i> | |
| The Research on Type Transformation of Multi-loop Spatial Linkages Kinematics Analysis | 705 |
| <i>Li Wenxing, Lian Liming</i> | |
| Semi-supervised Data Stream Ensemble Classifiers Algorithm Based on Cluster Assumption | 713 |
| <i>Wang Xuejun</i> | |
| Object-Oriented Design and Implementation of Embedded System IDE Software | 723 |
| <i>Guo Wushi, Yi Xin</i> | |
| Development and Application of Cloud Computing in China Information Resource Management Field | 729 |
| <i>Pang Jingxia</i> | |
| The Comparison between Two Control Methods about Simulation of Traction Load | 737 |
| <i>Jingyu Hu, Min Xu, Kai Guo, Yongqiang Cui</i> | |
| A Notable Problem on t-Distribution | 745 |
| <i>Shen Maoxing, Li Jun, Zhang Yaoping</i> | |
| SVM-Based Pornographic Images Detection | 751 |
| <i>Haiming Yin, Xiangqiong Huang, Yuanwang Wei</i> | |
| Secure Map Reduce Data Transmission Mechanism in Cloud Computing Using Threshold Secret Sharing Scheme | 761 |
| <i>Hua Yi Lin, Che-Yu Yang, Meng-Yen Hsieh</i> | |
| Superior-in-Status Analysis of Improved Genetic Algorithm for GTSP | 771 |
| <i>Tan Yang, Hao Zhi-feng, Cai Zhaoquan, Huang Han</i> | |
| The Development of Software and Its Infulence on Digital Art | 779 |
| <i>Cheng Yin, Shengyong Mu, Jing Liu, Maopo Yin</i> | |

| | |
|--|-----|
| An Evaluation Model for Synchronized Supply Chain Based on Entropy Weight and TOPSIS I | 787 |
| <i>Li Zhiyong, Wang Xiang, Zheng Nan</i> | |
| An Evaluation Model for Synchronized Supply Chain Based on Entropy Weight and TOPSIS II | 795 |
| <i>Li Zhiyong, Wang Xiang, Zheng Nan</i> | |
| Design of an Online Parameters Configuration Based on C8051F USB Processors Family | 801 |
| <i>Zhu Lei, Zhi Wenxia, Zhang Gaoqiang</i> | |
| Study on Theory of Model Test in Model Credibility Evaluation | 809 |
| <i>Hao Li, Shuo Tang, Xiaodong Yan</i> | |
| The Learning Evaluation of Learning Enterprises Based on BP Neural Network | 819 |
| <i>Yong Zhang</i> | |
| An Approximate Algorithm to Solve the Location-Selection of Wireless Network Problem | 827 |
| <i>Xiao-xu Lu, Wei Zhang</i> | |
| A New Supervised Discriminant Locality Preserving Projections Algorithm | 833 |
| <i>Jun Yu, Jintao Meng, Xiao-xu Lu</i> | |
| Theory and Application of Monte Carlo Method | 841 |
| <i>Liu Xinhua, Liu Yongzhi, Liu Hao</i> | |
| A New XML Database System Using Pagination Technique | 849 |
| <i>Xiaopeng Wang</i> | |
| Design and Implementation of Multimedia Format Converter Based on FFmpeg | 857 |
| <i>Yun Cheng, Qingtang Liu, Chengling Zhao, Xiaoliang Zhu, Guoqing Zhang</i> | |
| Design and Implementation of MediaPlayer Based on FFmpeg | 867 |
| <i>Yun Cheng, Qingtang Liu, Chengling Zhao, Xiaoliang Zhu, Guoqing Zhang</i> | |
| The Research and Implementation of Incident Response Information System | 875 |
| <i>Zuyi Chen, Hua Yong, Taixiang Zhao</i> | |
| The Application of Semantic Web in E-Learning Personalized Knowledge Query | 881 |
| <i>Zuyi Chen, Liu Fang, Taixiang Zhao</i> | |
| QoS-Based Web Service Selection Approach | 887 |
| <i>Guofeng Chang</i> | |

| | |
|--|-----|
| Development Strategy of the Digital Library in University Based on SWOT Analysis | 893 |
| <i>Lijuan Yang</i> | |
| Quality Evaluation on Troops Transport Capacity Based on Interval Grey Numbers Multi-attribute Decision-Making Method | 901 |
| <i>Lu Bangjun, Dai Xiliang</i> | |
| Explorations of and Reflections on the Educational Model of Software Schools | 909 |
| <i>Jingfeng Yan, Guoqing Li</i> | |
| The Research and Contrast of the Hybrid Intrusion Detection | 915 |
| <i>Meng JianLiang, Yu Yang</i> | |
| A New Adaptive Stopping Criterion for BICM-ID System | 921 |
| <i>Xiao Ying, Li Jianping, Cai Chaoshi</i> | |
| Data Integration of Heterogeneous Data Source in Multi-parameter Test Processing | 929 |
| <i>Wang Guitang, Liu Wenjuan, Jiang Yuelong</i> | |
| The Unbounded Parallel-Batching Scheduling Problem with Family Jobs to Minimize Makespan | 935 |
| <i>Jintao Meng, Zhiyong Li</i> | |
| Research on Layout Evaluation Indexes System of Dangerous Goods Logistics Port Based on AHP | 943 |
| <i>Zhang Peilin, Mao Jian, Yang Long</i> | |
| An Improved BP Neural Network Algorithm Embedded with Logistic Mapping and Its Application | 951 |
| <i>Wang Tao</i> | |
| The Analytics of College Sports Participation and Implementation of the Path of Sun Sports | 959 |
| <i>Zhang Run-tao</i> | |
| Using the Least Repeaters to Connect Radios | 965 |
| <i>Danying Chen, Dongping Yu, Zheng Wang, Yuehua Wan</i> | |
| Research on a Defogging Method of Fog-Degrade Image Based on Depth Region Segmentation | 973 |
| <i>Yi Yang</i> | |
| An Algorithm of Incremental Construction Voronoi Region | 979 |
| <i>Pengfei Xu, Zhigang Chen, Ming Chen</i> | |
| A Brief Analysis of Geocoding | 987 |
| <i>Lijing Zhang, Jing Yi</i> | |

| | |
|---|------|
| Study the Method of the Optimum Path Based on GIS Features | 995 |
| <i>Lijing Zhang, Jing Yi</i> | |
| Route Search Base on pgRouting | 1003 |
| <i>Lijing Zhang, Xuanhui He</i> | |
| Research on the Improved Algorithm of Node Split | 1009 |
| <i>Lijing Zhang, Xuanhui He, Yulan Ren</i> | |
| Modeling Research on Decision-Making Strategies in Human-Computer and Human-Human Ultimatum Game | 1015 |
| <i>Dan Xiao, Kerong He, Jianfeng Hu</i> | |
| Author Index | 1023 |

Analytic Solutions of a Second-Order Functional Differential Equation

Lingxia Liu

Department of Mathematics, Weifang University,
Weifang, Shandong 261061, P.R. China
llxmath@126.com

Abstract. This paper is concerned with the existence of analytic solutions of a second-order functional differential equation. As well as in many previous works, we reduce this problem to finding analytic solutions of a functional differential equation without iteration of the unknown function x . For technical reasons, in previous works the constant α given in the Schröder transformation is required to fulfil that α is off the unite circle or lies on the circle with the Diophantine condition. In this paper, we focus on those α on the unit circle, i.e. $|\alpha|=1$. We discuss not only those α 's at resonance, i.e. at a root of the unity, but also those α 's near resonance under the Brjuno condition, where the Diophantine condition is not required.

Keywords: Analytic solution, functional differential equation, resonance, Diophantine condition; Brjuno condition.

1 Introduction

In the last few years there has been a growing interest in studying functional differential equations with state dependent delay. Iterative differential equation, as a special type of functional differential equations, attracted the attention of researchers recently [1-10]. In [6-8], analytic solutions of the state dependent functional differential equations $x'(z) = x(az + bx(z))$, $\alpha z + \beta x'(z) = x(az + bx''(z))$ and

$$x''(z) = x(az + bx'(z)), \quad a, b \in C, b \neq 0 \quad (1)$$

are found. For Eq. (1), as in [8], let $y(z) = az + bx'(z)$, then for any number z_0 , we have

$x(z) = x(z_0) + \frac{1}{b} \int_{z_0}^z (y(s) - as) ds$, so $x(y(z)) = x(z_0) + \frac{1}{b} \int_{z_0}^{y(z)} (y(s) - as) ds$. Therefore, in

view of (1) and $x''(z) = \frac{1}{b}(y'(z) - a)$, we have

$$\frac{1}{b}(y'(z) - a) = x(z_0) + \frac{1}{b} \int_{z_0}^{y(z)} (y(s) - as) ds. \quad (2)$$

If z_0 is a fixed point of $y(z)$, i. e., $y(z_0) = z_0$, we see that $x(z_0) = \frac{1}{b}(y'(z_0) - a)$. Furthermore, differentiating both sides of (2) with respect to z , we obtain

$$y''(z) = [y(y(z)) - ay(z)]y'(z) \tag{3}$$

To find analytic solutions of (3), we first seek an analytic solution $g(z)$ of the auxiliary equation

$$\alpha g''(\alpha z)g'(z) = g'(\alpha z)g''(z) + (g'(z))^2 g'(\alpha z)[g(\alpha^2 z) - ag(\alpha z)] \tag{4}$$

satisfying the initial value conditions

$$g(0) = \mu, g'(0) = \eta \neq 0, \tag{5}$$

Where μ, η are complex numbers. For technical reasons, in previous works [8], the constant α given in the Schröder transformation is required to fulfil that (i) $0 < |\alpha| < 1$, (ii) $|\alpha| = 1$, α is not a root of unity in \mathbf{C} , and satisfies the Diophantine condition: $|\mu| = 1$, μ is not a root of unity, and $\log \frac{1}{|\mu^n - 1|} \leq T \log n$, $n = 2, 3, \dots$ for some positive constant T . In this paper, we break the restriction of Diophantine condition and study not only those α at resonance, i.e., at a root of the unity, but also those α near resonance under the Brjuno condition as in [7,9,10]. Now we discuss the analytic solutions of Eq. (4) in the following condition:

(H1) $\mu = e^{2\pi i \theta}$, where $\theta \in \mathbf{R} \setminus \mathbf{Q}$ is a Brjuno number [11,12], i.e., $B(\theta) = \sum_{k=0}^{\infty} \frac{\log q_{k+1}}{q_k} < \infty$, where $\{p_k / q_k\}$ denotes the sequence of partial fraction of the continued fraction expansion of θ , and is said to satisfy the Brjuno condition.

(H2) $\alpha = e^{2\pi i q^l / p}$ for some integers $p \in \mathbf{N}$ with $p \geq 2$ and $q \in \mathbf{Z} \setminus \{0\}$, and $\alpha \neq e^{2\pi i l / k}$ for all $1 \leq k \leq p-1$ and $l \in \mathbf{Z} \setminus \{0\}$.

Observe that α is a p -th unit root (or called p -order resonance) in the case of (H_2) , while the case (H_1) contains a part of α “near” resonance, neither being roots of the unity nor satisfying the Diophantine condition. In this paper, considering the Brjuno condition instead of the Diophantine one, we discuss not only the cases (H1) but also (H2) for iterative functional differential equation (1).

2 Analytic Solution of Auxiliary Equation

As stated in [13], for a real number θ we let θ denote its integer part and let $\{\theta\} = \theta - [\theta]$. Then every national number θ has a unique expression of the Gauss' continued fraction $\theta = a_0 + \theta_0 = a_0 + \frac{1}{a_1 + \theta_1} = \dots$, denoted simply by $\theta = [a_0, a_1, \dots, a_n, \dots]$, where a_j 's and θ_j 's are calculated by the algorithm: (a) $a_0 = [\theta]$, $\theta_0 = \{\theta\}$ and (b)

$a_n = \left[\frac{1}{\theta_{n-1}} \right], \theta_n = \left\{ \frac{1}{\theta_{n-1}} \right\}$ for all $n \geq 1$. Define the sequences $(p_n)_{n \in \mathbb{N}}$ and $(q_n)_{n \in \mathbb{N}}$ as follows

$$q_{-2} = 1, q_{-1} = 0, q_n = a_n q_{n-1} + q_{n-2}; \quad p_{-2} = 0, p_{-1} = 1, p_n = a_n p_{n-1} + p_{n-2}.$$

It is easy to show that $p_n / q_n = [a_0, a_1, \dots, a_n]$. Thus, for every $\theta \in \mathbf{R} \setminus \mathbf{Q}$ we associate, using its convergence, an arithmetical function $B(\theta) = \sum_{n \geq 0} \frac{\log q_{n+1}}{q_n}$. We say that θ is a Brjuno number or that it satisfies Brjuno condition if $B(\theta) < +\infty$. The Brjuno condition is weaker than the Diophantine condition. For example, if $a_{n+1} \leq c e^{a_n}$ for all $n \geq 0$, where $c > 0$ is a constant, then $\theta = [a_0, a_1, \dots, a_n, \dots]$ is a Brjuno number but is not a Diophantine number. So, the case (H1) contains both Diophantine condition and a part of μ "near" resonance.

In order to discuss the existence of the auxiliary equation (4) under (H1), we need to introduce Davie's Lemma. First, we recall some facts in [14] briefly. Let $\theta \in \mathbf{R} \setminus \mathbf{Q}$ and $(q_n)_{n \in \mathbb{N}}$ be the sequence of partial denominators of the Gauss's continued fraction for θ as in the Introduction. As in [13], let $A_k = \{n \geq 0 \mid \|n\theta\| \leq \frac{1}{8q_k}\}$, $E_k = \max(q_k, \frac{q_{k+1}}{4})$, $\eta_k = \frac{q_k}{E_k}$. Let A_k^* be the set of integers $j \geq 0$ such that either $j \in A_k$ or for some j_1 and j_2 in A_k , with $j_2 - j_1 < E_k$, one has $j_1 < j < j_2$ and q_k divide $j - j_1$. For any integer $n \geq 0$, define

$$l_k(n) = \max \left(\left(1 + \eta_k\right) \frac{n}{q_k} - 2, \left(m_n \eta_k + n\right) \frac{1}{q_k} - 1 \right),$$

Where $m_n = \max\{j \mid 0 \leq j \leq n, j \in A_k^*\}$. We then define function $h_k : \mathbf{N} \rightarrow \mathbf{R}_+$ as follows:

$$\begin{cases} \frac{m_n + \eta_k n}{q_k} - 1, & \text{if } m_n + q_k \in A_k^*, \\ l_k(n), & \text{if } m_n + q_k \notin A_k^*. \end{cases} \quad \text{Let } g_k(n) := \max \left(h_k(n), \left\lfloor \frac{n}{q_k} \right\rfloor \right),$$

and define $k(n)$ by the condition $q_{k(n)} \leq n \leq q_{k(n)+1}$. Clearly, $k(n)$ is non-decreasing. Then we are able to state the following result:

Lemma 1. (Davie's Lemma [14]) Let

$$K(n) = n \log 2 + \sum_{k=0}^{k(n)} g_k(n) \log(2q_{k+1}). \quad \text{Then}$$

(a) There is a universal constant $\gamma > 0$ (independent of n and θ) such that

$$K(n) \leq n \left(\sum_{k=0}^{k(n)} \frac{\log q_{k+1}}{q_k} + \gamma \right),$$

(b) $K(n_1) + K(n_2) \leq K(n_1 + n_2)$ for all n_1 and n_2 , and

(c) $-\log |\alpha^n - 1| \leq K(n) - K(n-1)$.

Now we state and prove the following theorem under Brjuno condition. The idea of our proof is acquired from [14].

Theorem 1. Suppose (H1) holds. Then for the initial value conditions (5), equation (4) has an analytic solution of the form

$$y(z) = \mu + \eta Z + \sum_{n=2}^{\infty} b_n z^n \tag{6}$$

in a neighbourhood of the origin.

Proof. As in [8], rewrite (4) in the form

$$\frac{\alpha g''(\alpha z)g'(z) - g'(\alpha z)g''(z)}{(g'(z))^2} = g'(\alpha z)[g(\alpha^2 z) - ag(\alpha z)],$$

or

$$\left(\frac{g'(\alpha z)}{g'(z)}\right)' = g'(\alpha z)[g(\alpha^2 z) - ag(\alpha z)].$$

Therefore, in view of $g'(0) = \eta \neq 0$, we have

$$g'(\alpha z) = g'(z) \left[1 + \int_0^z g'(\alpha s)(g(\alpha^2 s) - ag(\alpha s)) ds \right]. \tag{7}$$

We now seek a solution of (7) in the form of a power series (6). By defining $b_0 = \mu$ and $b_1 = \eta$ then substituting (6) into (7), we see that the sequence $\{b_n\}_{n=2}^{\infty}$:

$$(\alpha^{n+1} - 1)(n + 2)b_{n+2} = \sum_{k=0}^n \sum_{j=0}^{n-k} \frac{(k+1)(j+1)(\alpha^{2(n-k)-j} - a\alpha^{n-k})}{n-k+1} b_{k+1} b_{j+1} b_{n-k-j}, \quad n = 0, 1, \dots \tag{8}$$

is successively determined in a unique manner. Now we show that the power series (6) converges in a neighbourhood of the origin. First of all, note that

$$\left| \frac{(k+1)(j+1)(\alpha^{2(n-k)-j} - a\alpha^{n-k})}{(n+2)(n-k+1)} \right| \leq 1 + |a|,$$

then we have

$$|b_{n+2}| \leq \frac{M}{|\alpha^{n+1} - 1|} \sum_{k=0}^n \sum_{j=0}^{n-k} |b_{k+1}| \cdot |b_{j+1}| \cdot |b_{n-k-j}|, \quad n = 0, 1, 2, \dots, \tag{9}$$

where $M = 1 + |a|$. If we define a sequence $\{B_n\}_{n=0}^{\infty}$ by $B_0 = |\mu|$, $B_1 = |\eta|$ and

$B_{n+2} = M \sum_{k=0}^n \sum_{j=0}^{n-k} B_{k+1} B_{j+1} B_{n-k-j}$, $n = 0, 1, \dots$. Then in view of (8),

$$|b_{n+1}| \leq \frac{B_{n+1}}{|\alpha^n - 1|}, \quad n = 0, 1, 2, \dots. \tag{10}$$

As in [8], we define

$$G(z) = \sum_{n=0}^{\infty} B_n z^n, \tag{11}$$



Then

$$G^2(z) = (|\mu| + \sum_{n=0}^{\infty} B_{n+1} z^{n+1}) (\sum_{n=0}^{\infty} B_n z^n) = |\mu| \sum_{n=0}^{\infty} B_n z^n + \sum_{n=0}^{\infty} (\sum_{k=0}^n B_{k+1} B_{n-k}) z^{n+1},$$

and

$$\begin{aligned} G^3(z) &= (|\mu| + \sum_{n=0}^{\infty} B_{n+1} z^{n+1}) (|\mu| \sum_{n=0}^{\infty} B_n z^n + \sum_{n=0}^{\infty} (\sum_{k=0}^n B_{k+1} B_{n-k}) z^{n+1}) \\ &= |\mu|^2 \sum_{n=0}^{\infty} B_n z^n + 2|\mu| \sum_{n=0}^{\infty} (\sum_{k=0}^n B_{k+1} B_{n-k}) z^{n+1} + \sum_{n=0}^{\infty} (\sum_{k=0}^n \sum_{j=0}^{n-k} B_{k+1} B_{j+1} B_{n-k-j}) z^{n+2} \\ &= |\mu|^2 G(z) + 2|\mu| (G^2(z) - |\mu| G(z)) + \frac{1}{M} \sum_{n=0}^{\infty} B_{n+2} z^{n+2} \\ &= |\mu|^2 G(z) + 2|\mu| (G^2(z) - |\mu| G(z)) + \frac{1}{M} (G(z) - |\mu| - |\eta| z) \\ &= 2|\mu| G^2(z) + (\frac{1}{M} - |\mu|^2) G(z) - \frac{1}{M} (|\mu| + |\eta| z), \end{aligned}$$

that is,

$$G^3(z) - 2|\mu| G^2(z) - (\frac{1}{M} - |\mu|^2) G(z) + \frac{1}{M} (|\mu| + |\eta| z) = 0. \quad (12)$$

$$\text{Let } R(z, w, \mu, \eta, M) := R(z, w) = w^3 - 2|\mu| w^2 - (\frac{1}{M} - |\mu|^2) w + \frac{1}{M} (|\mu| + |\eta| z) \quad (13)$$

for (z, w) from a neighbourhood of $(0, |\mu|)$. Since $R(0, |\mu|)$ and $R'_w(0, |\mu|) = -\frac{1}{M} \neq 0$, there exists a unique function $w(z)$, analytic in a neighbourhood of zero, such that $w(0) = |\mu|$, $w'(0) = |\eta|$ and $R(z, w(z)) = 0$. By (10) and (11), we have $G(z) = w(z)$. It follows that the power series (11) converges in a neighbourhood of the origin, and there is a positive constant T such that

$$B_n < T^n, n = 1, 2, \dots$$

Now, we can deduce, by induction, that $b_n \leq B_n e^{k(n-1)}$ for $n \geq 1$, where $K: N \rightarrow R$ is defined in Lemma 1. In fact $|b_1| = |\eta| = c_1$, for inductive proof, we assume that $|b_j| \leq B_j e^{k(j-1)}$, $j \leq n$. From (9) and Lemma 1

$$\begin{aligned} |b_{m+1}| &\leq \frac{M}{|\alpha^m - 1|} \sum_{k=0}^{m-1} \sum_{j=0}^{m-k-1} |b_{k+1}| \cdot |b_{j+1}| \cdot |b_{m-k-j}| \\ &\leq \frac{M}{|\alpha^m - 1|} \sum_{k=0}^{m-1} \sum_{j=0}^{m-k-1} B_{k+1} B_{j+1} B_{m-k-j} e^{k(k+j+k(m-k-j-1))}. \end{aligned}$$

Note that

$$k(k) + k(j) + k(m-k-j-1) \leq k(m-1) \leq \log |\lambda^m - 1| + k(m),$$

then $|b_{m+1}| \leq e^{k(m)} \cdot M \sum_{k=0}^{m-1} \sum_{j=0}^{m-k-1} B_{k+1} B_{j+1} B_{m-k-j} = B_{m+1} e^{k(m)}$. Moreover, from Lemma 1, we know that

$k(n) \leq n(B(\theta) + \gamma)$ for some universal constant $\gamma > 0$. Then $|b_n| \leq T^n e^{(n-1)(B(\theta) + \gamma)}$, that is,

$$\limsup_{n \rightarrow \infty} (|b_n|)^{\frac{1}{n}} \leq \limsup_{n \rightarrow \infty} (T^n e^{(n-1)(B(\theta) + \gamma)})^{\frac{1}{n}} = T e^{B(\theta) + \gamma}.$$

This implies that the convergence radius of (6) is at least $(Te^{B(\theta)+\gamma})^{-1}$. This completes the proof.

In case (H2), the constant α is not only on the unite circle in \mathbf{C} , but also a root of unity. In such a case, both the Diophantine condition and the Brjuno condition are not satisfied. The idea of our proof is acquired from [14].

Theorem 2. Suppose that (H2) holds, then for $b_0 = \mu$ and $\eta \neq 0$ the system $b_l = \eta$ and

$$(\alpha^{n+1} - 1)(n + 2)b_{n+2} = \sum_{k=0}^n \sum_{j=0}^{n-k} \frac{(k+1)(j+1)(\alpha^{2(n-k)-j} - \alpha\alpha^{n-k})}{n-k+1} b_{k+1} b_{j+1} b_{n-k-j},$$

$$n = 0, 1, 2, \dots \tag{14}$$

has a solution $\{b_i\}_{i=1}^\infty$ such that $b_{lp+1} = 0$ and

$$\sum_{k=0}^{lp-1} \sum_{j=0}^{lp-k-1} \frac{(k+1)(j+1)(\alpha^{2(lp-k-1)-j} - \alpha\alpha^{n-k-1})}{lp-k} b_{k+1} b_{j+1} b_{lp-k-j-1} = 0,$$

$l = 1, 2, \dots$.

Then the initial value problem (4) and (5) has an analytic solution of the form

$$y(z) = \mu + \eta z + \sum_{\substack{n \neq lp+1, \\ l \in N}} b_n z^n, \quad N = \{1, 2, 3, \dots\}$$

in a neighbourhood of the origin.

Proof. If $\{b_i\}$ is a solution of system (14) such that $b_{lp+1} = 0$, then $y(z) = \sum_{n=0}^\infty b_n z^n$ is the formal solution of the auxiliary equation (4). Now we need to prove that the power series $y(z) = \sum_{n=0}^\infty b_n z^n$ is convergent. Let

$$\Gamma = \max \left\{ \frac{1}{|\alpha - 1|}, \frac{1}{|\alpha^2 - 1|}, \dots, \frac{1}{|\alpha^{(p-1)} - 1|} \right\} > 0,$$

from (8) we have

$$|b_{n+2}| \leq M\Gamma \sum_{k=0}^n \sum_{j=0}^{n-k} |b_{k+1}| \cdot |b_{j+1}| \cdot |b_{n-k-j}|, \quad n = 0, 1, 2, \dots$$

for all $n \neq lp - 1$, where $l = 1, 2, \dots$. In order to construct a majorant series, we consider the function

$$R(z, \varphi, \mu, \eta, M\Gamma) = 0, \tag{15}$$

where R is defined in (13). Moreover, similarly to the proof of Theorem 1, we can prove that (15) has a unique analytic solution

$$\varphi(z) = \sum_{n=0}^\infty D_n z^n \tag{16}$$

in a neighbourhood of the origin such that

$$D_0 = \varphi(0) = \mu, \quad D_1 = \varphi'(0) = \eta.$$

Substituting (16) into (15) we have

$$D_{n+2} = M\Gamma \sum_{k=0}^n \sum_{j=0}^{n-k} D_{k+1} D_{j+1} D_{n-k-j}, \quad n = 0, 1, 2 \dots$$

for $n \neq lp - 1$, where $l = 1, 2 \dots$. It is easy to show that, by induction,

$$|b_n| \leq D_n, \quad n = 1, 2 \dots$$

Therefore, the series (6) converges in a neighbourhood of the origin. This completes the proof.

3 Analytic Solution of (1)

Theorem 3. Suppose the condition of Theorem 1 and Theorem 2 are satisfied, then Eq. (3) has an analytic solution $y(z)$ of the form $y(z) = g(\alpha g^{-1}(z))$ in a neighborhood of the number μ , such that $y(\mu) = \mu$, $y'(\mu) = \alpha$, where $g(z)$ is an analytic solution of (4).

Proof. As in [8], in view of Theorem 1 and Theorem 2, we may find a sequence $\{b_n\}_{n=0}^{\infty}$ such that the function $g(z)$ of the form (6) is an analytic solution of (4) in a neighborhood of the origin. Since $g'(0) = \eta \neq 0$, the function $g^{-1}(z)$ is analytic in a neighborhood of $g(0) = \mu$. Let $y(z) = g(\alpha g^{-1}(z))$, then from (4) we have

$$\begin{aligned} y'(z) &= \alpha g'(\alpha g^{-1}(z))(g^{-1}(z))' = \frac{\alpha g'(\alpha g^{-1}(z))}{g'(g^{-1}(z))}, \\ y''(z) &= \frac{\alpha^2 g''(\alpha g^{-1}(z)) - \alpha g'(\alpha g^{-1}(z))g''(g^{-1}(z)) \cdot \frac{1}{g'(g^{-1}(z))}}{(g'(g^{-1}(z)))^2} \\ &= \frac{\alpha^2 g''(\alpha g^{-1}(z)) \cdot g'(g^{-1}(z)) - \alpha g'(\alpha g^{-1}(z))g''(g^{-1}(z))}{[g'(g^{-1}(z))]^3} \\ &= \frac{\alpha (g'(g^{-1}(z)))^2 g'(\alpha g^{-1}(z)) [g(\alpha^2 g^{-1}(z)) - \alpha g(\alpha g^{-1}(z))]}{[g'(g^{-1}(z))]^3} \\ &= \frac{\alpha g'(\alpha g^{-1}(z)) [g(\alpha^2 g^{-1}(z)) - \alpha g(\alpha g^{-1}(z))]}{g' [g^{-1}(z)]^3} \\ [y(y(z)) - ay(z)]y'(z) &= [g(\alpha^2 g^{-1}(z)) - \alpha g(\alpha g^{-1}(z))] \frac{\alpha g'(\alpha g^{-1}(z))}{g'(g^{-1}(z))} \\ &= \frac{\alpha g'(\alpha g^{-1}(z)) [g(\alpha^2 g^{-1}(z)) - \alpha g(\alpha g^{-1}(z))]}{g'(g^{-1}(z))}, \end{aligned}$$

so Eq.(3) has an analytic solution $g(z)$ of the form $g(z) = y(\alpha y^{-1}(z))$ in a neighborhood of the number μ such that

$$y(\mu) = g(\alpha g^{-1}(\mu)) = g(\alpha \cdot 0) = g(0) = \mu,$$

$$y'(\mu) = \frac{\alpha g'(\alpha g^{-1}(\mu))}{g'(g^{-1}(\mu))} = \frac{\alpha g'(0)}{g'(0)} = \alpha,$$

where $g(z)$ is an analytic solution of (4). The proof is complete.

Since the function $g(z)$ in (6) can be determined by (8), it can calculate the explicit form of $y(z)$, then an analytic solution of (1) can be obtained in a neighborhood of the fixed point μ of $y(z)$ by means of

$$x(z) = x(z_0) + \frac{1}{b} \int_{z_0}^z (y(s) - as) ds .$$

References

1. Eder, E.: The functional differential equation $x'(t)=x(x(t))$. J. Differential Equations 54, 390–400 (1984)
2. Fečkan, E.: On certain type of functional differential equations. Math. Slovaca 43, 39–43 (1993)
3. Ke, W.: On the equation $x'(t)=f(x(x(t)))$. Funkcialaj Ekvacioj 33, 405–425 (1990)
4. Si, J.G., Cheng, S.S.: Analytic solutions of a functional differential equations with state dependent argument. Taiwanese J. Math. 1(4), 471–480 (1997)
5. Si, J.G., Cheng, S.S.: Note on an iterative functional differential equations. Demonstratio Math. 31(3), 609–614 (1998)
6. Si, J.G., Li, W.R., Cheng, S.S.: Analytic solutions of an iterative functional differential equation. Computer Math. Applic. 33(6), 47–51 (1997)
7. Liu, L.X.: Local analytic solutions of a functional equation $az+\beta x'(z)=x(z+bz^n(z))$. Appl. Math. Compu. 215, 644–652 (2009)
8. Si, J.G., Wang, X.P.: Analytic solutions of a second-order functional differential equation with a state derivative dependent delay. Colloquium Math. 79(2), 273–281 (1999)
9. Si, J., Liu, T.: Local analytic solutions of a functional differential equation with a deviating argument depending on the state derivative near resonance. Comput. Math. Appl. 54, 750–762 (2007)
10. Liu, L.X.: The Exist Local Analytic Solutions of an Iterative Functional Differential Equation. In: Huang, D.-S., Zhang, X., Reyes García, C.A., Zhang, L. (eds.) ICIC 2010. LNCS (LNAI), vol. 6216, pp. 17–26. Springer, Heidelberg (2010)
11. Bjuno, A.D.: Analytic form of differential equations. Trans. Moscow Math. Soc. 25, 131–288 (1971)
12. Marmi, S., Moussa, P., Yoccoz, J.-C.: The Brjuno functions and their regularity properties. Comm. Math. Phys. 186(2), 265–293 (1997)
13. Carletti, T., Marmi, S.: Linearization of Analytic and Non-Analytic Germs of Diffeomorphisms of $(\mathbb{C}, 0)$. Bull. Soc. Math. 128, 69–85 (2000)
14. Davie, A.M.: The critical function for the semistandard map. Nonlinearity 7, 219–229 (1994)

Analytic Solutions of an Iterative Functional Differential Equation with State Dependent Delay

Lingxia Liu

Department of Mathematics, Weifang University,
Weifang, Shandong 261061, P. R. China
llxmath@126.com

Abstract. This paper is concerned with a functional differential equation with state dependent delay. By constructing a convergent power series solution of an auxiliary equation, analytic solutions for the original differential equation are obtain. For technical reasons, in previous work the constant α given in the Schröder transformation, is required to fulfill that α is off the unit circle or lies on the circle with the Diophantine condition. In this paper, we obtain results of analytic solutions in the case of α at resonance, i.e., at a root of the unity and the case of α near resonance under the Brjuno condition.

Keywords: Iterative functional equation, analytic solution, state dependent delay, Brjuno condition, resonance.

1 Introduction

Functional differential equations with stat dependent delay have attracted have attracted the attentions of many authors in the last few year [1-6]. In [5] and [6], the existence of analytic solution of the equations $x'(z) = x^m(z)$ and

$$\alpha z + \beta x'(z) = x(\alpha z + \beta x'(z)) \quad (1)$$

is established. In Eq.(1), the case $\alpha = a = 0, \beta = 1$ and $b = 1; b = 0, a \neq 0$ and $\beta \neq 0$ and in case $b = 0, a \neq 0$ and $\beta = 0$, the authors have given the analytic solutions in [6]. When $b \neq 0$, in order to construct analytic solutions of (1) in a systematic manner, we let

$$y(z) = \alpha z + \beta x'(z). \quad (2)$$

Then for any number z_0 , we have

$$x(z) = x(z_0) + \frac{1}{b} \int_{z_0}^z (y(s) - \alpha s) ds, \quad (3)$$

and $x(y(z)) = x(z_0) + \frac{1}{b} \int_{z_0}^{y(z)} (y(s) - \alpha s) ds$. Therefore, in view of (1) we have

$$x(\alpha z + \beta x'(z)) = x(y(z)) = x(z_0) + \frac{1}{b} \int_{z_0}^{y(z)} (y(s) - \alpha s) ds = \alpha z + \frac{\beta}{b} (y(z) - \alpha z), \text{ or}$$

$$bx(z_0) + \int_{z_0}^{y(z)} (y(s) - as)ds = \beta y(z) + (b\alpha - a\beta)z. \tag{4}$$

In case z_0 is a fixed point of $y(z)$, i.e., $y(z_0) = z_0$, we see that

$$bx(z_0) + \int_{z_0}^{y(z_0)} (y(s) - as)ds = \beta y(z_0) + (b\alpha - a\beta)z_0,$$

or

$$x(z_0) = \frac{1}{b}[b\alpha + (1 - a)\beta]z_0. \tag{5}$$

Furthermore, differentiating both sides of the Eq. (4) with respect to z , we obtain an iterative functional differential equation

$$[y(y(z)) - ay(z) - \beta]y'(z) = b\alpha - a\beta. \tag{6}$$

In [6], when $b\alpha - a\beta = 0$ and $b \neq 0$, an entire solution of (1) is obtain. When $b\alpha - a\beta \neq 0$, we reduce the equation (6) with the

Schröder transformation $y(z) = g(\mu g^{-1}(z))$ to the auxiliary equation

$$\mu g'(\mu z)[g(\mu^2 z) - ag(\mu z) - \beta] = g'(z)(b\alpha - a\beta). \tag{7}$$

We assume that Eq.(7) satisfies the condition

$$g(0) = s. \tag{8}$$

For technical reason, in [6], the constant μ given in the *Schröder* transformation, is required to fulfill that μ is off the unit circle ($0 < \mu < 1$) or μ is on the circle with the Diophantine condition: $|\mu| = 1$, μ is not a root of unity, and $\log \frac{1}{|\mu^n - 1|} \leq T \log n$, $n = 2, 3, \dots$ for some positive constant T . Roughly speaking, the Diophantine condition requires μ to be far from all roots of unity.

In this paper, we discuss not only those α at resonance, i.e., at a root of unity considered in [6], but also those α near resonance under the Brjuno condition consider in [7, 8]. Now we discuss the analytic solution of Eq. (7) in the following conditions:

(H₁) $\mu = e^{2\pi i\theta}$, where $\theta \in \mathbf{R} \setminus \mathbf{Q}$ is a Brjuno number ([9, 10]), i.e., $B(\theta) =$

$\sum_{k=0}^{\infty} \frac{\log q_{k+1}}{q_k} < \infty$, where $\{p_k / q_k\}$ denotes the sequence of partial fraction of the continued fraction expansion of θ , and is said to satisfy the Brjuno condition.

(H₂) $\alpha = e^{2\pi i q^l p}$ for some integers $p \in \mathbf{N}$ with $p \geq 2$ and $q \in \mathbf{Z} \setminus \{0\}$, and $\alpha \neq e^{2\pi i l / k}$ for all $1 \leq k \leq p-1$ and $l \in \mathbf{Z} \setminus \{0\}$.

Observe that μ is a p-th unit root (or called p-order resonance) in the case of (H₂), while the case (H₁) contains a part of α near resonance.



2 Analytic Solution of Auxiliary Equation

As stated in [11], for a real number θ we let θ denote its integer part and let $\{\theta\} = \theta - [\theta]$. Then every rational number θ has a unique expression of the Gauss' continued fraction $\theta = a_0 + \theta_0 = a_0 + \frac{1}{a_1 + \theta_1} = \dots$, denoted simply by $\theta = [a_0, a_1, \dots, a_n, \dots]$, where a_j 's and θ_j 's are calculated by the algorithm: **(a)** $a_0 = [\theta]$, $\theta_0 = \{\theta\}$ and **(b)** $a_n = \left[\frac{1}{\theta_{n-1}} \right]$, $\theta_n = \left\{ \frac{1}{\theta_{n-1}} \right\}$ for all $n \geq 1$. Define the sequences $(p_n)_{n \in \mathbb{N}}$ and $(q_n)_{n \in \mathbb{N}}$ as follows

$$p_{-2} = 1, q_{-1} = 0, q_n = a_n q_{n-1} + q_{n-2}; \quad p_{-2} = 0, p_{-1} = 1, p_n = a_n p_{n-1} + p_{n-2}.$$

It is easy to show that $p_n / q_n = [a_0, a_1, \dots, a_n]$. Thus, for every $\theta \in \mathbb{R} \setminus \mathbb{Q}$ we associate, using its convergence, an arithmetical function $B(\theta) = \sum_{n \geq 0} \frac{\log q_{n+1}}{q_n}$. We say that θ is a Brjuno number or that it satisfies Brjuno condition if $B(\theta) < +\infty$. The Brjuno condition is weaker than the Diophantine condition. For example, if $a_{n+1} \leq c e^{a_n}$ for all $n \geq 0$, where $c > 0$ is a constant, then $\theta = [a_0, a_1, \dots, a_n, \dots]$ is a Brjuno number but is not a Diophantine number. So, the case (H1) contains both Diophantine condition and a part of μ "near" resonance.

In order to discuss the existence of analytic solutions of the auxiliary equation (6) under (H1), we need to introduce Davie's Lemma. First, we recall some facts in [12] briefly. Let $\theta \in \mathbb{R} \setminus \mathbb{Q}$ and $(q_n)_{n \in \mathbb{N}}$ be the sequence of partial denominators of the Gauss's continued fraction for θ as in the Introduction. As in [11], let

$$A_k = \{n \geq 0 \mid \|n\theta\| \leq \frac{1}{8q_k}\}, \quad E_k = \max(q_k, \frac{q_{k+1}}{4}), \quad \eta_k = \frac{q_k}{E_k}.$$

Let A_k^* be the set of integers $j \geq 0$ such that either $j \in A_k$ or for some j_1 and j_2 in A_k , with $j_2 - j_1 < E_k$, one has $j_1 < j < j_2$ and q_k divide $j - j_1$. For any integer $n \geq 0$, define

$$l_k(n) = \max\left((1 + \eta_k) \frac{n}{q_k} - 2, (m_n \eta_k + n) \frac{1}{q_k} - 1 \right),$$

Where $m_n = \max\{j \mid 0 \leq j \leq n, j \in A_k^*\}$. We then define function $h_k : \mathbb{N} \rightarrow \mathbb{R}_+$ as follows:

$$\begin{cases} \frac{m_n + \eta_k n}{q_k} - 1, & \text{if } m_n + q_k \in A_k^*, \\ l_k(n), & \text{if } m_n + q_k \notin A_k^*. \end{cases} \quad \text{Let } g_k(n) := \max\left(h_k(n), \left\lfloor \frac{n}{q_k} \right\rfloor \right), \quad \text{and}$$

define $k(n)$ by the condition $q_{k(n)} \leq n \leq q_{k(n)+1}$. Clearly, $k(n)$ is non-decreasing. Then we are able to state the following result:



Lemma 1. (Davie’s Lemma [12]) Let

$$K(n) = n \log 2 + \sum_{k=0}^{k(n)} g_k(n) \log(2q_{k+1}).$$

Then

(a) There is a universal constant $\gamma > 0$ (independent of n and θ) such that

$$K(n) \leq n \left(\sum_{k=0}^{k(n)} \frac{\log q_{k+1}}{q_k} + \gamma \right),$$

(b) $K(n_1) + K(n_2) \leq K(n_1 + n_2)$ for all n_1 and n_2 , and

(c) $-\log |\alpha^n - 1| \leq K(n) - K(n-1)$.

Now we state and prove the following theorem under Brjuno condition. The idea of our proof is acquired from [11].

Theorem 1. Suppose (H_1) holds and $b\alpha - a\beta \neq 0$. Suppose further that when $\alpha = 1$, we have $\beta \neq 0$ and $\mu = (a\beta - b\alpha)/\beta$. Then Eq. (7) has an analytic solution of the form

$$g(z) = s + \eta z + \sum_{n=2}^{\infty} c^n z_n, \tag{9}$$

where s is arbitrary when $a = 1$, otherwise $s = \frac{\beta\mu + b\alpha - a\beta}{(1-a)\mu}$.

Proof. As in [6], we seek a solution of (7) in a power series of the form (9). Substituting the subsequent power series into (7) and comparing coefficients, we obtain

$$[\beta\mu + b\alpha - a\beta - (1-a)\mu s]c_1 = 0, \tag{10}$$

and

$$(a\beta - b\alpha)(\mu^n - 1)(n+1)c_{n+1} = \sum_{k=0}^{n-1} (k+1)(\mu^{2n-k+1} - a\mu^{n+1})c_{k+1}c_{n-k},$$

$$n = 1, 2, \dots. \tag{11}$$

In view of the definition of s , we have that $\beta\mu + b\alpha - a\beta - (1-a)\mu s = 0$, so that we can choose $c_1 = \eta \neq 0$. Once c_0 and c_1 are determined, the other terms of the sequence $\{c_n\}$ can be determined from (11) in a unique manner.

We need to show that the power series (9) converges in a neighborhood of the origin. First of all, note that

$$\left| \frac{(k+1)(\mu^{2n-k+1} - a\mu^{n+1})}{(a\beta - b\alpha)(n+1)} \right| \leq \frac{1+|a|}{|a\beta - b\alpha|} = M.$$

Then (11) can be changed into

$$|c_{n+1}| \leq \frac{M}{|\mu^n - 1|} \sum_{k=0}^{n-1} |c_{k+1}| |c_{n-k}|, \quad n \in \mathbb{Z}^+. \tag{12}$$

We define a power series

$$U(z) = \sum_{n=1}^{\infty} B_n z^n, \tag{13}$$



by $B_1 = |\eta|$, and

$$B_{n+1} = M \sum_{k=0}^{n-1} B_{k+1} B_{n-k}, \quad n = 1, 2, \dots, \tag{14}$$

then

$$\begin{aligned} U^2(z) &= \sum_{n=2}^{\infty} (B_1 B_{n-1} + B_2 B_{n-2} + \dots + B_{n-1} B_1) z^n \\ &= \sum_{n=1}^{\infty} (B_1 B_n + B_2 B_{n-1} + \dots + B_n B_1) z^{n+1} \\ &= \frac{1}{M} \sum_{n=1}^{\infty} B_{n+1} z^{n+1} = \frac{1}{M} (U(z) - |\eta|z). \end{aligned}$$

Define the function

$$G(M, \eta; z, U) = G(z, U) = U^2 - \frac{1}{M}U + \frac{|\eta|}{M}z = 0 \tag{15}$$

for (z, U) in a neighborhood of $(0, 0)$. Then $G(0, 0) = 0$, $G'_z(0, 0) = -\frac{1}{M} \neq 0$. Thus, there exists a unique function $U(z)$, analytic in a neighborhood of zero, such that $U(0) = 0$, $U'_z(0) = -\frac{G'_z(0, 0)}{G'_U(0, 0)} = |\eta| \neq 0$, and $G(z, U(z)) = 0$, so $U(z)$ can be expanded into a convergent power series (13). Hence there is a constant $T > 0$ such that $B_n \leq T^n$, $n = 1, 2, \dots$. Now we prove

$$|C_n| \leq B_n e^{K(n-1)}, \quad n = 1, 2, \dots, \tag{16}$$

where $K: N \rightarrow R$ is defined in Lemma 1. In fact, $|C_1| = |\eta| = B_1$, we assume that $|C_j| \leq B_j e^{K(j-1)}$, $j \leq n$. From Lemma 1 and (12) we have

$$\begin{aligned} |c_{n+1}| &\leq \frac{M}{|\mu^n - 1|} \sum_{k=0}^{n-1} |c_{k+1}| |c_{n-k}| \\ &\leq \frac{M}{|\mu^n - 1|} \sum_{k=0}^{n-1} B_{k+1} e^{K(k)} \cdot B_{n-k} e^{K(n-k-1)} \\ &\leq \frac{M}{|\mu^n - 1|} e^{K(n-1)} \sum_{k=0}^{n-1} B_{k+1} B_{n-k} = \frac{e^{K(n-1)}}{|\mu^n - 1|} B_{n+1}. \end{aligned}$$

Note that

$$K(n-1) \leq K(n) + \log |\mu^n - 1|,$$

then

$$|c_{n+1}| \leq B_{n+1} e^{K(n)}.$$

From (16), we have $|c_n| \leq B_n e^{K(n-1)} \leq T^n e^{K(n-1)}$. Note that $k(n) \leq n(B(\theta) + \gamma)$ for some universal constant $\gamma > 0$, then

$$|c_n| \leq T^n e^{(n-1)(B(\theta) + \gamma)},$$

that is,

$$\limsup_{n \rightarrow \infty} (|c_n|)^{\frac{1}{n}} = \limsup_{n \rightarrow \infty} (Te^{\frac{n-1}{n}(B(\theta)+\gamma)}) = Te^{B(\theta)+\gamma}.$$

This implies that the convergence radius of the series (9) is at least $(Te^{B(\theta)+\gamma})^{-1}$. This completes the proof.

In case (H_2) the constant μ is not only on the unit circle in C , but also a root of unity, i.e., the resonant of the order p . In such a case Diophantine condition or Brjuno condition is not satisfied.

Let $\{A_n\}_{n=1}^\infty$ be a sequence define by $A_1 = |\eta|$ and

$$A_{n+1} = M\Gamma \sum_{k=0}^{n-1} A_{k+1} A_{n-k}, \quad n = 1, 2, \dots, \tag{17}$$

where $\Gamma = \max \left\{ 1, \frac{1}{|\mu-1|}, \frac{1}{|\mu^2-1|}, \dots, \frac{1}{|\mu^{p-1}-1|} \right\}$, M is defined in Theorem 1.

Theorem 2. Suppose (H_2) holds and that $b\alpha - a\beta \neq 0$. Suppose further that when $\alpha = 1$, we have $\beta \neq 0$ and $\mu = (a\beta - b\alpha)/\beta$. p is given as above. Let $\{c_n\}_{n=1}^\infty$ be determined recursively by $c_0 = s$, $c_1 = \eta$ and

$$(a\beta - b\alpha)(\mu^n - 1)(n+1)c_{n+1} = \Theta(n, \mu), \quad n = 1, 2, \dots, \tag{18}$$

where

$$\Theta(n, \mu) = \sum_{k=0}^{n-1} (k+1)(\mu^{2n-k+1} - a\mu^{n+1})c_{k+1}c_{n-k}, \quad n = 1, 2, \dots.$$

If $\Theta(vp, \mu) = 0$ for all $v = 1, 2, \dots$, then Eq. (7) has an analytic solution $g(z)$ in a neighborhood of the origin such that $g(0) = s$, $g'(0) = 1$ and $g^{vp+1}(0) = (vp+1)! \tau_{vp+1}$, where all $\tau_{vp+1}'s$ is arbitrary constants satisfying the inequality $|\tau_{vp+1}| \leq A_{vp+1}$ and the sequence $\{A_n\}_{n=1}^\infty$ is defined in (17). Otherwise, if $\Theta(vp, \mu) \neq 0$ for some $v = 1, 2, \dots$, then Eq. (7) has no analytic solutions in any neighborhood of the origin.

Proof. As in the proof of Theorem 1, we seek a power series solution of the form (9). Then defining $C_0 = s$ and $C_1 = \eta$, (11) or (18) holds again.

If $\Theta(vp, \mu) \neq 0$ for some natural number v , then the equality in (11) or (18) does not hold for $n = vp$. This is because $\mu^{vp} - 1 = 0$, then such a circumstance equation (7) has no formal solutions. When $\Theta(vp, \mu) = 0$ for all natural number v , for each v the corresponding b_{vp+1} in (11) or (18) has infinitely many choices in C , that is, the formal series solutions (9) define a family of solutions with infinitely many parameters. Choose $c_{vp+1} = \tau_{vp+1}$ arbitrary such that $|c_{vp+1}| \leq A_{vp+1}$, $v = 1, 2, \dots$, where A_{vp+1} is defined



by (17). In what follows we prove the power series solution (9) converges in a neighborhood of the origin. For $n \neq \nu p$, we have $|\mu^n - 1| \leq \Gamma$, then

$$|c_{n+1}| \leq M\Gamma \sum_{k=0}^{n-1} |c_{k+1}| |c_{n-k}|, \quad n \neq \nu p, \quad \nu = 1, 2, \dots \quad (19)$$

We consider the implicit functional equation

$$G(M\Gamma, \eta; z, \omega) = 0, \quad (20)$$

where G is defined in (15). Similar to the proof of Theorem 1, we can prove that there exists a unique function $\omega(M\Gamma, \eta; z)$, analytic in a neighborhood of the zero such that $\omega(M\Gamma, \eta; 0) = 0$, $\omega'_z(M\Gamma, \eta; 0) = |\eta|$. So $\omega(M\Gamma, \eta; z)$ can be expanded into a convergent power series

$$\omega(M\Gamma, \eta; z) = \sum_{n=1}^{\infty} A_n z^n, \quad A_1 = |\eta|. \quad (21)$$

Moreover, by induction we can easily prove

$$|C_n| \leq A_n, \quad n = 1, 2, \dots$$

So the series (9) converges in a neighborhood of the origin.

3 Existence of Analytic Solution

Having results of analytic solutions for the auxiliary equation (7), we can give the main results to the original equation (1).

Theorem 3. Suppose one of the conditions in Theorem 1-2 is fulfilled. Then Eq. (6) has a solution of the form $y(z) = g(ug^{-1}(z))$ in a neighborhood of the number s , such that $y(s) = s, y'(s) = \mu$, where s is defined in (8), and g is an analytic solution of the auxiliary equation (7).

Proof. By Theorem 1-2, we can find an invertible analytic solution g of the auxiliary equation (7) in the form of (9) such that $g(0) = s$ and $g'(s) = \eta \neq 0$. Let

$$y(z) = g(ug^{-1}(z)),$$

which is analytic in a neighborhood of the origin. From (7) it is easy to see

$$\begin{aligned} & [y(y(z)) - ay(z) - \beta] y'(z) \\ &= [g(u^2 g^{-1}(z)) - ag(ug^{-1}(z)) - \beta] \frac{\mu g'(ug^{-1}(z))}{g'(g^{-1}(z))} \\ &= b\alpha - a\beta, \end{aligned}$$

and

$$\begin{aligned} y(s) &= g(ug^{-1}(s)) = g(0) = s, \\ y'(s) &= \frac{\mu g'(ug^{-1}(s))}{g'(g^{-1}(s))} = \frac{\mu g'(0)}{g'(0)} = \mu. \end{aligned}$$

This shows that the composite function $y(z)$ is a solution of (6) as desired and that s is a fixed point of $y(z)$. The proof is complete.

Since the function $g(z)$ in (9) can be determined by (11), it is possible to calculate the explicit form of $y(z)$. Under the additional condition that $b \neq 0$ and $b\alpha - a\beta \neq 0$ an explicit analytic solution of (1) exists in a neighborhood of the fixed point s of $y(z)$ by means of (3) and (5), so analytic solution of (1) exists.

References

1. Feckan, E.: On certain type of functional differential equations. *Math. Slovaca* 43, 39–43 (1993)
2. Ke, W.: On the equation $x'(t)=f(x(x(t)))$. *Funkcialaj Ekvacioj* 33, 405–425 (1990)
3. Si, J.G., Cheng, S.S.: Analytic solutions of a functional differential equations with state dependent argument. *Taiwanese J. Math.* 1(4), 471–480 (1997)
4. Si, J.G., Cheng, S.S.: Note on an iterative functional differential equations. *Demonstratio Math.* 31(3), 609–614 (1998)
5. Si, J.G., Li, W.R., Cheng, S.S.: Analytic solutions of an iterative functional differential equation. *Computer Math. Applic.* 33(6), 47–51 (1997)
6. Si, J.G., Li, W.R., Cheng, S.S.: Analytic solutions of a functional differential equations with a state derivative dependent delay. *Aequationes Math.* 57, 75–86 (1999)
7. Liu, L.X.: Local analytic solutions of a functional equation $az+\beta x'(z)=x(az+bx''(z))$. *Appl. Math. Compu.* 215, 644–652 (2009)
8. Si, J., Liu, T.: Local analytic solutions of a functional differential equation with a deviating argument depending on the state derivative near resonance. *Comput. Math. Appl.* 54, 750–762 (2007)
9. Bjuno, A.D.: Analytic form of differential equations. *Trans. Moscow Math. Soc.* 25, 131–288 (1971)
10. Marmi, S., Moussa, P., Yoccoz, J.-C.: The Brjuno functions and their regularity properties. *Comm. Math. Phys.* 186(2), 265–293 (1997)
11. Carletti, T., Marmi, S.: Linearization of Analytic and Non-Analytic Germs of Diffeomorphisms of $(\mathbb{C}, 0)$. *Bull. Soc. Math.* 128, 69–85 (2000)
12. Davie, A.M.: The critical function for the semistandard map. *Nonlinearity* 7, 219–229 (1994)

The Study of Optimal Conditions of Electroporation in *Gluconacetobacter Xylinum**

Huixia Zhu^{1,2}, Shiru Jia^{1,**}, Hongjiang Yang¹, Lin Yan¹, and Jing Li¹

¹ Key Laboratory of Industrial Microbiology, Ministry of Education,
Tianjin University of Science and Technology, Tianjin, China
jiashiru@mail.tust.cn

² Department of Life Science, Heng Shui University, Hengshui, Hebei, China
zhuhuixia@126.com

Abstract. The effect of electroporation conditions of *Gluconacetobacter xylinum* CGMCC 2955 was systematically investigated in this paper. The enzymolysis condition for removing bacterial cellulose of *Gluconacetobacter xylinum* was optimized by Response Surface Method, the result showed as follows: cellulase concentration 1046.4 IU/mL, enzymolysis duration 93.8 min at 30.6° C. Optimum conditions including growth stage of the strain, electroshock voltage, concentration and preservation of competent cell were defined for the electroporation of CGMCC 2955 with plasmid pDN19Lac Ω . The result was showed that the highest transformation efficiency was up to 8.76×10^5 CFU/ μ g DNA under the optimum conditions in which the competent cells were collected at 2d and concentration to about 1×10^8 CFU/mL, the mixture of the competent cells and plasmid pDN19Lac Ω was eletroporated at 2.4kV. The transformation efficiency is significantly reduced on account of -70° C preservation of competent cell.

Keywords: *Gluconacetobacter xylinum*, electroporation, pDN19Lac Ω , Response Surface Method, bacterial cellulose.

Gluconacetobacter xylinum was first discovered the bacteria biosynthesized bacterial cellulose (BC) and used extensively up to now [1,2]. BC was widely applied in fields of material and medical, chemical, food and bio-engineering because of high production and purity. So the genetic operations of *G. xylinum* were studied in a deep-going way and the wild strains were constructed by gene engineering. All of these were set up to obtain the new strains of higher-yield or better BC [3].

The electroporation technique has the advantages of simple, rapid and high transformation frequency and was applied from the end of 1980s. The transforming efficiency of electroporation is increased 10 ~ 100 times as against chemical

* This work was funded by the National Natural Science Foundation of China under No.20976133 and the Foundation for Excellent Doctoral Dissertations of TianJin University of Science and Technology under No. B201002.

** Corresponding author.

transforming, so it was often used in test with charging high transforming efficiency [4,5,6,7].

The electroporation techniques of *G. xylinum* have not made a systematical and profound study for the present. In addition, plenty of BC biosynthesized in process of culture lead to competent problems of *G. xylinum*. In this study, the enzymolysis condition for removing bacterial cellulose of *G. xylinum* was optimized by Response Surface Method; optimum transmission conditions including growth stage of the strain, electroshock voltage, concentration and preservation of competent cell were defined for the electroporation with plasmid pDN19Lac Ω . All of these pave the way to introducing external-DNA into *G. xylinum* CGMCC No.2955.

1 Materials and Methods

1.1 Strain and Plasmid

Gluconacetobacter xylinum CGMCC No.2955 (*G. xylinum*) was isolated in the Key Laboratory of Industrial Microbiology, Tianjin University of Science and Technology, and stored in China General Microbiological Culture Collection Center. The plasmid pDN19Lac Ω (containing spectinomycin gene and streptomycin gene) was provided by Professor Hongjiang Yang at Tianjin University of Science and Technology.

1.2 Culture Medium and Reagents

The culture medium for BC agitation fermentation consisted of carbon source 25 g/l (Yingbo Biochemical Reagent Co. Ltd., China), peptone 10 g/l (Sigma, America), yeast extract 7.5 g/l (Sigma, America) and disodium phosphate 10 g/l (Sigma, America) at the initial pH value 6.0 [8, 9].

According to the above medium, the solid medium was added agar 20 g/l.

Antibiotic was purchased from Beijing Newprobe Co. Ltd. Plasmid DNA purification Kit, restriction enzymes and DNA marker were purchased from Takara Co. Ltd.

1.3 Culture, Enzymolysis and Competent Cell Preparation

For seed culture, *G. xylinum* was inoculated into culture medium (100 ml) in a 500 ml flask on a rotary shaker (Donglian Co. Ltd., China) with 160 rpm of agitation speed and incubated at 30 °C for 24 hours. 200 μ l liquid seed was taken out spread plate and incubated at 30 °C for 2 days. Then small amount of sterile water was poured into plates. The culture on plate was scraped off and added cellulase degradation.

The corresponding optimum enzymolysis condition was studied by Response Surface Analysis of SAS software. The design of analytical experiments was in Table 1.

Table 1. Level and code of experiment variables

| | X ₁ , cellulase concentration (U/ml) | X ₂ , enzymolysis time (min) | Temperature (°C) |
|----|---|---|------------------|
| -1 | 500 | 60 | 25 |
| 0 | 1000 | 90 | 30 |
| 1 | 1500 | 120 | 35 |

Cellulase (Novozymes, Denmark) was added to the culture for enzymolysis at 30 °C for 2 hours. The aim of cellulase hydrolysis was to degrade cellulose in the seed liquid, so as to liberate enwrapped cells; therefore uniform BC spheres were produced in the subsequent fermentation stage. *G. xylinum* cells were centrifugally washed with deionized water for three times to remove residual cellulose. Then the competent cells were prepared as follows: the cells were centrifuged washing two times by 10% glycerin-MgCl₂ solution and washing two times by 10% glycerin solution. Finally, the competent cells were suspended in 10% glycerin solution. Some competent cells were transformed directly, others were preserved it in -70 °C refrigerator.

1.4 Electroporation Method

100 µl competent cells of different concentrations and 50 ng plasmid were taken to 0.2 cm ice-chilled gene pulser cup (Bio-Rad, America), while control sample were not add plasmid. All gene pulser cups were electrocuted (200 Ω, 25 µFd) by electroporation (Bio-Rad, America) apparatus under different voltage. Then LB medium at 37 °C were mixed immediately with electrocuted competent cells and cultured 2 h at 30 °C with a shaking speed of 160 rpm. The optimum dilution culture liquid was taken 100 µl to plates containing Sp 10 µg/ml and Sm10 µg/ml while plates were incubated for 4 d at 37 °C. Then number of transformant was calculated transforming efficiency.

Transforming efficiency (CFU/µg DNA) = transformant number × dilution multiplication × 10/ DNA quantity

1.5 Screening and Identification of Transformant

Several transformants were randomly picked from screening plate and the plasmid DNA abstracted from transformants was determined with agarose gel electrophoresis.

2 Results

2.1 Effect of Enzymatic Hydrolysis Conditions on Transforming Efficiency

We had analyzed the enzymatic hydrolysis conditions with SAS software package. This method was very sensitive (<0.05), and showed high veracity and linearity: the regression equation was

$$y=8.14+0.36125\times X_1+0.15625\times X_2+0.23\times X_3-2.2162\times X_1\times X_1-0.0325\times X_1\times X_2+0.01\times X_1\times X_3-0.90625\times X_2\times X_2+0.04\times X_2\times X_3-1.24375\times X_3\times X_3, R^2=0.9945.$$

Table 2. Response surface analysis

| Source | DF | SS | MS | F | Pr > F |
|--------------------------------|----|----------|----------|----------|----------|
| X ₁ | 1 | 1.044013 | 1.044013 | 36.72867 | 0.001765 |
| X ₂ | 1 | 0.195313 | 0.195313 | 6.871152 | 0.047025 |
| X ₃ | 1 | 0.4232 | 0.4232 | 14.8883 | 0.011898 |
| X ₁ ×X ₁ | 1 | 18.13574 | 18.13574 | 638.0209 | 0.0001 |
| X ₁ ×X ₂ | 1 | 0.004225 | 0.004225 | 0.148637 | 0.715699 |
| X ₁ ×X ₃ | 1 | 0.0004 | 0.0004 | 0.014072 | 0.91019 |
| X ₂ ×X ₂ | 1 | 3.032452 | 3.032452 | 106.6826 | 0.000146 |
| X ₂ ×X ₃ | 1 | 0.0064 | 0.0064 | 0.225154 | 0.655141 |
| X ₃ ×X ₃ | 1 | 5.711683 | 5.711683 | 200.9387 | 0.0001 |
| Model | 9 | 25.70965 | 2.856628 | 100.497 | 0.0001 |
| Linear | 3 | 1.662525 | 0.554175 | 19.49604 | 0.003445 |
| Quadratic | 3 | 24.0361 | 8.012033 | 281.8657 | 0.0001 |
| Cross Product | 3 | 0.011025 | 0.003675 | 0.129288 | 0.938614 |
| Error | 5 | 0.142125 | 0.028425 | | |
| Lack of fit | 3 | 0.092725 | 0.030908 | 1.25135 | 0.473026 |
| Pure Error | 2 | 0.0494 | 0.0247 | | |
| Total | 14 | 25.85177 | | | |

The enzymolysis condition for removing bacterial cellulose of *G. xylinum* was optimized by Response Surface Method, X₁ =0.092795, X₂ =0.127377, X₃ =0.123143. The results showed as follows: cellulase concentration 1046.4 U/ml, enzymolysis duration 93.8 min at 30.6 °C. The highest transformation efficiency was up to 8.17×10⁵ CFU/μg DNA under the optimum conditions.

2.2 Effect of Ambient Temperature on Transforming Efficiency

Two kind of competent cells were prepared differently at 4 °C and 25 °C and else facts under the optimum conditions. The competent cells prepared at 4 °C were poured in ice-chilled gene pulser cup, while the competent cells prepared at 25 °C were poured in 25 °C gene pulser cup.

As shown in Figure 1a, transforming efficiency was temperature dependent which strongly decreased at higher temperatures. So the procedure of obtained cells and electroporation could do reached the highest transformation efficiency at 4 °C.

2.3 Effect of Cell Culture Time on Transforming Efficiency

Figure1b showed that the highest transformation efficiency was obtained when the culture time was 2 d, but the efficiency decreased when exceeded to 2 d. As a result, the following trails' culture was all 2 d.

2.4 Effect of Electrocuted Voltage on Transforming Efficiency

Figure1c showed that the highest transformation efficiency was obtained when the electrocuted voltage was 2.4 kV, but the efficiency decreased when exceeded to 2.4 kV.

2.5 Effect of Cells Concentration on Transforming Efficiency

As shown in Figure 1d, the transformation efficiency was very low when cells concentrations were low than 1×10^8 CFU/ml, while the highest transformation efficiency was obtained when cells concentration was justly 1×10^8 CFU/ml. The cells concentrations remain increased, the transformation efficiency was decreased.

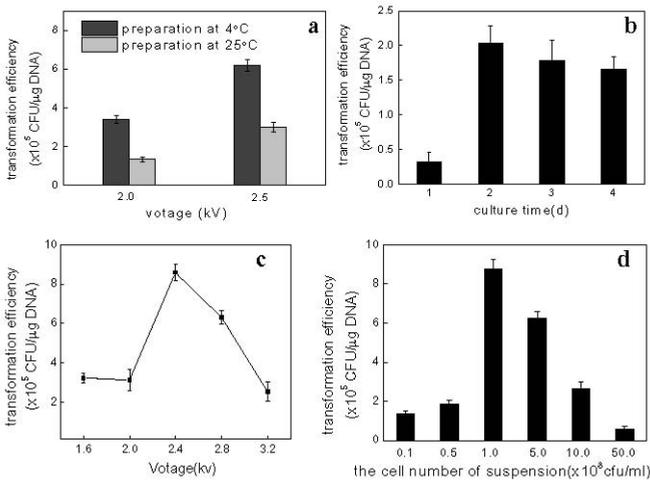


Fig. 1. The effects of temperature, culture time, voltage and cell concentration to electroporation

2.6 Effect of Competent Cells Preserved at -70° C on Transforming Efficiency

The competent cells were prepared on the same conditions. The transforming efficiency preserved at -70° C was less than without preserved by 33.3% (2.4kV) and 81.99% (2.8kV), respectively.

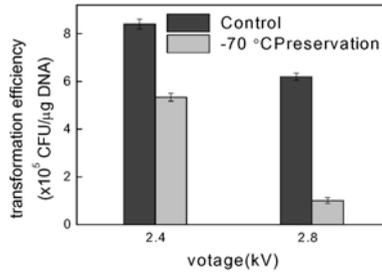


Fig. 2. The effects of -70°C preservation of competent cells to electroporation

2.7 Identification of Transformant

Control plates were not colonies generation. As shown in Figure 3a, transformant was cultured and biosynthesized BC membrane in LB liquid medium contained spectinomycin $10\ \mu\text{g}/\text{ml}$ and streptomycin $10\ \mu\text{g}/\text{ml}$ while wild type strain was not growth.

Plasmid pDN19Lac Ω digested by *Hind* III was divided into two fragments with 6 kb and 2 kb respectively. Several transformants were randomly picked from screening plate and the plasmid DNA was abstracted from transformants. 8 kb fragment was extracted from agarose gel electrophoresis and digested by *Hind* III. Figure 3b proved the following results: Plasmid pDN19Lac Ω was transformed into *Gluconacetobacter xylinum* CGMCC No.2955 and transformants were screened and identified.

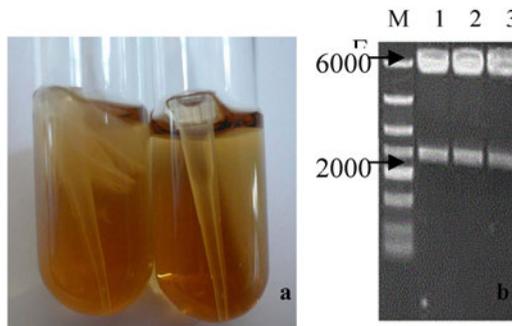


Fig. 3. The electroporation results. (a: the left is the fermented result by positive transformant; the right is control. b: M is DNA Marker; 1: pDN19Lac Ω cut out by *Hind* III; 2 and 3: the plasmid of positive transformants cut out by *Hind* III.)

3 Discussion

This study shows that plentiful bacterial cellulose could be produced by *Gluconacetobacter xylinum* in the process of solution culture, because the somatic cell were encapsulated by BC, the competent cell and efficiency of electroporation were effected. The BC obtained by scraping solid plate culture was less than that of solution

culture. BC was digested into glucose and cellobiose by cellulase, then the glucose cellobiose and cellulase was removed through washing cells for three times. So the competent cells could be prepared.

When the competent cells were placed in electric field, the cell membrane played a role of resistor. Electric current can not pass through cell, but cell membrane could be polarized and potential difference could produce on both sides of the cell membrane with voltage rising. After the potential difference exceeded cutoff, the part of cell membrane was punctured and came into being instantaneous hole, a lot of molecules could pass in and out cell through those holes [10]. Because of looseness of cell wall structure, the cell to be in exuberant growth period was sensitive to electric shocking and fit for electroporation. The result indicated that the efficiency of electroporation was highest when culturing *Gluconacetobacter xylinum* to 2 d. But the electric breakdown voltage strength could affect the result for electroporation, so it was important that the suitable electric breakdown voltage strength obtained.

It had been proposed that the highest transformation efficiency of electroporation and obtaining competent cell was in 4 °C for majority of bacteria [11,12]. This study was in agreement with this conclusion. The speculating cause as follows: first, electroporation could reduce heat injury and increase cell livability under low temperature. Second, the opening time for cell holes prolonged and the foreign DNA probability of F-to-enter increased under low temperature.

When electrical field strength exceeded threshold limit value (TLV), the cell membrane could produce holes, the computing formula for the TLV of electrical field strength as follows: $E=10000/1.5 r$.

E refers to electrical field strength (V/cm); r refers to the cell radius (μm)^[13]. Because the cell of *Gluconacetobacter xylinum* radius was $1.5 \times 4.0 \mu\text{m}$, the TLV of electrical field strength should be 1.67~4.44 kV/cm. So the study was defined from 1.6 to 3.2 kV/cm as range of study. As a result, 2.4kV was suitable.

4 Conclusion

The effect of electroporation conditions was systematically investigated. The enzymolysis condition for removing bacterial cellulose of *G. xylinum* was optimized by Response Surface Method; the results showed as follows: cellulase concentration 1046.4 U/ml, enzymolysis duration 93.8 min at 30.6 °C. Optimum transmission conditions including growth stage of the strain, electroshock voltage, concentration and preservation of competent cell were defined for the electroporation with plasmid pDN19Lac Ω . The result was shown that the highest transformation efficiency was up to 8.76×10^5 CFU/ μg DNA under the optimum conditions in which the competent cells were collected at 2 d and concentration to about 1×10^8 CFU/mL, the mixture of the competent cells and plasmid pDN19Lac Ω was eletroporated at 2.4kV. The transformation efficiency was significantly reduced if competent cells were preserved at -70° C.

References

1. Wan, Y.Z., Hong, L., Jia, S.R.: Synthesis and characterization of hydroxyapatite– bacterial cellulose nanocomposites. *Compos. Sci. Technol.* 66, 1825–1832 (2006)
2. Yoshiaki, Y., Shin, K., Kenji, T.: Structural characterization of the *Acetobacter xylinum* endo- β -1, 4-Glucanase CMCax required for cellulose biosynthesis. *Proteins* 64(4), 1069–1077 (2006)
3. Romling, U.: Molecular biology of cellulose production in bacteria. *Research in Microbiology* 153, 205–212 (2002)
4. Becker, D.M., Guarente, L.: High-efficiency transformation of yeast by electroporation. *Methods Enzymol.* 194, 182 (1991)
5. Dower, W.J., Miller, J.F., Ragsdale, C.W.: High efficiency transformation of *E.coli* by high voltage electroporation. *Nucleic Acids Res.* 16(13), 6127–6145 (1988)
6. Delorme, E.: Transformation of *Saccharomyces Cerevisiae* by Electroporation. *Appl. Environ. Microbiol.* 55(9), 2242–2246 (1989)
7. David, S., Simons, G., De, W.M.: Plasmid transformation by electroporation of *euconostoc paramesenter-oides* and its use in molecular cloning. *Appl. Environ. Microbiol.* 55(6), 1483–1489 (1989)
8. Jia, S.R., Ou, H.Y., Chen, G.: Cellulose production from *Gluconobacter oxydans* TQ-B2. In: *Biotechnology and Bioprocess Engineering*, pp. 166–170 (2004)
9. Tang, W.H., Jia, S.R., Yang, H.J.: The influence of fermentation conditions and post-treatment methods on porosity of bacterial cellulose membrane. *World Journal Microbiology Biotechnology* 26, 125–131 (2010)
10. Qin, Y.J., Jin, J.L., Bao, X.M.: Transformation of *Scerevisiae* by electroporation. *Journal of Shan Dong University* 34(2), 236 (1999)
11. Dower, W.J.: Electroporation of bacteria: a general approach to genetic transformation. *Genet. Eng. (N Y)* 12, 275–295 (1990)
12. Ling, L.J., Zhu, Y., Wang, H.M.: Effect of storage time and temperature on transformation rate of *E. coli* competent cells. *Letters in Biotechnology* 15(1), 51–52 (2004)
13. Shimogawara, K., Fujiwara, S., Grossman, A.: High-efficiency transformation of *chlamydomonas reinhardtii* by electroporation. *Genetics* 148, 1821–1828 (1998)

Study on the Mechanism of Apoptosis in the Swine Small Intestine Epithelium Treated by Heat Stress

Dan Jia¹, Fenghua Liu^{1,*}, Weili Luan¹, Kaijun Guo^{1,*}, An Lu¹, and Jin Yu²

¹ College of Animal Science and Technology, Beijing University of Agriculture, Beijing, China, 102206

² College of Veterinary Medicine, China Agricultural University, Beijing, China, 100193
{liufenghua1209, kjguo126}@126.com

Abstract. This research was to investigate the alteration of the expression of apoptotic genes and proteins in the small intestine of swine, so as to provide the indexes of apoptosis treated by heat stress. Based on the swine heat stress model, this research detected the expression of the genes include Caspase-3 and Caspase-8 by RT-PCR, and the protein Caspase-3 by immunohistochemistry. The results showed that, compared to the control group, the gene expressions of Caspase-3 and Caspase-8 in the heat stress group improved significantly, the Caspase-3's protein expression in the head of the intestinal villi increased significantly. The result indicated that the level of apoptosis in the cells of the small intestine got higher after heat stress, especially in the head of the intestinal villi, and the extrinsic pathway was activated. This study provided the factors to evaluate the apoptotic state on the heat stress model, and furthermore, provided the drug's functional targets and evaluative factors.

Keywords: apoptosis, Caspase-3, Caspase-8, swine.

1 Introduction

The pigs are unresistant to heat. The Subcutaneous fat under the skin are very deep and sweat glands are poor developed. The pigs they can't regulate the body temperature through skin[1]. Under heat stress, the animals' blood begins to redistribute, peripheral blood flow is increased to dissipate internal body heat, thereby resulting a significant reduction in blood flow to the small intestine[2,3]. In the condition of heat stress, the intestine is subjected to ischemia and hypoxia first and recovers latest, which lead it easily to be hurt. Research showed that the swine intestine had the most severe injuries on the third day after heat stress. The mucous epithelial at the top of the villus fell off and the basal layer naked. The height of the villus and the depth of the crypt decreased, and the injury at the jejunum was more serious than those at duodenum and ileum[4].

There are three kinds of cell death: apoptosis, autophagy and necrosis. Apoptosis is physiological death in the control of regulator [5]. It plays an important role in function of cell number, development of embryo and immunity, also included in the happening,

* Corresponding author.

development, curing of the tumor [6]. In this process, Caspase-3 is considered as the most important apoptotic executor and its activation is the symbol of cell developing into unreversible stage [7]. Caspase-8 can active almost all caspases and is thought to be at the top of apoptosis process [8]. There is few research to describe the apoptotic pathway in swine intestine.

Based in the swine heat stress model established previously, this research were to detect the genes of Caspase-3 and Caspase-8 by RT-PCR, to find the apoptotic pathway in the pigs, and at the same time, detect Caspase-3 by immunol histochemistry to make sure the distribution of apoptosis in the intestine.

2 Materials and Methods

2.1 Animals

Ten 2 months old Chinese experimental mini-pigs (CEMP), weighing 20.0 ± 2.0 kg, were purchased from experimental pig farm of China Agricultural University. The pigs were treated according to the protocol approved by the Committee for Experimental Animals Care and Use at Beijing University of Agriculture.

2.2 Treatment

Pigs were subdivided into control group and heat stressed group, 5 heads each. Pigs in the control group were housed in a controlled environment (25°C , 60% relative humidity (RH)); while pigs in the heat stressed group were housed under control group conditions, but exposed to 40°C and 60% relative humidity between 4:00 am and 9:00 pm daily. On day 3, five pigs from each group were sacrificed immediately following the 5-h heat exposure time period.

2.3 Fixing Intestinal Sections and Staining

Small intestine tissue samples were promptly rinsed with physiological saline and immediately fixed in 10% buffered formalin phosphate following removal from the animal. The formalin-fixed samples were embedded in paraffin and transversely sectioned (5 mm thick). After deparaffinisation and dehydration, the sections of duodenum, jejunum and ileum (HyClone, Logan, Utah, USA) were stained with hematoxylin and eosin (Sigma, St. Louis, MO, USA). Microstructures of the small intestine were observed using a BH2 Olympus microscope (DP71, Olympus, Tokyo, Japan) and analysed using an Olympus Image Analysis System (Olympus 6.0).

2.4 Total RNA Expression and Reverse Transcription

Total RNA was isolated from the small intestine using a phenol and guanidine isothiocyanatebased TRIZOL reagent (Invitrogen). RNA concentration and purity were assessed using a spectrophotometer (SmartSpec plus, BIO-RAD) utilising the OD260/OD280 ratio. Total RNA was reverse transcribed as follows: 2 mg of RNA isolated from each sample was added to 25 mL of reaction solution comprising 2 mL oligo-dT18, 5 mL dNTP, 1 mL RNase inhibitor, 1 mL MLV transcriptase, 5

mLM-MLV reverse transcription (RT) reaction buffer (Promega, Madison, Wisconsin, USA) and 11 mL RNase-free water. The parameters for the reverse-transcription procedure, based on the manufacturer's instructions (Promega), were: 70°C for 5 min followed by 42 °C for 2 h. The RT products (cDNA) were stored at -20°C for subsequent PCR. Validation of growth factor mRNA expression using real time PCR.

To corroborate the relative changes in gene expression in response to heat stress both in vivo and in vitro identified by the oligonucleotide microarrays, reverse transcriptase-PCR (RT-PCR) was employed. The cDNA of each sample was subjected to real-time RT-PCR using the primer pairs listed in Table 1. Cycling conditions were: 40 cycles of 94°C for 30 s, 56°C for 30 s and 72°C for 40 s. Dissociation was initiated by 95°C for 1 min.

2.5 Statistical Analysis

Statistical analysis was performed by independent-sample T-tests using SPSS 12.0 software. A p-value of less than 0.05 was considered to indicate a significant difference.

Table 1. Primer sequences of PCR product

| Name | Oligonucleotide Sequences | Size (bp) | annealing temperature (°C) | source |
|--------------------------------------|---|-----------|----------------------------|--------|
| Caspase-3 Forward Reverse | 5-TGAGCATGGAAACAAT ACATGG-3 5-ATTCTTGGCGAAATTC AAAGG-3 | 198 | 60 | NCBI |
| Caspase-8 Forward Reverse | 5-TCCTGAGCCTGGACTA CAT-3 5-CTCCTCCTCATTGGTTT CC-3 | 188 | 60 | NCBI |
| β -actin Forward Reverse | 5-GCGGCATCCACGAAAC TAC -3 5-AGAAGCATTTGCGGTG GAC-3 | 312 | 60 | NCBI |

3 Results and Discussion

3.1 The Changes of Genes Caspase-3, Caspase-8 in Heat Stress

Compared with the control group, the expression of genes Caspase-3, Caspase-8 increased in the duodenum, jejunum, ileum in the heat stressed group (Fig. 1, 2, 3).



Fig. 1. The results of PCR of Caspase-3 and β -actin (from left to right:CD,SD,CI,SI,CJ,SJ. CD-control group duodenum, SD-heat stressed group duodenum, CI- control group jejunum, SI- heat stressed group jejunum, CJ- control group ileum, SJ- heat stressed group ileum)

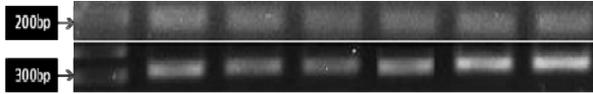


Fig. 2. The results of PCR of Caspase-8 and β -actin (from left to right:CD,SD,CI,SI,CJ,SJ. CD-control group duodenum, SD-heat stressed group duodenum, CI- control group jejunum, SI- heat stressed group jejunum, CJ- control group ileum, SJ- heat stressed group ileum)

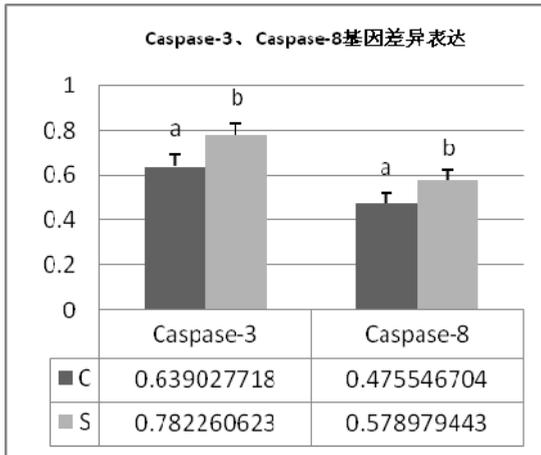


Fig. 3. The Caspase-3 and Caspase-8 mRNA genes expression (C- control group, S- heat stressed group, $P < 0.05$)

3.2 The Changes of Protein Caspase-3 in Heat Stress

Compared with the control group, the expression of genes Caspase-3, Caspase-8 increased in the duodenum, jejunum, ileum in the heat stressed group. The expression at the top of the villus was evident (Fig. 4, 5).

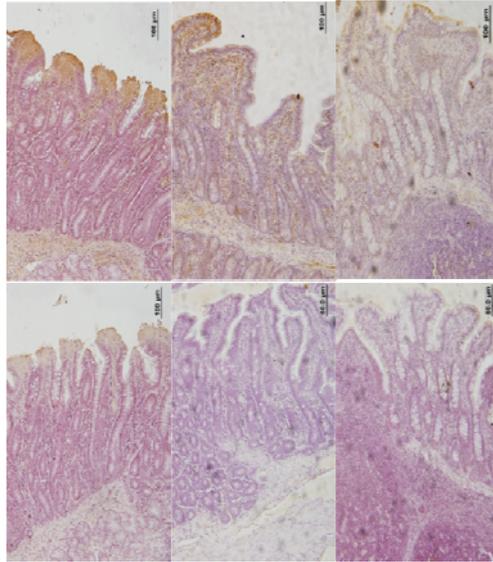


Fig. 4. The caspase-3 protein expression by immuno-histochemistry (first level: CD, SD; second level: CI,SI; third level : CJ, SJ. CD-control group duodenum, SD-heat stressed group duodenum, CI- control group jejunum, SI- heat stressed group jejunum, CJ- control group ileum, SJ- heat stressed group ileum. The deep area is the caspase-3 positive area)

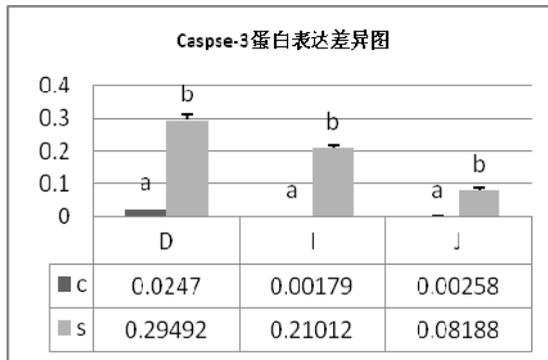


Fig. 5. The caspase-3 protein expression (D- duodenum, I- jejunum, J- ileum, C- control group, S- heat stressed group, $P < 0.05$)

4 Discussion

4.1 The Influence of Heat Stress on the Gene Expression of Caspase-3

There are two pathways of apoptosis according to the signal source: the death receptor pathway and the mitochondria pathway. These two pathways converge into downstream caspases. The caspase-3 is considered as the most important apoptotic

executor. The activation of caspase-3 is recognized as symbol of unreversibility of apoptosis [7].

There was few research about heat stress reaching molecular level. This research detected during heat stress process, the gene expression of caspase-3 increased, the apoptosis aggravated and become irreversible. This result indicated that the apoptosis aggravation is one of major reasons for intestine injury. This research revealed the mechanism of injury in the molecular level in heat stress for the first time.

4.2 The Influence of Heat Stress on the Gene of Caspase-8

The death receptor pathway of apoptosis is signal transmission intermediated by death receptors (Fas, TNFR and TGF-BR), which interact with their partners, and then activate the down-stream proteins. The Fas interacts with FasL. Through death functional area, they can activate FADD, which in turn interacts with DED and procaspase-8, inducing procaspase-8 into caspase-8, which activates down-stream caspases. Caspases-8 is at the top of caspase process [8].

This research revealed the gene expression of caspase-8 increased after heat stress, proving that the death receptor pathway was activated during heat stress. Caspase-8 could be considered as a indicator to evaluate the apoptotic model, as well as a target of medicine action [9, 10].

4.3 The Influence of Heat Stress on the Protein Expression of Caspase-3

Many diseases involve the abnormal apoptosis caused by the activation of caspase-3 [11]. Research showed that the inhibitor of caspase-3 has obvious cure reaction in many diseases [12, 13]. As a result, caspase-3 was thought a target for medicine development.

This research show that the caspase-3 expression increased significantly in the duodenum, jejunum, ileum in the heat stress group, and the top of the villus was more evident. Both at gene and protein level, caspase-3 increased significantly. Caspase-3 could be taken as a indicator of evaluation heat stress model and a target for medicine development.

5 Conclusion

Heat stress can aggregate the apoptotic condition of the swine intestine, and the apoptosis was more evident at the top of villus. The gene and protein level of caspase-3 increased significantly, while the gene level of caspase-8 increased. This research could provide not only evaluating mechanism of apoptosis in heat stress model, but also a target indicator for medicine development.

References

1. Chu, Y., Wei, Q., Hu, L., Yang, Q., et al.: The development, prevention and curation of pigs in summer. *Xiandai Nongye Keji* 11, 23–25 (2007)
2. Samain, A.: Digestive hemorrhages during stress. *Acta Chir. Belg.* 65, 574–576 (1966)

3. Radwan, M.A., EI-Gendy, K.S., Gad, A.F.: Oxidative stress biomarkers in the digestive gland of *Theba pisana* exposed to heavy metals. *Arch. Environ. Contam. Toxicol.* 58, 828–835 (2010)
4. Yu, J., Yin, P., Liu, F., et al.: Effect of heat stress on the porcine small intestine: A morphological and gene expression study. *Comparative Biochemistry and Physiology, Part A* (2010)
5. Hengartner, M.O.: Apoptosis: corralling the corpses. *Cell* 104, 325–328 (2001)
6. Johnstone, R.W., Ruefli, A.A., Lowe, S.W.: Apoptosis: a link between cancer genetics and chemotherapy. *Cell* 108, 153–164 (2002)
7. Cryns, V., Yuan, J.: Protease to die for. *Genes. Dev.* 129(11), 1551–1570 (1998)
8. Kukhta, V.K., Marozkina, N.V., Sokolchik, I.G., et al.: Molecular mechanisms of apoptosis. *Ukr Biokhim Zh* 75(6), 5–9 (2003)
9. Fesik, S.W., Shi, Y.: Structural biology. Controlling the caspases. *Science* 294, 1477–1478 (2001)
10. Kumar, S., Vaux, D.L.: Apoptosis. A cinderella caspase takes center stage. *Science* 297, 1290–1291 (2002)
11. O'Brien, T.: Prospects for caspase inhibitors. *Mini Rev. Med. Chem.* 4, 153 (2004)
12. Ye, H., Cande, C., Stephanou, N.C., et al.: DNA binding is required for apoptogenic action of apoptosis inducing factor. *Nat. Struct. Biol.* 9, 680–684 (2002)
13. Wang, X., Zhang, X., Xia, M., et al.: The activation of Caspase. *Xiandai Shengwu Yixue Jinzhan* 6(3), 53–55 (2006)

Effect of Spleen Deficiency Syndrome on the Expression of TLR2 and TLR4 in Rats^{*}

Tao Liu^{1,**}, An Lu¹, Xiaoyu Zhu^{2,**}, Nuowei Zhang², Guilin Cheng¹, Ning Wang¹, Fenghua Liu^{1,2}, Kaijun Guo^{1,2,***}, and Huichuan Wang^{3,***}

¹ College of Animal Science and Technology, Beijing University of Agriculture, Beijing 102206, China

² Key Laboratory of Development and Evaluation of the Chemical and Herbal drugs for Animal Use, Ministry of Agriculture, The People's Republic of China, Beijing 100193, China

³ Department of Basic Science, Beijing University of Agriculture, Beijing 102206, China
kjguo126@126.com, whch56221@sohu.com

Abstract. [Objective]The aim of the paper was to study the expression of TLR2 and TLR4 in the rat's spleen deficiency syndrome (SDS) induced by reserpine. [Method]Twelve SD rats were divided into the control group and SDS group. After 7 days, the rats were sacrificed. The duodenum, jejunum and ileum was removed and washed in phosphatebuffered saline (PBS). One part of each intestine sections was fixed and paraffin embedded. The TLR2 secretion were observed by immunohistochemical staining. The other part was to measure the expressions of TLR2 and TLR4 were measured using RT-PCR. [Result]The results indicated that comparing with control group, the TLR2 secretion of SDS group was significantly decreased. The expression of TLR2 and TLR4 of SDS group in jejunum were significantly decreased ($P < 0.01$), but TLR4 in ileum was increased ($P < 0.01$). [Conclusion]In conclusion, these results indicated that SDS induced damage in jejunum, which resulted in the down-regulation of the expression of TLR2 and TLR4.

Keywords: Spleen deficiency syndrome, small intestine, TLR2, TLR4.

1 Introduction

In Traditional Chinese Medicine (TCM), "spleen" is a physiological and pathological concept, rather than simply the anatomic concept. Its function in fact includes digestion and absorption, energy metabolism, endocrine, blood system, nervous and immune system function in modern medicine [1]. In TCM, the most basic

^{*} The work was partially supported by grants from National Natural Science Foundation of (No. 30771591), National Eleven-five Technological Supported Plan of China (No. 2008BADB4B01; 2008BADB4B07), Funding Project for Academic Human Resources Development in Institutions of Higher Learning Under the Jurisdiction of Beijing Municipality.

^{**} These authors contribute equally to the work.

^{***} Corresponding author.

physiological functions of "spleen" is that it governs transportation and transformation. In recent 20 years, the research on spleen deficiency syndrome mechanism is mainly concerning this theory [2]. The results showed in patients with spleen deficiency syndrome and animal models, there were different degrees of digestive system dysfunction, mainly demonstrating gastrointestinal motility disorder, reduced nutrition absorption, material metabolism disorders, gastrointestinal hormone abnormalities, immune disorders, etc. [3]. Intestinal mucosal homeostasis and Toll receptor (TLRs) were closely correlated [4]. TLRs are the innate immune receptors found in recent 10 years. TLRs are known as "a major breakthrough in the field of Immunology" because TLRs makes the people to know and position of the important role of innate immune system in immune initiation and regulation [5]. At present 11 of TLRs family have been cloned [6], and various pathogen-associated molecular patterns (PAMPs) could be identified [7]. TLRs could be expressed in intestinal mucosal epithelium, lamina propria macrophages, dendritic cells, fibroblasts and vascular endothelial cells. Identification of various PAMPs of symbiotic bacteria or pathogen expression has close relation with mucosal tolerance and protection [8]. Spleen deficiency can down-regulate the effect immune responses, to damage the body's mucosal immune system [9].

TLRs differentially recognize PAMPs, which exerts the important role of connecting innate immunity and acquired immunity. The study on TLRs is becoming the hot point in life science field [11]. However, there is few research on intestinal mucous immunity and TLRs expression of animal with Spleen Deficiency Syndrome.

The aim of this experiment was to observe the expression of TLR4 TLR2 in the duodenum, jejunum, and ileum using by spleen deficiency model of rat induced by reserpine.

2 Materials and Methods

2.1 Materials

2.1.1 Animals

Twelve Sprague-Dewley (SD) adult male rats (250 ± 20 g), from Beijing Weitong Lihua Laboratory Animal Technology Ltd. Co., Beijing, China, were used for the experiment.

2.1.2 Drugs and Reagents

Reserpine injection (1 mg/mL) was purchased from Guangdong Bangmin Pharmaceutical Factory, Guangzhou, China. TRIZOL Reagent was purchased from Invitrogen, Carlsbad, CA, USA. Taq DNA polymerase, M-MLV Reverse Transcriptase and Rnase Inhibitor were from Promega, Madison, WI, USA. SYBR Green Real-time PCR kit was purchased from Stratagene.

2.1.3 Main Instruments

Protein nucleic acid detector (Bio-Rad), BH2 Olympus bio-microscope (Olympus), Image-Pro Plus 5.1 analysis system (Media Cybernetics), Mx3000P Real-time PCR (Stratagene).

2.1.4 Primer Sequences

Primers were supplied by Beijing Aoke bio-tech Co. LTD

TLR2 Forward: 5'-CAGGGATACAGGCCGTC AAG-3',
Reverse: 5'-CCAACACCTCCAGGGTCTGA-3'' ;

TLR4 Forward: 5'-GAATCCCTGCATAGAGGTA CTTC-3',
Reverse: 5'-TGATCCATGCATTGGTAGGTAATATTA-3' ;

β -actin Forward: 5'-TTGTCCCTGTATGCCTCTGG-3',
Reverse: 5'-ATGTCACGCACGATTTCCC-3'.

2.2 Methods

2.2.1 Model Preparation

Spleen deficiency rat model was established by reserpine injection verified previously [12]. Twelve male SD rats were randomly divided into 2 groups (6 / group): control group (CG) and reserpine-treated group (RG). CG rats were intraperitoneally injected with 0.5 mL/kg normal saline. Rats in RG was intraperitoneally injected with reserpine 0.5 mL/kg for 7 consecutive days. On the 8th day, rats were fastened for 6 h without restriction of water and then sacrificed according to the protocol approved by the Committee for Experimental Animals Care and Use at Beijing University of Agriculture.

2.2.2 Observation of SIgA and TLR2 by Immunohistochemical Staining

Method

On the 8th of the experiment, the middle of the duodenum, jejunum and ileum were collected separately, then rinsed with cold saline. One part of the tissue was used for locating observation of SIgA and TLR2 expression using immunohistochemical staining method. The remaining tissue was frozen at - 80 °C for PCR examination.

2.2.3 Total RNA Isolation

Total RNA was isolated from 100 mg of various tissues using Trizol reagent. And its concentration was detected by protein nucleic acid detector. The RNA was reconstituted in 100 μ L of Rnase-free water. The quality of RNA was determined by measuring the OD260/280 ratio with UV-vis spectrophotometry. RNA purified by this method resulted in OD260/280 ratio of 1.8-2.1. Total RNA was stored in - 70 °C for use.

2.2.4 Real Time Quantitative PCR

1 μ g RNA was used to synthesis cDNA using MMLV reverse transcriptase. For each of concerning genes and housekeeping genes, PCR reactions were performed to prepare the gradient dilution DNA template for drawing standard curve. The total volume of reaction system and composition were determined according to the previous research [12]

2.2.5 Drawing Standard Curve

The PCR products were gradiently diluted by 10 times. Assuming the PCR product concentration is 1, the cDNA was diluted into 5 gradient (10^{-3} , 10^{-4} , 10^{-5} , 10^{-6} , 10^{-7}).

SYBR Green real-time quantitative PCR reaction was performed for target gene TLR2, TLR4 and reference gene β -actin. Accordingly, standard curve was drawn and the amplification efficiency was detected.

2.2.6 Calculation of Relative Content by Quantitative Fluorescence

The relative expressions of TLR2, TLR4 mRNA were analyzed using Pfaffl method[13]. The calculation formula is as follows: the relative content = $2^{-[\Delta\Delta Ct]}$,
 $\Delta\Delta Ct = RG \Delta Ct$ (target genes - β -actin) - $CG\Delta Ct$ (target genes - β -actin)

2.2.7 Statistical Analysis

The data were analyzed using a standard software package (SPSS 11.0; SPSS Inc., Chicago, IL, USA). All data were reported using means \pm standard deviations.

3 Result

3.1 Weight Gain

Weight changes in rats were significantly different. The weight of CG rats on 8th day was significantly increased compared with that of the first day (245 ± 8 g vs 275 ± 5 g, $P < 0.01$). The weight of RG rats on 8th day was significantly reduced compared with that of the first day (246 ± 4 g vs 180 ± 11 g, $P < 0.01$).

3.2 Spleen Deficiency on Rat Duodenum, Jejunum and Ileum TLR2 Expression Influence

As shown in Fig. 1, the brown color were TLR2 protein expression position. TLR2 in CG expressed in brush border and within the cytoplasm of rat intestinal villus tip cell, as well as in the submucosa. TLR2 in RG expressed at the same position of CG, but the color is weak.

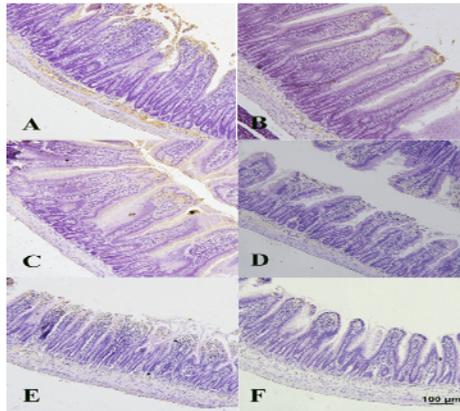


Fig. 1. TLR2 Protein Expression Changes in intestine induced by reserpine. A, C, E(100 \times) show duodenum, jejunum, ileum of control group, B, D, F(100 \times) show duodenum, jejunum, ileum of reserpine treated group.

3.3 Standard Curve and Amplification Curve of RT - PCR

CT values of gradiently diluted cDNA were calculated and drawn diagram. Standard curve of TLR2, TLR4 and β -actin genes were then generated. The slope of their standard curve were 3.096, 3.002 and 3.087 respectively. The correlation coefficients were 0.995, 0.998 and 0.999 and the reaction efficiency is higher than 0.99. these data indicated that the system can be used for fluorescence quantitative PCR system, and the results are accurate and reliable.

After reverse transcription from RNA to cDNA, the latter functioned as a template for SYBR GreenI RT PCR. Fig. 2 showed TLR2, TLR4 and β -actin gene amplification stage and platform stage were obvious, indicating ideal amplification curves. The results implied that the experimental data could be used for relative quantitative analysis of TLR2, TLR4 and β -actin gene.

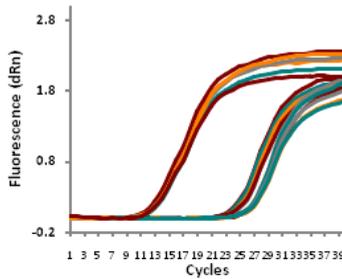


Fig. 2. Amplification Plots of TLR2, TLR4 and β -actin

3.4 Melting Curve

Through the melting curve analysis, it could be found that there were single peaks for β -actin gene at 73°C, TLR2 gene at 79°C, and TLR4 at 87°C, indicating that these genes were specifically amplified.

3.5 Relative Expression of TLR2 and TLR4 mRNA in RG Rat Small Intestine

Compared with CG, the relative expressions of TLR2 and TLR4 mRNA in RG rat jejunum were significantly lower ($P < 0.01$), TLR4 mRNA expression of in RG rat ileal was significantly lower ($P < 0.01$). TLR2 and TLR4 mRNA expression in the duodenum had no significant difference (Fig. 3).

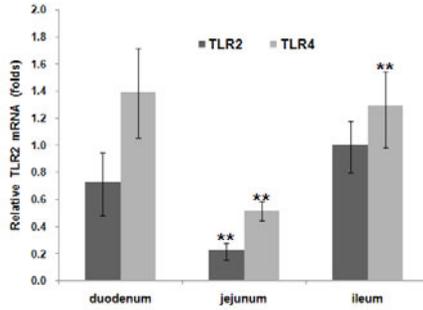


Fig. 3. Changes in the relative gene expression of TLR2 and TLR4 in rat intestine (**: $P < 0.01$)

4 Discussions

Intestinal mucosal immunity is the first line of defense against infection, and spleen has a close relationship with low gut mucosal immune function[14]. Research in recent years showed the TLRs family had its indispensable importance in the innate immunity. The TLRs family plays a very important role in the innate immunity and induction of acquired immune responses. TLR2, TLR4 are two important members in TLR family. Its existence has an important protective function for intestinal mucosa.

Under normal conditions, TLR2 and TLR4 distributed in non specific antigen presenting cells such as intestinal mucosa epithelial cells. Through the identification of pathogens, intestinal mucosa epithelial cells were induced to produce cytokines, chemokines and antimicrobial peptides. Chemokines secreted by epithelial cells can spread to the vascular endothelial surface of surrounding tissue and lymphoid tissue, involved in cell recruitment [15]. This experiment adopts SYBR Green I RT-PCR assay for detection of TLR2, TLR4 mRNA expressions in different intestinal segments. The results showed TLR2 and TLR4 expression in jejunum are significantly reduced, resulting in the reduction of jejunal antigen recognition ability and affecting intestinal mucosal innate immune function. Data display TLR plays an important role in tissue injury, tissue repair and regeneration process. In the intestinal injury models caused by chemical factors, radiological factors and infection factors, the TLR signaling pathway played an important role in maintaining tissue integrity, repairing the injured tissues [16].

Study on the expression of TLR2 and TLR4 can help to reveal the deficiency of mucosal immune status, especially the inner mechanism of the spleen deficiency influencing of natural immune. TLRs can be used as an important cutting point in animal intestinal immune system with spleen deficiency. Further study of its signaling pathways could provide new ideas and methods for the clinical prevention and treatment of spleen deficiency.

Acknowledgements. The work was partially supported by grants from National Natural Science Foundation of China (No. 30771591), National Eleven-five Technological Supported Plan of China (No. 2008BADB4B01; 2008BADB4B07), Funding Project for Academic Human Resources Development in Institutions of Higher Learning Under the Jurisdiction of Beijing Municipality.

References

1. Peng, J.Q., Li, Q.M., Yuan, H.P.: Mechanism of acupuncture on intestinal tract local immunity of splenasthenic syndrome. *Journal of Changchun University of Traditional Chinese Medicine* 23(2), 84–85 (2007) (in Chinese)
2. Wu, Y.F., Qi, Q.H.: The study of spleen-deficiency syndrome. *Journal of Tianjin Medical University* 1, 135–137 (2003) (in Chinese)
3. Ren, P., Huang, X., Zhang, H.X., Huang, Y.X., Yang, X.Z., Zhao, B.M., Zhao, Y.L.: Significant changes of gastric acid secretion, gastroduodenal movement and electrogastrography in patients with spleen-asthenia. *World Chinese Journal of Digestology* 12(3), 228–231 (2004) (in Chinese)
4. Schenk, M., Mueller, C.: The mucosal immune system at the gastrointestinal barrier. *Journal of Best Practice and Research Clinical Gastroenterology* 22(3), 391–409 (2008)
5. Li, M., Carpio, D.F., Zheng, Y., Bruzzo, P., Singh, V., Ouaz, F., Medzhitov, R.M., Beg, A.A.: An essential role of the NF-kappa B/Toll-like receptor pathway in induction of inflammatory and tissue-repair gene expression by necrotic cells. *Journal of Immunology* 166, 7128–7135 (2001)
6. O'Neill, L.A.: Immunology: After the Toll Rush. *Science* 303, 1481–1482 (2004)
7. Moynagh, P.N.: TLR signalling and activation of IRFs: revisiting old friends from the NF-kappaB pathway. *Trends Immunology* 26, 469–476 (2005)
8. Harris, G., Kuolee, R.: Role of Toll-like receptors in health and diseases of gastrointestinal tract. *World Journal of Gastroenterology* 14, 11–22 (2006)
9. Xu, S.J., Xiao, C., Zhou, G.Q., Lv, A.P., Wang, A.M.: Studies on Functions of Mucosal Immune in Collagen-induced Arthritis Superimposed on Spleen Deficiency in Mice. *Chinese Journal of Basic Medicine in Traditional Chinese Medicine* 9(12), 35–38 (2003) (in Chinese)
10. Kawai, T., Akira, S.: Toll-like receptor downstream signaling. *Arthritis Research and Therapy* 7, 12–19 (2005)
11. Uematsu, S., Akira, S.: Toll-like receptor s and type I interferons. *Journal of Biological Chemistry* 282(21), 15319–15323 (2007)
12. Cheng, F., Liu, F.H., Zhu, X.Y., Gan, J.Y., Wang, W., Cheng, G.L., Song, X.P., Xu, J.Q.: Gastrointestinal secretin and gastric inhibitory peptide mRNA expressions and their effects on pancreatic external secretion in spleen deficiency rats. *Chinese Journal of Veterinary Medicine* 45(7), 3–6 (2009) (in Chinese)
13. Yu, X.W., Liu, H.Y., Luo, L.J.: Analysis of Relative Gene Expression Using Different Real-Time Quantitative PCR. *Acta Agronomica Sinica* 7, 1214–1218 (2007) (in Chinese)

Reserpine Caused Gastrointestinal Damages and Hormonal Changes in Rats*

Kaijun Guo^{1,**}, Fei Cheng^{2,3,**}, Xiaorui Wang², Xiaoping Song^{2,***}, Xiaoyu Zhu³, Jingyi Gan³, Fenghua Liu¹, and Jianqin Xu^{3,***}

¹ College of Animal Science and Technology, Beijing University of Agriculture, Beijing 102206, P.R China

² College of Veterinary Medicine, Northwest A&F University, Yangling, Shaanxi, 712100, P.R China

³ College of Veterinary Medicine, China Agricultural University, Beijing 100193, P.R. China
{xpsong126, jianqinxucau}@126.com

Abstract. The aim of the present study is to investigate the changes of gastrointestinal (GI) hormones in reserpine-treated rat. Reserpine (0.5 mg/kg body weight) was injected once per day for continuous 14 days. The histological changes of GI were observed and GI hormones variation was measured. The results showed that comparing with the control group, the reserpine-treated rats had obvious damages. In the GI damaged duodenum mRNA and protein content of gastrin (GAS), cholecystokinin (CCK), vasoactive intestinal polypeptide (VIP) were significantly decreased and somatostatin (SS) protein content were significantly increased comparing the control group. In the gastric antrum of reserpine-treated rats, mRNA and protein content of GAS increased significantly while those of CCK were significantly decreased comparing the control group. These results suggest that the GI hormones changes were associated with gastro and small intestine damage in reserpine-treated rats.

Keywords: Reserpine-treated rat, GI damage, GI hormone.

1 Introduction

GI hormones are chemical messengers that regulate the physiological functions of the intestine, including secretion, motility, absorption and digestion [1]. In addition to these well-defined physiological effects, GI hormones can stimulate the proliferation of intestinal mucosa and protect the intestinal mucosal barrier [2]. Gastrin (GAS), somatostatin (SS), cholecystokinin (CCK), vasoactive intestinal polypeptide (VIP)

* The work was partially supported by grants from National Natural Science Foundation of China (No. 30771591), National Eleven-five Technological Supported Plan of China (No. 2008BADB4B01; 2008BADB4B07), Funding Project for Academic Human Resources Development in Institutions of Higher Learning Under the Jurisdiction of Beijing Municipality.

** These authors contributed equally to this work.

*** Corresponding author.

were important GI hormones, which belong to the brain-intestinal peptides and are closely related with the GI digestion and absorption functions. GAS is well established as a growth factor for the gastro-intestinal tract and the pancreas, stimulating the proliferation of the enterochromaffin-like cells in the stomach and the proximal duodenum, and of pancreatic epithelial cells [3]. SS functions as a potent inhibitor of hormone and prevents GI damage [4]. CCK and VIP also has a dual role as a neurotransmitter, in both the enteric and central nerve system. CCK can increase secretion of pancreatic enzymes and bicarbonate, inhibits gastric acid secretion, slows gastric emptying [5].

Reserpine-induced gastric damage in rodents has been extensively described. Reserpine is thought to induce gastric ulcers mainly via the depression of adrenergic activity with an increase in the cholinergic tone [6]. GI functions are collaboratively affected and regulated by nerves and endocrine system. Some reports indicate that the enterochromaffin cells have been changed in intestine of reserpine-treated Guinea pigs [7]. While there is few information about the GI hormones changes caused by reserpine.

The aim of this study was to investigate the changes of GI hormones (GAS, SS, VIP, CCK) expression in reserpine-treated rat.

2 Materials and Methods

2.1 Drugs and Reagents

Reserpine injection (1 mg/mL) was purchased from Guangdong Bangmin Pharmaceutical Factory, Guangzhou, China. Haematoxylin and eosin was procured from Sigma Chemical Co., St. Louis, MO, USA. TRIZOL Reagent was purchased from Invitrogen, Carlsbad, CA, USA. Taq DNA polymerase, M-MLV Reverse Transcriptase and Rnase Inhibitor were from Promega, Madison, WI, USA. Agarose was from Promega, Beijing, China. Goat anti-Rat Gastrin TH was purchased from ADI, Texas, USA; Rat Somatostatin ELISA from R&B, Calabasas, California, USA; Rat VIP ELISA from R&B, Calabasas, California; Rat CCK ELISA from R&B, Calabasas, California.

2.2 Animals and Protocol

Forty eight Sprague-Dewley (SD) adult male rats (240 ± 20 g), from Beijing Weitong Lihua Laboratory Animal Technology Ltd. Co., Beijing, China, were used for the experiment and maintained under standard laboratory conditions at about 25°C, relative humidity of about 45% for a 3-day adaptation period before tests. Standard laboratory chow diet and water were provided *ad libitum* at fixed time every morning. These rats were randomly divided into 2 groups: control group (CG) and reserpine-treated group (RG), 24 rats each. Rats in RG was intraperitoneally injected with reserpine 0.5 mL/kg and CG rats with normal saline 0.5 mL/kg. On the 3, 7, 10, 14 days of treatment, 6 rats of each group randomly sacrificed according to the protocol approved by the Committee for Experimental Animals Care and Use at Beijing University of Agriculture.

2.3 Clinical Signs and Body Weight

The clinical signs and body weight of each rat in two experimental groups were assessed and recorded at 8 o'clock every morning on day 3, 7, 10 and 14.

2.4 Histological Observation

Six subjects were randomly selected from each group on days 3, 7, 10 and 14 and sacrificed immediately. Tissues from stomach and small intestine were removed and immediately washed clear in normal saline, and then every two full-thick tissue masses were obtained respectively from gastric antrum $2 \times 1 \text{ cm}^2$ (the ventrolateral wall of gastric antrum from pylorus to gastric pit angular incisure), 2 cm duodenum (5 cm from the pylorus), 2 cm jejunum (the middle part of jejunum), and 2 cm ileum (5 cm proximal to the ileocecal orifice). Each tissue was divided into two parts, one was used for histological analysis and another was stored in the liquid nitrogen at -80°C for the detection of GAS and SS mRNA expression and protein content.

GI tissue sections histological analysis were immediately fixed in 10% neutral formalin for 24 h and embedded in paraffin wax. Briefly, paraffin-embedded tissues were cut at a thickness of $4 \mu\text{m}$ using RM2126 rotary microtomes (LEICA, Wetzlar, Germany) and stained with haematoxylin and eosin. Haematoxylin staining time was for 5 min and eosin for 45 sec. Histological changes in the sections of various tissues were observed under BX51 microscope (Olympus, Tokyo, Japan) (10×20 magnification).

2.5 Total RNA Isolation and RT-PCR

Total RNA Isolation: Total RNA was isolated from 100 mg of various tissues using Trizol reagent. And its concentration was detected by nucleic acid-protein detector. The RNA was reconstituted in $100 \mu\text{L}$ of Rnase-free water. The quality of RNA was determined by measuring the OD260/280 ratio with UV-vis spectrophotometry. RNA purified by this method resulted in OD260/280 ratio of > 1.7 . The quality of RNA was further determined by formaldehyde agarose gel electrophoresis. The presence of 18S and 28S rRNA indicated high quality with low degradation.

RNA Reverse Transcription (RT): It was performed as described previously. RNA reverse transcription reaction system was $25 \mu\text{L}$: $1 \mu\text{g}$ total RNA template, $5 \mu\text{L}$ Oligo-dT18 primer ($0.5 \mu\text{g}/\mu\text{L}$) was added to reaction system, adjusted to a volume of $13 \mu\text{L}$ with DEPC-treated water and mixed gently. The reaction was incubated at 70°C for 5 min, then cooled on ice immediately and according to the order the following reagents were added: $1.25 \mu\text{L}$ dNTPs (10 mmol/L), $0.5 \mu\text{L}$ Rnasin ($40 \text{ U}/\mu\text{L}$), $5.0 \mu\text{L}$ M-MLV $5 \times$ RT reaction buffer, $1.0 \mu\text{L}$ M-MLV Reverse Transcriptase ($200 \text{ U}/\mu\text{L}$), and $0.5 \mu\text{L}$ Rnase-inhibitor ($40 \text{ U}/\mu\text{L}$), and DEPC-treated water to $25 \mu\text{L}$. The reactions were mixed gently, incubated at 42°C for 60 min and terminated by incubating at 70°C for 10 min. The cDNA products was collected by brief centrifugation and used immediately for PCR or stored at -20°C .

PCR amplification of GAS, SS, VIP, and CCK mRNA: Two pairs of primers GAS, SS, VIP, and CCK were designed (Table 1) using Premier 5.0 software (Premier, Canada) with reference to GenBank rat gene sequence. PCR was performed

in a 96 well plate. Total volume of PCR reaction system was 20 μ L: 1.0 μ L cDNA template, 1.0 μ L forward and reverse primers (10 μ mol/L), 1.2 μ L MgCl₂ (25 mmol/L), 1.6 μ L dNTPs (2.5 mmol/L), 0.1 μ L Taq DNA polymerase (5 U/ μ L), 2.0 μ L 5 \times PCR buffer, and 13.1 μ L sterile ultrapure water were combined for a total volume of 20 μ L per well. The contents of the well were mixed by pipetting gently. PCR amplified conditions: After pre-denaturing at 94°C for 5 min, there were 34 cycles of PCR amplification performed consisting of denaturing at 94°C for 40 s, annealing at 56°C for 30 s, and extension at 72°C for 30 s and finally extending at 72°C for 8 min. PCR products were stored at -20°C.

The PCR products (15 μ L) were electrophoresed on a 1.2% agarose gel stained with ethidium bromide at a voltage of 100 V for 40 min, photographed by Digital Gel Imaging Analysis System (IS-2200; Alpha, Chicago, USA), an Image Analysis System (Olympus 6.0, Tokyo, Japan) was used to analyze the image of electrophoresis with samples and β -actin, the ratio of samples to β -actin indicated the quantity of nuclear acid.

2.6 Total Protein Extraction and Assays of GAS, SS, VIP, CCK

Total proteins were separately extracted from gastric antrum and small intestine of rats on experimental day 7 and their contents were measured by the BCA method (Pierce, Rockford, Illinois, USA). GAS, SS, VIP, CCK protein contents were assayed by ELISA with Model 680 Microplat Reader (Bio-Rad, Hercules, CA, USA), according to the instructions.

2.7 Statistical Analysis

Statistic evaluations were performed using a standard software package (SPSS 11.0; SPSS Inc., Chicago, IL, USA). All data were reported using means and standard deviations. Student's paired t-test was used to examine the group differences at three sampling time points. P values < 0.05 are considered statistically significant.

Table 1. Primer sequences of PCR product

| Names | | Oligonucleotide Sequences |
|----------------|---------|------------------------------|
| β -actin | Forward | 5'-TTGTCCCTGTATGCCTCTGG-3' |
| | Reverse | 5'-ATGTCACGCACGATTCCCC-3' |
| SS | Forward | 5'-CTGTCCTGCCGTCTCCAGT-3' |
| | Reverse | 5'-GCTCCAGCCTCATCTCGTC-3' |
| GAS | Forward | 5'-AACCTCGTCCCAGCTACA-3' |
| | Reverse | 5'-CCATCCGTATGCTTCCTCTT-3' |
| VIP | Forward | 5'-AGAATGACACGCCCTATTATGA-3' |
| | Reverse | 5'-TTCCGAGATGCTACTGCTGAT-3' |
| CCK | Forward | 5'-CAGCAGGTCCGCAAAGC-3' |
| | Reverse | 5'-GGCCCACTACGATGGGTAT-3' |

3 Results

3.1 Clinical Signs and Body Weight Changes

Three days after receiving reserpine, rats from RG began to appear a series of serious signs such as visible diarrhea, semiliquid stool with perianal dirty, poor appetite, inactivity, lethargy, crowding, haggard hair, body weight loss, etc. Rat weight was significantly decreased in RG than that in CG after the 7th day, and especially at day 10 (155 ± 8 g vs 290 ± 10 g). The subsequent analysis were carried out using the 7th day sample.

3.2 Histological Changes of GI Tissue

Microscopical analysis of gastric antrum and small intestine in RG rats and CG rats were shown in Fig. 1. Comparing with CG, RG had gastric evidences of mucosa damages (Fig. 1B); duodenum mucosal epithelium abscission, duodenal glands atrophy (Fig. 1D).

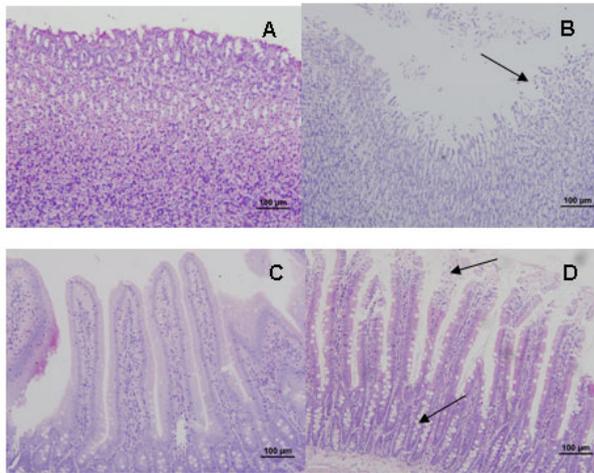


Fig. 1. Morphological examination of gastric mucosa and small intestine of rats (20×). (A) Gastric section in control group; (B) Gastric section in reserpine-treated group; (C) Control Duodenum section in control group; (D) Duodenum section in reserpine-treated group. There are evidences of mucosa damages in B, Duodenum mucosal epithelium abscission, duodenal glands atrophy in D.

3.3 GAS, SS, VIP and CCK mRNA Expression

Fig.2 shows GAS, SS, VIP, CCK mRNA expression in gastric antrum, duodenum, jejunum and ileum of control group rats ($n = 4$) and reserpine-treated group rats. Compared with CG, GAS and VIP mRNA expression in the duodenum of RG were significantly lower than those of CG (Fig. 2A, C). CCK mRNA expressions in both gastric antrum and duodenum of RG had been reduced significantly than those of CG (Fig. 2D). All others values were not significant different even though there were

some variations between RG and CG. Since no values of Jejunum and Ileum were significantly different between RG and CG, protein content of GAS, SS, VIP, CCK were analyzed only for those of gastric antrum and duodenum.

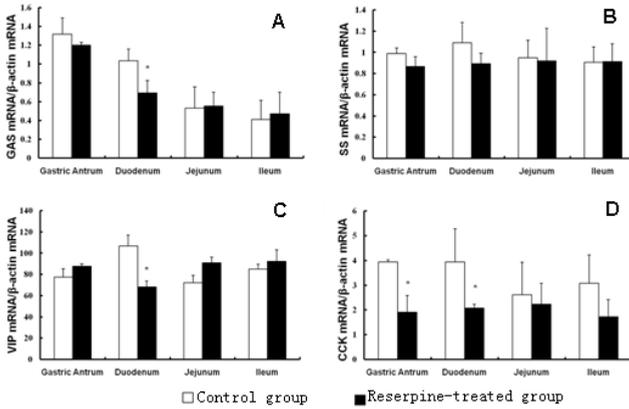


Fig. 2. GAS, SS, VIP, CCK mRNA expression in gastric antrum and small intestine of control group rats (n = 4) and reserpine-treated group rats (n = 4). * means the values of reserpine-treated group significantly different ($P < 0.05$) from those of control groups.

3.4 GAS, SS, VIP and CCK Protein Contents

Comparing with CG, RG GAS protein content were significantly increased in gastric antrum and decreased in the duodenum (Fig. 3A); SS protein contents in RG were increased in the duodenum (Fig. 3B); VIP and CCK protein contents in RG were reduced in gastric antrum (Fig. 3C, D).

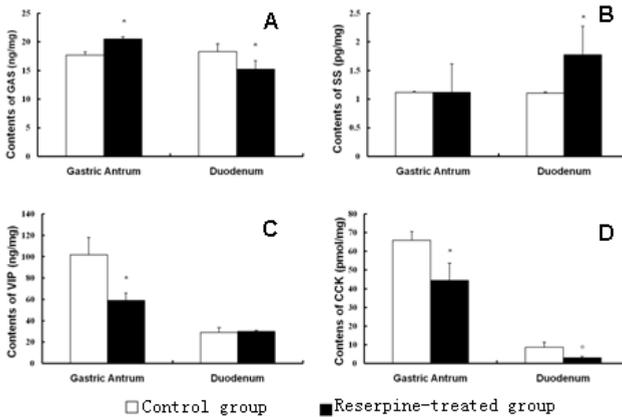


Fig. 3. Difference comparison of GAS, SS, VIP, CCK protein expression of gastric antrum and duodenum between control group (n = 4) and reserpine-treated group (n = 4). (A) GAS protein expression. (B) SS protein expression. (C) VIP protein expression. (D) CCK protein expression. * means the values of reserpine-treated group significantly different ($P < 0.05$) from those of control groups.

4 Discussion

4.1 Effects of Reserpine on Rats GI Tract Structure

In this research, the results showed that intraperitoneal injection of reserpine (0.5mg/kg) in rat, can induce gastro ulcer, duodenal glands atrophy, gastric and duodenal mucosa with different degrees. No obvious inflammation was seen in the GI mucosa. It is known that Gastric ulcer has multifactorial etiology, and gastric acidity, pepsin secretion, gastric motility and gastric mucosal blood flow are involved in the development of ulcers [8]. Clear symptoms of digestive diseases such as Diarrhea were observed in the reserpine-treated rats.

4.2 Reserpine-Induced GI Hormones Changes

GAS is secreted by the G cells located in the gastric antrum and duodenum. It's plays an important role in modulating various GI tract functions, including acid secretion, motility and cell proliferation. In the present study, the gastric antrum GAS mRNA expression was significantly up-regulated in RG comparing CG, which resulted in the increase of GAS secretion. The increased gastric antrum GAS could increase gastric acid secretion, which is an important reason for tissue damage. The duodenal GAS mRNA expression was significantly down-regulated in RG respective to CG, which resulted the decrease of GAS secretion. Studies had shown that duodenal GAS reduction was related to duodenal injury [9]. GAS decrease in duodenum caused slow proliferation of intestinal cell. Although the changing of GAS in the duodenum was not the reason of intestinal injury, but it may plays a role in the repair process.

Somatostatin (SS) is a straight-chain peptide in the brain and GI tract. It has an inhibitory action on the secretion functions of almost all the GI exocrine glands, including the stomach, small intestine and pancreatic exocrine glands, and affects the GI digestion and absorption. Previous researches had confirmed that SS can protect gastric mucosa, promote gastric mucus generation [7]. In the present study, the duodenum SS protein content after reserpine treatment increased significantly, while the corresponding SS mRNA expression had no significant change. This result was in agreement with the study of Yao et al [10].

VIP can promote pancreatic juice and bicarbonate secretion, is conducive to the formation of GI mucosal barrier mucus and repair, and can stimulate the intestinal fluid secretion, inhibit gastric acid and GAS secretion to reduce the incidence of peptic ulcer and promote ulcer repair [11]. In our study, VIP protein in the gastro reduced significantly in RG than in CG, causing gastro-intestinal barrier function and repair capacity decreased. Further study is needed to explain the discordances of mRNA and protein in gastro and duodenum (Fig. 2 and Fig. 3).

This study showed that CCK mRNA was significantly decreased in gastric antrum and duodenum in RG than those in CG. CCK can inhibit gastric acid secretion and gastric emptying. The reduction of gastric CCK may accelerate gastric emptying and the peristalsis of distal duodenum and jejunum, and enhance ileum movement [12]. Therefore, the reduction of CCK is another reason of gastric acid increase, indirectly induce gastric damage.

In summary, the present study showed that Reserpine can cause gastrointestinal damages; the GI hormones GAS, SS, VIP, CCK changed were associated with gastro and small intestine damage in reserpine-treated rats.

Acknowledgements. The authors deeply appreciate the members of CAU and BUA TCVM Teaching and Research team for giving considerable help by means of suggestion, comments and criticism.

References

1. Little, T.J., Feltrin, K.L., Horowitz, M., Smout, A.J., Rades, T., Meyer, J.H., et al.: Dose-related effects of lauric acid on antropyloroduodenal motility, gastrointestinal hormone release, appetite, and energy intake in healthy men. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 289(4), R1090–R1098 (2005)
2. Thomas, R.P., Hellmich, M.R., Townsend Jr., C.M., Evers, B.M.: Role of gastrointestinal hormones in the proliferation of normal and neoplastic tissues. *Endocr. Rev.* 24(5), 571–599 (2003)
3. Seva, C., Dickinson, C.J., Yamada, T.: Growth-promoting effects of glycine-extended progastrin. *Science* 265(5170), 410–412 (1994)
4. Ancha, H., Ojeas, H., Tedesco, D., Ward, A., Harty, R.F.: Somatostatin-induced gastric protection against ethanol: involvement of nitric oxide and effects on gastric mucosal blood flow. *Regul. Pept.* 110(2), 107–113 (2003)
5. Forster, E.R., Green, T., Elliot, M., Bremner, A., Dockray, G.J.: Gastric emptying in rats: role of afferent neurons and cholecystokinin. *Am. J. Physiol.* 258(4 Pt 1), G552–G556 (1990)
6. Cantarella, G., Martinez, G., Di Benedetto, G., Loreto, C., Musumeci, G., Prato, A., et al.: Protective effects of amylin on reserpine-induced gastric damage in the rat. *Pharmacol. Res.* 56(1), 27–34 (2007)
7. Singh, I.: On argyrophile and argentaffin reactions in individual granules of enterochromaffin cells of the Human gastro-intestinal tract. *J. Anat.* 98, 497–500 (1964)
8. Tanemura, A., Yano, T., Tamaki, H., Sanda, T., Ohashi, N., Ishihara, A., et al.: Ectopic pancreas in the minor duodenal papilla presenting as upper-GI bleeding. *Gastrointest Endosc.* 62(2), 324–326 (2005)
9. Anlauf, M., Perren, A., Henopp, T., Rudolf, T., Garbrecht, N., Schmitt, A., et al.: Allelic deletion of the MEN1 gene in duodenal gastrin and somatostatin cell neoplasms and their precursor lesions. *Gut.* 56(5), 637–644 (2007)
10. Yao, Y.L., Xu, B., Zhang, W.D., Song, Y.G.: Gastrin, somatostatin, and experimental disturbance of the gastrointestinal tract in rats. *World J. Gastroenterol.* 7(3), 399–402 (2001)
11. Karmeli, F., Eliakim, R., Okon, E., Rachmlewitz, D.: Role of vasoactive intestinal peptide (VIP) in pathogenesis of ethanol-induced gastric mucosal damage in rats. *Dig. Dis. Sci.* 38(7), 1210–1219 (1993)
12. Giralt, M., Vergara, P.: Inhibition by CCK of ascending contraction elicited by mucosal stimulation in the duodenum of the rat. *Neurogastroenterol. Motil.* 12(2), 173–180 (2000)

Heat Stress Induced Apoptosis in Rat Intestinal Epithelial Cell Line-6 (IEC-6)

Kaijun Guo^{1,*}, Weili Luan^{1,*}, Huichuan Wang¹, Jin Yu², Ning Wang¹, Guilin Cheng¹, and Fenghua Liu^{1,**}

¹ College of Animal Science and Technology, Beijing University of Agriculture, Beijing 102206, P.R. China

² College of Veterinary Medicine, China Agricultural University, Beijing 100193, P.R. China

liufenghua1209@126.com

Abstract. The aim of this paper is to study the effect of heat stress on the apoptosis of rat intestinal epithelial cell line-6 (IEC-6), using fluorescent dyeing technology to observe its morphological changes, flow cytometry to detect the apoptosis cells rate, RT-PCR to measure the apoptosis relevant genes. The results showed that there were evident morphological changes in the IEC-6 treated by the heat stress (42 °C 4h) and the cell membrane permeability was significantly enhanced. The apoptosis cells rate of early and late stages were significantly increased ($p < 0.05$). RT-PCR measurement revealed that the apoptosis promoting gene BAX expression was were significantly increased ($p < 0.05$) and the apoptosis inhibiting gene bcl-2 expression was were significantly decreased ($p < 0.05$). Meanwhile, the caspase-3 gene expression was were significantly increased ($p < 0.05$). This preliminary result showed the heat stress can induce the apoptosis of IEC-6.

Keywords: Heat stress, IEC-6, apoptosis, bax, bcl-2, caspase-3.

Apoptosis is a cell positive death process, and closely related with the cellular growth, aging, and maintaining a stable inner environment. Apoptosis plays an important role in body development, pathology and disease treatment. In serious cases, it can cause shock and acidosis. Given the important role of apoptosis, much research has been done concerning its trigger factors and mechanism. The studies revealed that many factors, such as high temperature, hypoxia, ray, oxidative stress etc., can cause the cells apoptosis. High temperature induced in vitro cell apoptosis [1, 2]. Our previous research indicated that the heat stress can cause intestinal tissue injury [3, 4], impair the antioxidant function of IEC-6 cells [5], influence the generation of reactive oxygen species, and consequently change the mitochondrial membrane permeability and affect the normal cell function. The aim of this paper is to study the effect of heat stress on the apoptosis of rat intestinal epithelial cell line-6 (IEC-6), using fluorescent dyeing technology to observe its morphological changes, flow cytometry to detect the apoptosis cells rate, RT-PCR to measure the apoptosis relevant genes.

* These authors contributed equally to this work.

** Corresponding author.

1 Materials and Methods

1.1 Materials

1.1.1 Cell Line

IEC-6 (CRL21592), purchased from China Union Medical University of science and technology cell preservation chamber.

1.1.2 Drugs

DMEM Culture medium, HyClone, U.S. Lot. NTL0126; fetal bovine serum, Lot. 060106; trypsin - EDTA solution, GIBCO, U.S. Lot. 495139; phosphate buffer solution (D-PBS), GIBCO, U.S. Lot. 8108037; penicillin sodium for injection, Chinese chino-e Pharmaceutical Co., Ltd, Lot. 06029468; streptomycin sulfate for injection, North China Pharmaceutical Corporation, Lot. 040507; DMSO, Kehaoze Biologic Technology Co., Ltd, Beijing, China. Amresco0231 packaging; ethidium bromide, Kehaoze Biologic Technology Co., Ltd, Beijing, China. Amresco0492 packaging; acridine orange, sigma company, Lot. A6014. Annexin V-PI keygen batch; apoptosis detection kit.

1.1.3

Main instruments single double-sided purification worktable, China Suzhou purification equipment Co., Ltd; carbon dioxide incubator, SANYO, Japan. MCO-18AIC.

1.2 Cell Treatment

IEC-6 culture: cell culture medium, including 88.3%DMEM medium + 10% ultra-filtrated fetal bovine serum + 0.7% insulin + 1% penicillin and streptomycin was stored at 4 for use. When the fusion rate of IEC-6 more than 95%, IEC-6 cells were digested with 0.25% trypsin – EDTA, then added DMEM complete culture solution. The inverted microscope was used to count the cell number and cell viability. Subsequently, the cell concentration was adjusted to 1×10^5 cell / mL, 100 μ L / well suspension was added into 96 well cell culture plate. The plate was placed in 5% CO₂ culture box. Heat stress conditions were set as 42 °C, 4h.

1.3 Fluorescent Staining

Clean cover glass slices were put in 6 well cell culture clusters after sterilization. Passage cells were added and cultured at 37 °C, 5%CO₂. When cells fusion reached 80% of the bottom, cover glass slices were removed and washed three times with PBS. The side with cells was pressured in the 1% dye solution made of 1: 1 acridine orange (AO) and ethidium bromide (EB) and stained for 10 min under dark condition. Photos were taken using fluorescent microscope camera. AO showed green under blue excitation light, while EB displayed red using green excitation light. The photos was processed using image overlaying techniques.

1.4 Flow Cytometric Detection of Cell Apoptosis

After high temperature stimulation, cells was digested using trypsin and collected. Cells were put into 1.5mL centrifugal tube and centrifuged at 2000 rpm for 5 min, then

washed 3 times with PBS and collected 5×10^5 cells. 500 μ L Binding Buffer cell suspension was added into the cells. Then 5 μ L Annexin V-FITC was added and mixed. Subsequently 5 μ L Propidium Iodide was added and well mixed. The solution was kept at room temperature for 15 min under condition for reaction. Cells without apoptosis induced treatment were used as control for fluorescence compensation regulation. Flow cytometry was set as the excitation wavelength $Ex = 488$ nm and the emission wavelength of $Em = 530$ nm. Cell apoptosis was analyzed using Multicycle software and the early and late apoptosis rate (%) were calculated.

1.5 RT-PCR Detection of Apoptosis Related Genes

The treated cells, after abandoned the culture medium, was rinsed twice with PBS. Total RNA was extracted using Trizol (Invitrogen, Life Technologies) one step extraction method. 5 μ L total RNA was dissolved in 95 μ L DEPC treated water. Total RNA concentration was measured using protein nucleic acid analyzer (Eppendorf Biophotometer). The A260 / A280 value was used to detected the purity. Taking about 2 μ g RNA, cDNA was synthesized according to the reverse kit manual and stored at - 20 °C.

Table 1. Primer sequences of apoptosis related genes

| Names | | Oligonucleotide sequences | Length |
|----------------|---------|---------------------------------|--------|
| Bax | Forward | 5'- CCGAGAGGTCCTTCTTCCGTGTG -3' | 319bp |
| | Reverse | 5'- GCCTCAGCCCATCTTCTTCCA -3' | |
| Bcl-2 | Forward | 5'- CAAGCCGGGAGAACAGGGTA -3' | 450bp |
| | Reverse | 5'- CCCACCGAACTCAAAGAAGGC -3' | |
| Caspase-3 | Forward | 5'-CGGAAAGCAAGATCCATGAG-3' | 216bp |
| | Reverse | 5'-TAAGACGCCTCAGTTCACAG-3' | |
| β -actin | Forward | 5'- CTTTCTACAATGAGCTGCGTG -3' | 218bp |
| | Reverse | 5'- TCATGAGGTAGTCTGTCAGG -3' | |

Amplification conditions were 94 °C for 5min, 34 cycles (each cycle included 94 °C 30s, 58 °C 30s and 72°C 40s). The last 72°C was 8min. Each sample is repeated at least 3 independent RT-PCR processes. Amplified products were expressed as relative value using agarose gel (1.2%) the band intensity corrected by the β -actin intensity.

1.6 Statistical Analysis

The results were expressed as mean \pm SD. The data were analyzed by LSD method of SPSS10.0.

2 Results

2.1 Morphological Changes of Heat Stressed Cells

Under the ordinary microscope (Fig.1), IEC-6 cells in control group (37 °C) were polygonal, pavement-like monolayer. Cells adhered tightly the walls with clear boundary, no overlapping each other. The nucleus is large, round or pebble shape, with

abundant cytoplasm. Nuclear chromatin was sparse and bright, with 1-2 nucleolus. Cells were connected to each other, showing exuberant proliferation activity. Morphology of cells treated with high temperature (42°C, 4 h) has the certain change. large amount of cells shrank and aggregated in round shape. Their volume were reduced, so that the cells shed, floating in the cell culture supernatant. The residual cells adhered the walls with disordered distribution. The cell morphology changed from paving stone shape to irregular shape, with dispersed growth and blurred edge.

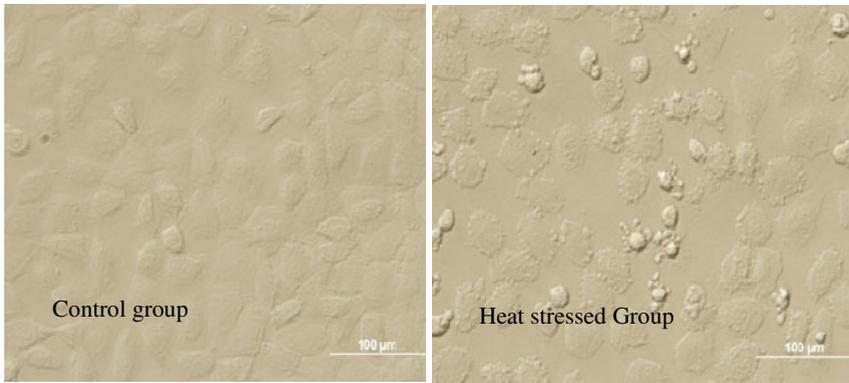


Fig. 1. Morphological changes of heat stressed cells

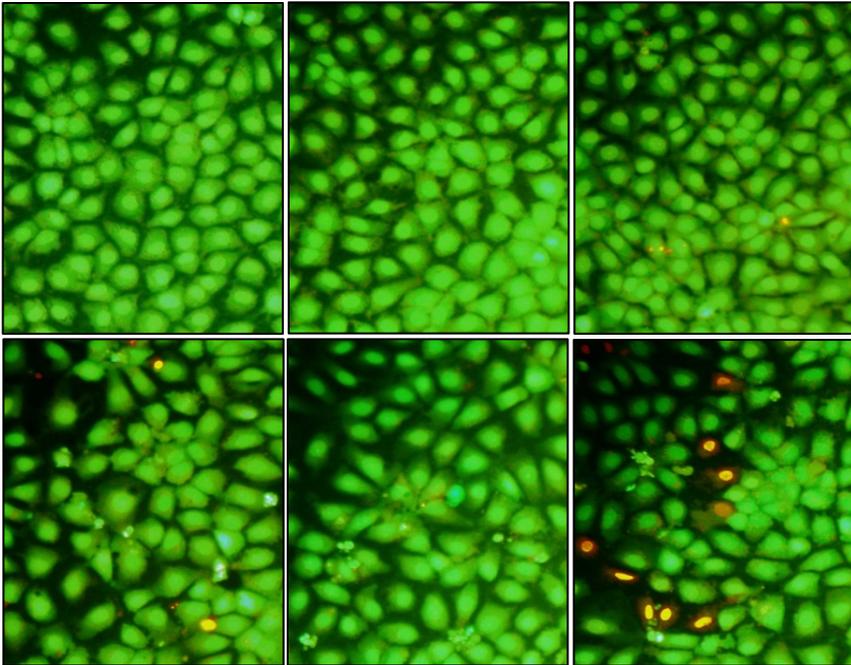


Fig. 2. Fluorescence double staining showed apoptotic cells

In cells double stained by AO and EB (Fig. 2), fluorescent dye AO can pass through an intact cell membrane, embedding nuclear DNA, to emit green fluorescence, while EB can pass through the damaged cell membrane, embedding DNA, to emit red light. After the image superimposed, cells in control group showed a uniform green fluorescence as normal cell. After heat stress, cells showed different fluorescence degree, enhancing staining and making fluorescence brighter, and appearing round, pyknosis, crumb structure. Several stages of cells could be observed: early apoptotic cells appeared green pyknosis, late apoptotic cells lumpy orange, and the dead cells red.

2.2 Apoptotic Rate Changes of IEC-6 after Heat Stress

Table 2 shows that apoptotic rate both in early stage and in late stage of heat stressed cells were higher than those in control group.

Table 2. The influence of heat stress on apoptotic rate of IEC-6

| Group | Normal cell rate | Apoptotic rate in early stage | Apoptotic rate in late stage |
|-------------|------------------|-------------------------------|------------------------------|
| Control | 94.03±0.87 | 5.13±0.73 ^a | 0.60±0.20 ^a |
| Heat stress | 80.53±0.70 | 17.80±0.79 ^b | 1.33±0.21 ^b |

2.3 Changes of Apoptosis Related Gene Expression after Heat Stress

The results showed compared with the control group, the pro-apoptosis gene BAX expression in heat stressed group was significantly increased ($p < 0.05$) and the antiapoptosis gene bcl-2 expression was significantly decreased ($p < 0.05$). Meanwhile, the caspase-3 gene expression in heat stressed group was significantly increased ($p < 0.05$) compared with the control group (Fig. 3).

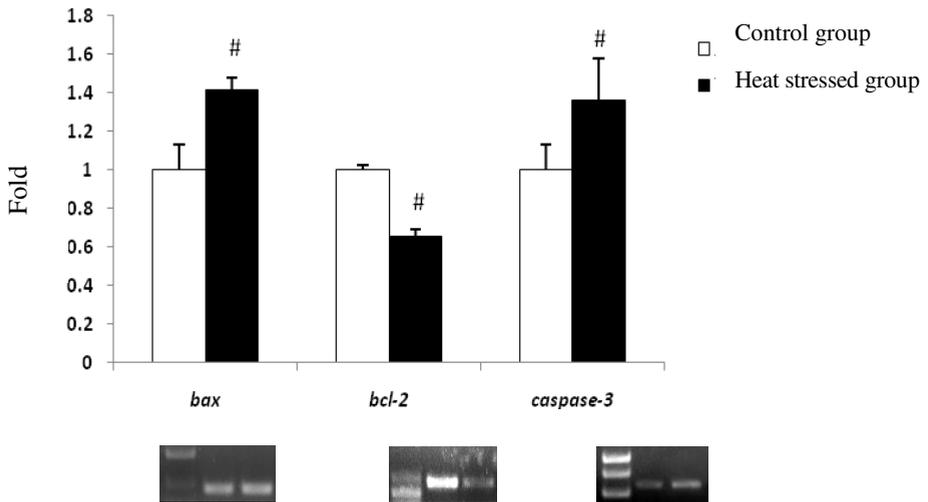


Fig. 3. Effects of heat stress on apoptosis related gene expression

3 Discussions

IEC-6 is stem cells of rat intestinal crypt, which is similar with other crypt stem cells *in vivo* from the histological point of view. It not only maintains undifferentiated properties of small intestinal epithelial stem cells, but also can be differentiated into mature intestinal epithelial cells in certain circumstances [6]. The IEC-6 heat stress model established in the previous studied has important implications and provides a suitable *in vitro* model for the study of intestinal epithelial cell growth, differentiation, cell proliferation, and cell cycle [7, 8].

Previous studies showed that, heat stress can induce *in vitro* apoptosis. High temperature can make intracellular reactive oxygen species increased and change mitochondrial permeability [9]. After stimulated by apoptotic signal, cytoplasm, nuclei and cells membrane will be subject to a series of biological chemical and physical changes [10]. In this paper, morphological changes had been observed in IEC-6 cells treated by 42 °C for 4 h. Observed with ordinary microscope, cells grew dispersedly and the edges blurred. After the AO and EB double staining, there were yellow apoptotic cells in early and late stages and dead cells under fluorescence microscope photography. Flow cytometry measurements showed that apoptosis rate in the early stage was significantly increased in heat stressed cells compared with that in the control group. RT-PCR detection verified that compared with the control group, the heat stressed group had the pro-apoptotic gene Bax expression significantly increased and the antiapoptotic gene Bcl-2 expression significantly reduced. The terminal performing enzyme of apoptosis, caspase-3, was also significantly elevated after heat stress. All these results demonstrated that the heat stress induced apoptosis of IEC-6 cells.

Acknowledgements. This work was partially supported by grants from the National Twelve-Five Technological Supported Plan of China (No. 2011BAD34B01), Ministry of Agriculture, Public Service Sectors Agriculture Research Projects (No. 201003060-9/10), the Funding Project for Academic Human Resources Development in Institutions of Higher Learning under the Jurisdiction of Beijing Municipality, and the Beijing Key Laboratory of Traditional Chinese Veterinary Medicine.

References

1. Li, Z.H., Wang, L.: Influence of Heat Stress on Lymphocyte Apoptosis of Holstein Cow. Hubei Agricultural Sciences 48(12), 35–42 (2009)
2. Zhang, S.X., Hua, R.H., Bao, E.D.: Effect of Heat Stress on Apoptosis in Chicken Spleen and Bursa Cells and Their Regulation. China Agricultural Science 38(1), 46–51 (2005)
3. Liu, F.H., Yin, J.D., Du, M., Yan, P.S., Xu, J.Q., Zhu, X.Y., Yu, J.: Heat Stress induced Damage to Porcine Small Intestinal Epithelium associated with Down-Regulation of Epithelial Growth Factor (EGF) Signaling. Journal of Animal Science 6(87), 1941–1949 (2009)
4. Yu, J., Yin, P., Liu, F.H., Zhu, X.Y., Cheng, G.L., Lu, A., Luan, W.L., Xu, J.Q.: Effect of heat stress on porcine small intestine: a study of morphological and gene expression. Comparative Biochemistry and Physiology, Part A 156, 119–128 (2010)

5. Song, X.Z., Wang, Z.H., Mao, S., Liu, F.H., Wang, T.: Effect of traditional Chinese medicine additives on serum antioxidant status in piglets under high temperature. *Animal Husbandry & Veterinary Medicine* 40(5), 5–8 (2008)
6. Quaroni, A., Wands, J., Trelstad, R.L., et al.: Epithelioid cell cultures from rat small intestine: characterization by morphologic and immunologic criteria. *J. Cell Biol.* 80, 248–265 (1979)
7. Thomas, C., Oates, P.S.: IEC-6 Cells Are an Appropriate Model of Intestinal Iron Absorption in Rats. *J. Nutr.* (32), 680–687 (2002)
8. Raveendran, N.N., Silver, K., Freeman, L., et al.: Drug-induced alterations to gene and protein expression in IEC-6 cells suggest a role for calpains in the GI toxicity of non-steroidal anti-inflammatory agents (2008), doi:1124-1277
9. Cheng, J.Z., Sharma, R., Yang, Y.S., et al.: Accelerated Metabolism and Exdusion of 4-Hydroxynonenal through Induction of RLIP76 and Hgst5. 8 Is an Early AdaPtive Response of Cells to Heat and Oxidative Stress. *The Journal of Biological Chemistry* 276(44), 41213–41223 (2001)
10. Arnoult, D., et al.: The murine cytomegalovirus cell death suppressor m38.5 binds Bax and blocks Bax-mediated mitochondrial outer membrane permeabilization. *Apoptosis* 13(9), 1100–1110 (2008)

Evaluation of Chinese New Manufacturing Development Based on Factor Analysis

Wei Sun

Nanjing University of Information Science and Technology, Nanjing, China
Sunwei16@nuist.edu.cn

Abstract. By selection of a series of indicators, this paper study evaluation of Chinese new manufacturing development. Factor analysis method is used to evaluate the status of the development of new manufacturing. The theoretical framework and practical application of the evaluation system is studied in detail in the paper. Empirical analysis is used in the paper based on statistics of manufacturing industry from 2003 to 2008. The results show that manufacturing industry is sustainable development. Research results are consistent with the actual situation. It shows that this study is scientific and objective.

Keywords: Evaluation, manufacturing, factor analysis.

1 Introduction

New manufacturing is industry of high technology content, good economic returns, low resources consumption and little environmental pollution. Chinese manufacturing industry has different levels of development each year. The new evaluation index system of the manufacturing industry can be established from three aspects. The three aspects are economic creativity, technological innovation and environmental protection. Evaluation of the economic creativity includes 7 indicators. These indicators are total value, employed population, the proportion of manufacturing value added in GDP, foreign trade dependence degree, labor productivity, total profit and manufacturing efficiency index. Evaluation of the technological innovation includes 5 indicators. These indicators are R&Dfunding, R&D staff working full-time equivalent, digestion and absorption of expenditures, patent applications per capita and output rate of new products. Evaluation of the environmental protection includes 8 indicators. These indicators are waste water discharge, wastewater discharge per unit of output, exhaust emissions, emissions per unit output value, solid waste generation, solid waste generation per unit of output, total energy consumption, and energy consumption per unit output value. New manufacturing individual capacity of from 2003 to 2008 can be evaluated from these three different aspects.

2 Evaluation of Economic Creativity

Chinese manufacturing industry data from 2003 to 2008 can be used to do statistical analysis. According to multivariate statistical analysis, confidence coefficient of

processing results is 100% through the use of mathematical software SPSS12.0. The first two principal components of the cumulative variance contribution rate is 94.45%. Processing accuracy can meet the requirements.

Table 1. The correlation matrix of principal components and indicators

| indicators | No. | Principal component 1 | Principal component 2 |
|--|-----|-----------------------|-----------------------|
| total value (billion) | A1 | 0.972 | 0.233 |
| Employed population (million) | A2 | 0.988 | 0.137 |
| the proportion of manufacturing value added in GDP | A3 | 0.939 | 0.211 |
| Foreign trade dependence degree | A4 | -0.951 | -0.109 |
| Labor Productivity (yuan / person. years) | A5 | 0.493 | 0.779 |
| Total profit (million) | A6 | 0.908 | 0.406 |
| Manufacturing efficiency index (%) | A7 | 0.033 | 0.968 |

According to the principal component coefficients and related indicators of factors. It is found that the coefficient of A1, A2, A4 in the first principal component and the coefficient of A5, A7 in the second principal component are greater than the coefficient of other indicators in the column. It shows that the first principle component is the scale of the situation in the manufacturing and the second principle component is the economic level in the manufacturing. According to two principal components, sort results summarized as the following table

Table 2. Evaluation results of manufacturing's economic creativity

| year | F1 | F2 | complex | Ranking |
|------|----------|----------|----------|---------|
| 2008 | 1.45896 | 0.01257 | 0.998267 | 1 |
| 2007 | 0.53325 | 1.32098 | 0.710471 | 2 |
| 2006 | 0.28363 | -0.36119 | 0.0986 | 3 |
| 2005 | -0.03745 | -1.5765 | -0.43943 | 4 |
| 2004 | -1.12255 | 0.77985 | -0.5608 | 5 |
| 2003 | -1.11584 | -0.17572 | -0.8071 | 6 |

It shows that Chinese manufacturing industry is growing from 2003 to 2008 according to overall condition of economic creativity.

3 Evaluation of Scientific and Technological Innovation

Based on indicators from B5 to B7, cumulative variance contribution rate of the first five principal components is 96.709%. Accuracy is high. Weight index is balanced.

Table 3. Correlation coefficient matrix of principal components and indicators

| indicators | No. | Principal component 1 | Principal component 2 |
|--|-----|-----------------------|-----------------------|
| R&DFunding (million) | B1 | 0.987 | -0.023 |
| R&D Staff working full-time equivalent (million / year) | B2 | 0.961 | -0.174 |
| Digestion and absorption of Expenditures (million) | B3 | 0.965 | 0.092 |
| Patent applications per capita (Pieces / million People) | B4 | 0.985 | 0.027 |
| Output rate of new Products | B5 | -0.014 | 0.998 |

According to the correlation coefficient of principal components and indicators, the coefficient of B1、B4 is higher. It show that investment in research has great impact on technology innovation. The coefficient of B5 is higher in the second principal component. It is the main representative of scientific and technological innovation reflected in the proportion of new products. According to the above five indicators, scientific and technological innovation in all sample areas on the comprehensive summary of the two principal components are shown in the following table.

Table 4. Comprehensive evaluation of manufacturing's technological innovation

| Year | F1 | F2 | complex | Ranking |
|------|----------|----------|----------|---------|
| 2008 | 1.21009 | 0.02265 | 0.924359 | 1 |
| 2007 | 1.02129 | 0.07002 | 0.790681 | 2 |
| 2006 | 0.20985 | -0.2593 | 0.105788 | 3 |
| 2004 | -0.80054 | 1.86505 | -0.22218 | 4 |
| 2005 | -0.33982 | -0.79172 | -0.42222 | 5 |
| 2003 | -1.30087 | -0.9067 | -1.17643 | 6 |

It shows that manufacturing's technological innovation is sustainable development from 2003 to 2004. The development of manufacturing technology innovation is high-speed in 2004. The situation in 2004 is slightly higher than the development in 2005. Manufacturing technology innovation is in the continuous development from 2005 to 2008.

4 Evaluation of Environmental Resources Protection

To C1 ~ C8-based indicators, The results before the two principal components of the cumulative variance contribution rate of 98.562%. Processing precision can meet the requirements.

Table 5. Principal Components and indicators of the correlation matrix

| indicators | No. | Principal component 1 | Principal component 2 |
|--|-----|-----------------------|-----------------------|
| Waste Water Discharge (tons) | C1 | -0.824 | 0.537 |
| Wastewater discharge per unit of output (tons / million) | C2 | 0.760 | -0.642 |
| Exhaust emissions (million standard cubic meters) | C3 | -0.926 | 0.353 |
| Emissions per unit output value (m / yuan) | C4 | 0.505 | -0.851 |
| Solid waste generation (tons) | C5 | -0.336 | 0.937 |
| Solid waste generation per unit of output (tons / million) | C6 | 0.888 | -0.450 |
| Total energy consumption (million tons of standard coal) | C7 | -0.748 | 0.658 |
| Energy consumption per unit output value (ton / million) | C8 | 0.917 | -0.387 |

According to the correlation coefficient of principal components and indicators, the coefficient of C3、C8 is higher in the first Principal component. It show that the principal component represent the emissions per unit of output energy situation and the degree of influence on the capacity for environmental protection.

The coefficient of C4、C5 is higher in the second principal component. The second principal component represent emissions and solid waste impact on the environment. According to the eight indicators, sort results summarized as the following table.

Table 6. Manufacturing Comprehensive Evaluation of environmental and resource protection

| year | F1 | F2 | complex | Ranking |
|------|----------|----------|----------|---------|
| 2007 | 1.04616 | -0.13985 | 0.554862 | 1 |
| 2008 | 1.19786 | -0.4238 | 0.529441 | 2 |
| 2006 | 0.0327 | -0.36586 | -0.12782 | 3 |
| 2005 | -0.66568 | -0.61676 | -0.63646 | 4 |
| 2003 | -0.20187 | 2.01634 | 0.691777 | 5 |
| 2004 | -1.40917 | -0.47007 | -1.0118 | 6 |

This shows that from the overall analysis, the manufacturing sector continue to strengthen in protection of environmental resources from 2004 to 2007. Manufacturing protection of environmental resources develop well in 2003. The development of 2003 is better than the situation in 2004. In 2008 the development of the manufacturing sector is good in protection of environmental resources in 2008. The development of 2008 is slightly higher than the situation of 2004.

5 Manufacturing New Degree Evaluation

Through the regional manufacturing sector evaluation of new extent, we can grasp the reality and development potential of manufacturing industries. Integration of all 20 indicators calculate with the principal components. The result shows that the correlation of 20 indicators entirely get through testing. The first two principal components of the cumulative variance contribution rate is 97.066%. Processing accuracy can meet the requirements.

Table 7. Principal components and indicators of the correlation matrix

| indicators | Principal component 1 | Principal component 2 | Principal component 3 |
|------------|-----------------------|-----------------------|-----------------------|
| A1 | 0.896 | 0.411 | 0.144 |
| A2 | 0.874 | 0.460 | 0.062 |
| A3 | 0.798 | 0.579 | 0.129 |
| A4 | -0.969 | -0.150 | 0.006 |
| A5 | 0.556 | 0.084 | 0.751 |
| A6 | 0.881 | 0.355 | 0.302 |
| A7 | 0.202 | -0.064 | 0.909 |
| B1 | 0.888 | 0.413 | 0.170 |
| B2 | 0.946 | 0.293 | 0.070 |
| B3 | 0.723 | 0.639 | 0.187 |
| B4 | 0.827 | 0.463 | 0.258 |
| B5 | -0.398 | 0.399 | 0.779 |
| C1 | -0.854 | -0.455 | -0.159 |
| C2 | 0.814 | 0.571 | 0.056 |
| C3 | -0.939 | -0.263 | -0.202 |
| C4 | 0.586 | 0.803 | -0.003 |
| C5 | -0.408 | -0.899 | -0.140 |
| C6 | 0.921 | 0.366 | 0.094 |
| C7 | -0.796 | -0.585 | -0.146 |
| C8 | 0.952 | 0.304 | 0.017 |

According to the principal component coefficients and related indicators of factors. It is found that the coefficient of A4, B2, C3, C6, C8 in the first principal component are greater than the coefficient of other indicators in the column. It shows that the first principle component represent resources capacity for environmental protection and the situation of foreign attachment. the coefficient of A5,C4,C5 in the second principal component are greater than the coefficient of other indicators in the column. It shows that the second principle component mainly reflects the emissions of waste and labor productivity. The third principal component is representative of A7. It is the indicator of the manufacturing effectiveness. The cumulative contribution of the above three principal components is more than 97%. According to three principal components, sort results summarized as the following table.

Table 8. Manufacturing "New" Comprehensive Evaluation of the degree

| year | F1 | F2 | F3 | complex | Ranking |
|------|----------|----------|----------|----------|---------|
| 2008 | 1.27435 | 0.31812 | 0.07621 | 0.877133 | 1 |
| 2007 | 0.91517 | 0.00493 | 0.95905 | 0.686989 | 2 |
| 2006 | 0.10213 | 0.38742 | -0.36419 | 0.108284 | 3 |
| 2005 | -0.45606 | 0.71404 | -1.56724 | -0.30961 | 4 |
| 2004 | -1.48142 | 0.55743 | 1.18428 | -0.6575 | 5 |
| 2003 | -0.35418 | -1.98194 | -0.28812 | -0.7053 | 6 |

This shows that from the overall ranking, Chinese manufacturing industry new development is getting better from 2003 to 2008. The overall manufacturing sector is sustainable development.

The work is partly supported by the Public Sector (Meteorology) Funding Research Project (GYHY200806017), the National Natural Science Foundation of China (No. 70873063), the major funds project of philosophy and social sciences key research base of Jiangsu Province (2010JDXM028), the major funds project under the Ministry of Education of key research base of humanities and sciences (08JJD630006), the National Natural Science Foundation of China (No. 70973057), the National Natural Science Foundation of China (No.70901043), the funds project under the Ministry of Education of the PRC for young people who are devoted to the researches of humanities and social sciences (09YJC630130), the funds project of Philosophy and Social Sciences of Jiangsu province (09SJB630044), the funds project under the Ministry of Education (10YJC790284).

References

1. Hu, S.-l.: Fuzzy Mathematics and Its Applications. Sichuan University Press, Chengdu (1994)
2. Yanglunbiao, Gaoyingyi: Principle and Application of Fuzzy Mathematics. Publishing House of South China University of Technology (2001)
3. Li, L., Du, Z.: A Research Report on the Development of Chinese Manufacturing Industry. Science Press (2008)
4. Liu, J., Li, J.: Summary of abroad research for manufacturing servitization. Economist (3) (2007)

Research of Evaluation of New Computer Complex Talents

Wei Sun

Nanjing University of Information Science and Technology, Nanjing, China
Sunwei16@nuist.edu.cn

Abstract. The paper analyzes the meaning of new computer complex talents. Evaluation indicator of new computer complex talents is established in the article. Fuzzy mathematics is used to build the evaluation model of new computer complex talents in the paper. The case is analyzed in the paper. It shows that the evaluation system of new computer complex talents is scientific and objective.

Keywords: computer, complex talents, evaluation.

1 The Meaning of New Computer Complex Talents

New computer complex talent is administered by a multiple subject background, taking control of new computer knowledge, and effective integration of comprehensive use of creative talents. The essence of new computer complex talent is breaking the professional boundaries between the closely guarded.

2 Evaluation Indicator about New Computer Complex Talents

There are three primary indicators in the evaluation system of new computer complex talents. The first indicator is the level of knowledge. The second indicator is ability condition. The third indicator is comprehensive quality.

3 The Level of Knowledge

There are four secondary indicators in the evaluation of the level of knowledge. The first indicator is the mastery of new computer knowledge. The second indicator is the mastery of the other knowledge. The third indicator is the effective level of integration on knowledge. The fourth indicator is the comprehensive understanding of the knowledge.

4 Ability Condition

There are four secondary indicators in the evaluation of ability condition. The first indicator is computer level. The second indicator is language ability. The third indicator is the actual ability to use. The fourth indicator is self-learning ability.

5 Comprehensive Quality

There are five secondary indicators in the evaluation of comprehensive quality. The first indicator is the moral quality. The second indicator is the physical quality. The third indicator is the intellectual quality. The fourth indicator is the psychological quality. The fifth indicator is the professional quality. Therefore, we can construct the following table of the evaluation index system.

Table 1. New computer complex talent evaluation index system

| Primary indicators | Secondary indicators |
|--------------------|--|
| Level of knowledge | The mastery of new computer knowledge |
| | The mastery of the other knowledge |
| | The effective level of integration on knowledge |
| | The comprehensive understanding of the knowledge |
| Ability condition | Computer level |
| | Language ability |
| | The actual ability to use |
| | Self-learning ability |
| Overall quality | The moral quality |
| | The intellectual quality |
| | The physical quality |
| | The psychological quality |
| | The professional quality |

6 Evaluation Methods of New Computer Complex Talents

Fuzzy comprehensive appraisal can be used to evaluate new computer complex talents. Steps of evaluation are as follows:

The set of factors is determined. We can suppose the factors set U . It is $U = \{X_1, X_2, \dots, X_n\}$. The factors $U = \{X_1, X_2, \dots, X_n\}$ can be divided into several groups $u = \{u_1, u_2, \dots, u_n\}$. The first level factor set is $u = \{u_1, u_2, \dots, u_n\}$. The second level factor set is $U = \{X_1, X_2, \dots, X_n\}$. X_i is refer to the fact affect the object. X_i can be divided into several groups u_i . We can suppose that the factors set of new computer complex talents evaluation is $U = \{U_1, U_2, U_3\}$. Then, $U_1 = \{u_{11}, u_{12}, u_{13}, u_{14}\}$, $U_2 = \{u_{21}, u_{22}, u_{23}, u_{24}\}$, $U_3 = \{u_{31}, u_{32}, u_{33}, u_{34}, u_{35}\}$.



Weight of each indicator is determined. Different indicator should be given different weights depending on the importance to the object in the established system of evaluation of new computer complex talents. Because the impact is difficult to quantify, the expert analysis method (Delphi method) is used to determine the weight of each indicator in the index system. The assumption that the expert analysis in accordance with the calculated weight of the indicators are as follows:

The weight of U indicator $W = \{W_1, W_2, W_3\}$. The weight of U_1 indicator $W_1 = \{w_{41}, w_{42}, w_{43}, w_{44}\}$. The weight of U_2 indicator $W_2 = \{w_{21}, w_{22}, w_{23}, w_{24}\}$. The weight of U_3 indicator $W_3 = \{w_{31}, w_{32}, w_{33}, w_{34}, w_{35}\}$.

Evaluation rating scale is determined. We can rate the object into different class which named set of appraisal $V : V = \{v_1, v_2, \dots, v_m\}$. Assuming the set of asset appraisal of patent is $V = \{v_1, v_2, v_3, v_4, v_5\} = \{\text{Highest, higher, medium, lower, lowest}\}$

The first fuzzy comprehensive appraisal of single index is performed. Appraisal of single the first level indicator is a separate appraisal to single indicator u_{ij} of subset U_i . In the evaluation of new computer complex talents, each of the second level indicator of the first level indicator U_1 buildup the appraisal matrix R_1 . And the result of U_1 fuzzy comprehensive appraisal is B_1 . Other results of (U_2, U_3) fuzzy comprehensive appraisal are B_2, B_3 by using the same method.

The second fuzzy comprehensive appraisal is performed. The result of fuzzy comprehensive appraisal U_i buildup the higher level U appraisal matrix $R = (B_1, B_2, B_3)^T$. By use of fuzzy synthetic method, the result of U Fuzzy comprehensive appraisal is $B = W \bullet R = (w_1, w_2, w_3) \bullet (B_1, B_2, B_3)^T$.

7 Case Analysis

The example of evaluation of new computer complex talents is as follow. The set of facts is $U = \{U_1, U_2, U_3\}$. U_1 is state of knowledge. U_2 is capacity status. U_3 is comprehensive quality. We can suppose $U_1 = \{u_{11}, u_{12}, u_{13}, u_{14}\}$; u_{11} is mastery of new computer knowledge. u_{12} is mastery of the other knowledge. u_{13} is effective level of integration on knowledge. u_{14} is comprehensive understanding of the knowledge. We can suppose $U_2 = \{u_{21}, u_{22}, u_{23}, u_{24}\}$; u_{21} is computer level. u_{22} is

language ability. u_{23} is the actual ability to use. u_{24} is self-learning ability. We can suppose $U_3 = \{u_{31}, u_{32}, u_{33}, u_{34}, u_{35}\}$; u_{31} is moral quality. u_{32} is intellectual quality. u_{33} is physical quality. u_{34} is psychological quality. u_{35} is Professional quality.

Using analytical hierarchy process, the calculated weight of the indicators is as follows:

Weight of the U indicator is $W = \{W_1, W_2, W_3\} = \{0.36, 0.21, 0.43\}$. Weight of the U_1 indicator is $W_1 = \{w_{11}, w_{12}, w_{13}, w_{14}\} = \{0.22, 0.19, 0.15, 0.44\}$, Weight of the U_2 indicator is $W_2 = \{w_{21}, w_{22}, w_{23}, w_{24}\} = \{0.31, 0.22, 0.14, 0.33\}$, Weight of the U_3 indicator is $W_3 = \{w_{31}, w_{32}, w_{33}, w_{34}, w_{35}\} = \{0.19, 0.11, 0.15, 0.21, 0.34\}$, The set of asset appraisal of patent is $V = \{v_1, v_2, v_3, v_4, v_5\} = \{\text{Highest, higher, medium, lower, lowest}\}$. Experts are invited to rate each indicator u_{ij} into one class. The subjection rate of appraisal set V comes from the agreement indicator rate of experts.

In U_1

$$R_1 = \begin{pmatrix} 0.27 & 0.29 & 0.17 & 0.12 & 0.15 \\ 0.22 & 0.38 & 0.20 & 0.11 & 0.09 \\ 0.19 & 0.28 & 0.20 & 0.16 & 0.17 \\ 0.21 & 0.39 & 0.15 & 0.11 & 0.14 \end{pmatrix}$$

In U_2

$$R_2 = \begin{pmatrix} 0.12 & 0.39 & 0.18 & 0.20 & 0.11 \\ 0.23 & 0.36 & 0.15 & 0.14 & 0.12 \\ 0.20 & 0.28 & 0.17 & 0.15 & 0.20 \\ 0.31 & 0.38 & 0.12 & 0.11 & 0.08 \end{pmatrix}$$

In U_3

$$R_3 = \begin{pmatrix} 0.24 & 0.27 & 0.18 & 0.17 & 0.14 \\ 0.22 & 0.35 & 0.21 & 0.12 & 0.10 \\ 0.18 & 0.29 & 0.21 & 0.15 & 0.17 \\ 0.26 & 0.33 & 0.12 & 0.13 & 0.16 \\ 0.17 & 0.36 & 0.20 & 0.16 & 0.11 \end{pmatrix}$$

Then comes,

$$B_1 = W_1 \circ R_1 = (0.36, 0.21, 0.43) \circ \begin{pmatrix} 0.27 & 0.29 & 0.17 & 0.12 & 0.15 \\ 0.22 & 0.38 & 0.20 & 0.11 & 0.09 \\ 0.19 & 0.28 & 0.20 & 0.16 & 0.17 \\ 0.21 & 0.39 & 0.15 & 0.11 & 0.14 \end{pmatrix}$$

$$= (0.22, 0.39, 0.19, 0.15, 0.15)$$

Similarly

$$B_2 = W_2 \circ R_2 = (0.31, 0.22, 0.14, 0.33) \circ \begin{pmatrix} 0.12 & 0.39 & 0.18 & 0.20 & 0.11 \\ 0.23 & 0.36 & 0.15 & 0.14 & 0.12 \\ 0.20 & 0.28 & 0.17 & 0.15 & 0.20 \\ 0.31 & 0.38 & 0.12 & 0.11 & 0.08 \end{pmatrix}$$

$$= (0.31, 0.33, 0.18, 0.20, 0.14)$$

$$B_3 = W_3 \circ R_3 = (0.19, 0.11, 0.15, 0.21, 0.34) \circ \begin{pmatrix} 0.24 & 0.27 & 0.18 & 0.17 & 0.14 \\ 0.22 & 0.35 & 0.21 & 0.12 & 0.10 \\ 0.18 & 0.29 & 0.21 & 0.15 & 0.17 \\ 0.26 & 0.33 & 0.12 & 0.13 & 0.16 \\ 0.17 & 0.36 & 0.20 & 0.16 & 0.11 \end{pmatrix}$$

$$= (0.21, 0.34, 0.20, 0.17, 0.16)$$

Then begin level 2 comprehensive estimate.

the F matrix from $U_1 U_2 U_3$ to V is

$$R = \begin{pmatrix} 0.22 & 0.39 & 0.19 & 0.15 & 0.15 \\ 0.31 & 0.33 & 0.18 & 0.20 & 0.14 \\ 0.21 & 0.34 & 0.20 & 0.17 & 0.16 \end{pmatrix}$$

So the second level comprehensive judges

$$B = W \circ R = (0.36, 0.21, 0.43) \circ \begin{pmatrix} 0.22 & 0.39 & 0.19 & 0.15 & 0.15 \\ 0.31 & 0.33 & 0.18 & 0.20 & 0.14 \\ 0.21 & 0.34 & 0.20 & 0.17 & 0.16 \end{pmatrix}$$

$$= (0.22, 0.36, 0.20, 0.20, 0.16)$$

Normalized : $B' = (0.19, 0.32, 0.18, 0.18, 0.13)$

According to the principle of maximal subjection, the evaluation of new computer complex talents is higher. The fuzzy comprehensive Appraisal can be an effective method in evaluation of new computer complex talents.

The work is partly supported by the Public Sector (Meteorology) Funding Research Project (GYHY200806017), the National Natural Science Foundation of China (No. 70873063), the major funds project of philosophy and social sciences key research base of Jiangsu Province (2010JDXM028), the major funds project under the Ministry of Education of key research base of humanities and sciences (08JJD630006), the National Natural Science Foundation of China (No. 70973057), the National Natural Science Foundation of China (No.70901043), the funds project under the Ministry of Education of the PRC for young people who are devoted to the researches of humanities and social sciences (09YJC630130), the funds project of Philosophy and Social Sciences of Jiangsu province (09SJB630044), the funds project under the Ministry of Education (10YJC790284).

References

1. Hezili, Shenyaping: New Training Model of Complex Talents. Chinese Higher Education Research (September 2006)
2. Yanglunbiao, Gaoyingyi: Principle and Application of Fuzzy Mathematics. Publishing House of South China University of Technology (2001)
3. Zhongqiquan: The New Curriculum Ideas and Innovation. Higher Education Press (2003)
4. Hu, S.-L.: Fuzzy Mathematics and Its Applications. Sichuan University Press, Chengdu (1994)

Optimal Iterative Learning Control for Nonlinear Discrete-Time Systems

Hong-wei Xu

College of Mechanical & Electrical Engineering, China Jiliang University, Hangzhou, China
xhw@cjlju.edu.cn

Abstract. In this paper, the problem about optimal iterative learning control of general nonlinear discrete-time systems has been studied. Based on the sufficient conditions of the existence of optimal iterative learning control in general nonlinear discrete-time systems, in view of the practical application, propose an approximate iterative algorithm, and prove the approximate iterative control restraining to the optimum control.

Keywords: Iterative Learning Control, general nonlinear discrete-time systems, Approximate Algorithm, Optimal Control.

1 Introduction

Iterative Learning Control (ILC for short) was setup in [1] in 1984 firstly, it is an effective method for the dynamic systems which have uncertainties and run repeatedly, such as industrial robots, disk motivating mechanism system. Optimal ILC is an important ILC designing method based on some criteria. In [2, 3, 4, 5], ILC of discrete-time linear systems with an optimality criterion was studied. In [6, 7, 8, 9], ILC of continuous-time linear systems with a criterion was given. In [10], optimal ILC about general nonlinear systems has been studied.

In this paper, based on the paper [10], an approximate algorithm optimal ILC about general nonlinear discrete-time systems is considered. It is proven that the approximate iterative control restraining to the optimum control.

2 Description of the Problem

Consider general nonlinear discrete-time system of the form Eq.1

$$\begin{cases} x_k(t+1) = f(x_k(t), u_k(t), t) \\ y_k(t) = g(x_k(t), u_k(t), t) \end{cases} \quad (1)$$

Where $x_k(t) \in R^n$, $y_k(t) \in R^m$, $u_k(t) \in R^r$ are, respectively, the state, output, input of the k th iteration, $t = [0, 1, 2, \dots, T]$ and $f: R^n \times R^m \rightarrow R^n$; $g: R^n \times R^m \rightarrow R^r$ are the state, output function of the k th iteration.

Considering A1, A2 are met:

A1 The partial derivatives $\partial f / \partial u, \partial g / \partial u$ are continuous and bounded;

A2 There be an unique control $u_d(t), t = [0, 1, 2, \dots, T]$, Satisfies the following Eq.2

$$\begin{cases} x_d(t+1) = f(x_d(t), u_d(t), t) \\ y_d(t) = g(x_d(t), u_d(t), t) \end{cases} \quad (2)$$

$y_d(t), t = [0, 1, 2, \dots, T]$ and simultaneously the optimality criterion $J_{k+1}(u_{k+1})$ tend to minimum

$$J_{k+1}(u_{k+1}) = \|e_{k+1}\|_Q^2 + \|u_{k+1} - u_k\|_R^2 \quad (3)$$

Where Q, R are diagonal matrix, $Q > 0, R > 0$. u_k, y_k, x_k are, respectively, the super vector of input, output and state of the k th iteration, y_d is the super vector of desired track, $e_k = y_d - y_k$ is the super vector of output error.

We can get the derivative of Eq.3. The control input u_{k+1} satisfies $(\frac{\partial y_{k+1}}{\partial u_{k+1}})^T Q e_{k+1} = -R(u_{k+1} - u_k)$. Rewrite it in form of integral

$$y_{k+1} - y_k = \int_0^1 \frac{\partial y_{k+1}(u_k + \sigma(u_{k+1} - u_k))}{\partial u_{k+1}} d\sigma(u_{k+1} - u_k) \quad (4)$$

Define the symbol $\Delta u = u_{k+1} - u_k$, then Eq.4 can be rewrite

$$[R + (\frac{\partial y_{k+1}(u_k + \Delta u)}{\partial u_{k+1}})^T Q \int_0^1 \frac{\partial y_{k+1}(u_k + \sigma \Delta u)}{\partial u_{k+1}} d\sigma] \Delta u = (\frac{\partial y_{k+1}(u_k + \Delta u)}{\partial u_{k+1}})^T Q e_k$$

Denoting $A((\Delta u)) = (\frac{\partial y_{k+1}(u_k + \Delta u)}{\partial u_{k+1}})^T Q \int_0^1 \frac{\partial y_{k+1}(u_k + \sigma \Delta u)}{\partial u_{k+1}} d\sigma$, $B((\Delta u)) = \frac{\partial y_{k+1}(u_k + \Delta u)}{\partial u_{k+1}}$.

Theorem 1: Consider the nonlinear general discrete-time systems as Eq.1, assuming the following ①, ② and ③ are met:

① $\{\Delta u^i \mid i = 0, 1, 2, \dots\}$ is Δu iterative serial satisfied Eq.4;

② $A(\Delta u), B(\Delta u)$ is limited and satisfied Lipschiz conditions, Lipschiz constant is

L_1 ;

③ Define

$$\alpha = L_1 \|R^{-1}\| \cdot \|\sqrt{Q}\| \cdot \|\sqrt{Q}e_0\| \cdot \|(I + R^{-1}A(\Delta u^i))^{-1}\| \cdot \|(I + R^{-1}A(\Delta u^{i-1}))^{-1}\| \cdot \|R^{-1}B^T(\Delta u^i)\| + \|(I + R^{-1}A(\Delta u^{i-1}))^{-1}\|$$

, $\alpha < 1$; there exist an unique casual $u_{k+1} = \varphi(e_k, u_k)$ satisfied Eq.4. The proof can be seen in paper [10].

In the theorem 1, $u_{k+1} = \varphi(e_k, u_k)$ can't be use directly. We need the causes and effects form.

3 Optimal Iterative Control Approximate Algorithm

From the theorem 1, we can see if $\frac{\partial y_{k+1}}{\partial u_{k+1}}$ is full rank, $\alpha < 1$ is true at all the time. Let's to inspect the convergence of iteration learning algorithm by using $\frac{\partial y_k}{\partial u_k}$ substitute for $\frac{\partial y_{k+1}}{\partial u_{k+1}}$.

We have the following theorem 2.

Theorem 2: In the general nonlinear discrete-time system as Eq.1, using iteration learning algorithm $u_{k+1} = u_k + L_k e_k$, iterative output uniformly converge to the desired output, iterative control uniformly converge to the optimal control. Where

$$L_k = \left(\left(\frac{\partial y_k}{\partial u_k} \right)^T Q \frac{\partial y_k}{\partial u_k} + R \right)^{-1} \left(\frac{\partial y_k}{\partial u_k} \right)^T Q.$$

Proof: We use approximate expression to define $y_{k+1} - y_k$

$$y_{k+1} - y_k \triangleq \frac{\partial y_k}{\partial u_k} (u_{k+1} - u_k)$$

It means the output error

$$e_{k+1} \triangleq e_k - \frac{\partial y_k}{\partial u_k} (u_{k+1} - u_k) \quad (5)$$

Thus $\left(\frac{\partial y_{k+1}}{\partial u_{k+1}} \right)^T Q e_{k+1} = -R(u_{k+1} - u_k)$ can be rewrite as follows

$$\left(\frac{\partial y_k}{\partial u_k} \right)^T Q e_{k+1} = R(u_{k+1} - u_k) \quad (6)$$

Substituting e_{k+1} from above relation Eq.5

$$\left(\frac{\partial y_k}{\partial u_k} \right)^T Q \left(e_k - \frac{\partial y_k}{\partial u_k} (u_{k+1} - u_k) \right) = R(u_{k+1} - u_k)$$

If $\left(\frac{\partial y_k}{\partial u_k} \right)^T Q \frac{\partial y_k}{\partial u_k} + R$ is reversible, then

$$u_{k+1} - u_k = \left(\left(\frac{\partial y_k}{\partial u_k} \right)^T Q \frac{\partial y_k}{\partial u_k} + R \right)^{-1} \left(\frac{\partial y_k}{\partial u_k} \right)^T Q e_k \quad (7)$$

Let's inspect the convergence.

If $\left(\left(\frac{\partial y_k}{\partial u_k} \right)^T Q \right)^{-1}$ is reversible, substitute Eq.7 into the relation Eq.5

$$e_{k+1} = \left(\left(\frac{\partial y_k}{\partial u_k} \right)^T Q \right)^{-1} R \left(\left(\frac{\partial y_k}{\partial u_k} \right)^T Q \frac{\partial y_k}{\partial u_k} + R \right)^{-1} \left(\frac{\partial y_k}{\partial u_k} \right)^T Q e_k = \left(\frac{\partial y_k}{\partial u_k} \right)^T R^{-1} \left(\frac{\partial y_k}{\partial u_k} \right)^T Q + I)^{-1} e_k$$

Takes the norm, we have $\|e_{k+1}\| \leq \theta \|e_k\|$, where $\theta = \left\| \left(\frac{\partial y_k}{\partial u_k} R^{-1} \left(\frac{\partial y_k}{\partial u_k} \right)^T Q + I \right)^{-1} \right\|$.

Since Q, R is diagonal matrix and $Q > 0, R > 0$, $\frac{\partial y_k}{\partial u_k} R^{-1} \left(\frac{\partial y_k}{\partial u_k} \right)^T Q + I$ is always a symmetric matrix and $\frac{\partial y_k}{\partial u_k} R^{-1} \left(\frac{\partial y_k}{\partial u_k} \right)^T Q > 0$. Therefore the following relation is achieved

$$\theta = \left\| \left(\frac{\partial y_k}{\partial u_k} R^{-1} \left(\frac{\partial y_k}{\partial u_k} \right)^T Q + I \right)^{-1} \right\| < 1$$

Thus the output error monotonically converge to zero, $\|e_{k+1}\| \rightarrow 0, k \rightarrow \infty$. At the same time, we can find $u_{k+1} - u_k \rightarrow 0, k \rightarrow \infty$ from Eq.7.

Now evaluate the error of Eq.7.

Define $L_k \triangleq \left(\left(\frac{\partial y_k}{\partial u_k} \right)^T Q \frac{\partial y_k}{\partial u_k} + R \right)^{-1} \left(\frac{\partial y_k}{\partial u_k} \right)^T Q$. We compare the optimal control $u_{k+1} - u_k = L_{k+1} e_k$ and the approximate control $u_{k+1} - u_k = L_k e_k$. For convenience discrimination, use v to instead u , then $v_{k+1} - v_k = L_k e_k$.

The $k+1$ iterative control input deviation

$$u_{k+1} - v_{k+1} = u_k - v_k + H_{k+1} e_k - L_k e_k = u_k - v_k - L_k \frac{\partial y_k}{\partial u_k} (u_k - v_k) + (L_{k+1} - L_k) e_k$$

Takes the norm, we got

$$\|u_{k+1} - v_{k+1}\| \leq \left\| I - L_k \frac{\partial y_k}{\partial u_k} \right\| \cdot \|u_k - v_k\| + \|L_{k+1} - L_k\| \cdot \|e_k\|$$

$$\text{Define } \beta = \left\| I - L_k \frac{\partial y_k}{\partial u_k} \right\| = \left\| I - \left(\left(\frac{\partial y_k}{\partial u_k} \right)^T Q \frac{\partial y_k}{\partial u_k} + R \right)^{-1} \right\| < 1$$

Considering $\|e_k\| \rightarrow 0, k \rightarrow \infty$, we will have $\|u_{k+1} - v_{k+1}\| \rightarrow 0, k \rightarrow \infty$.

Therefore the control of approximate algorithm is converge to the optimal control.

4 Simulation Example

In order to illustrate the performance of the proposed optimal iterative learning law, a numerical example in [11] is presented.

Consider a single-link manipulator which is described by

$$\tau(t) = ml^2 \frac{d^2 \theta(t)}{dt^2} + V \frac{d\theta(t)}{dt} + mgl \cos(\theta(t)) \quad (8)$$

The length, mass and friction coefficient are $l = 1m$, $m = 2.0kg$ and $v = 1.0kgm^2/s$, respectively.

Where h is sample period, $\theta(t)$ is the single-link manipulator angular movement, $\tau(t)$ is the driving moment, g is gravity acceleration.

Eq.8 can describe the model by using the Euler method, as follows

$$\begin{cases} x_1((j+1)h) = x_2(jh) \\ x_2((j+1)h) = (2 - Vh/Ml^2)x_2(jh) + (Vh/Ml^2 - 1)x_1(jh) - gh^2/l \cos x_1(jh) + h^2/Ml^2 u(jh) \\ y(jh) = x_2(jh) \end{cases}$$

Where $j \in [0, 1, \dots, 10]$, $h=0.1$. The initial state $x_1(0) = 0$, $x_2(0) = 0$, the initial control $u_0(jh) = 0$. The desired output $y_d(t) = \sin(5\pi jh/7)$. Where $Q = \text{diag}(800)$, $R = \text{diag}(3)$ in the optimality criterion $J_{k+1}(u_{k+1})$. $\frac{\partial y}{\partial u}$ is lower triangular matrix, can get by recursion.

Using the approximate algorithm Eq.7

$$u_{k+1} - u_k = \left(\left(\frac{\partial y_k}{\partial u_k} \right)^T Q \frac{\partial y_k}{\partial u_k} + R \right)^{-1} \left(\frac{\partial y_k}{\partial u_k} \right)^T Q e_k$$

The output error can be observed in Fig.1, and it is noted that the error monotonically converge to zero.

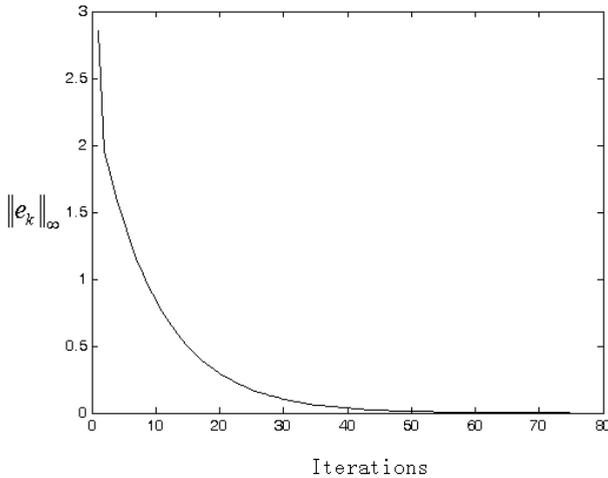


Fig. 1. The output error $\|e_k\|_\infty$ curve, $Q = \text{diag}(800)$, $R = \text{diag}(3)$

The convergent rate is relate to Q, R . The convergent curve of $Q = \text{diag}(800)$, $R = \text{diag}(1.5)$ and $R = \text{diag}(1)$ is shown in Fig.2. It can observe that the error of $R = \text{diag}(1)$ is more rapid.

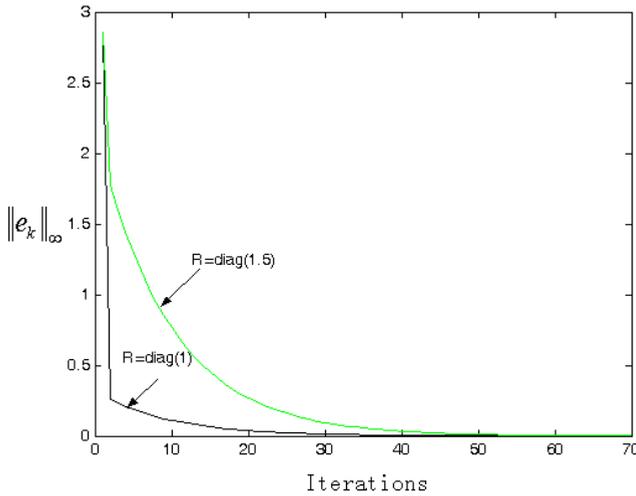


Fig. 2. Convergent rate of $R=\text{diag}(1)$ and $R=\text{diag}(1.5), Q=\text{diag}(800)$

We can obtain from the simulation that iterative output is monotonous restraining and the tracking speed is related to R value. R is bigger, restraining can slow down.

5 Summary

In this paper we studied the iterative learning control problem from an optimal control viewpoint for the general discrete-time nonlinear system, and a new iterative learning law was established. If the partial derivative of output equation to input is exist, the iterative initial state is strictly redundant, and the weight R is big enough, the iterative output restrain monotonously using the causes and effects iteration study law.

References

1. Arimoto, S., Kawamura, S., Miyazaki, F.: Bettering operation of robots by learning. *Journal of Robotic Systems* 1, 123 (1984)
2. Li, H., Hao, X., Xu, W.: A Fast Parameter Optimal Iterative Learning Control Algorithm. In: *ICISS*, p. 375 (2008)
3. Li, H., Hao, X., Zhang, L.: A Clonal Selection Algorithm Based Optimal Iterative Learning Control Algorithm. In: *Proceedings of the 7th World Congress on Intelligent Control and Automation*, p. 927 (2008)
4. Fang, X., Chen, P., Shao, J.: Optimal higher-order iterative learning control of discrete-time linear systems. *IEE Proc.-Control Theory Appl.* 152, 43 (2005)
5. Hätönen, J., Owens, D.H., Feng, K.: Basis functions and parameter optimisation in high-order iterative. *Automatica* 42, 287 (2006)
6. Gunnarsson, S., Norrlof, M.: On the Design of ILC Algorithms Using Optimization. *Automatica* 37, 2011 (2001)

7. Vrabie, D., Pastravanu, O., Abu-Khalaf, M., Lewis, F.L.: Adaptive optimal control for continuous-time linear systems based on policy iteration. *Automatica* 45, 477 (2009)
8. Nasiri, M.R.: An Optimal Iterative Learning Control for Continuous -Time Systems. In: 32nd IECON, p. 114 (2006)
9. Xu, J.-X., Tan, Y., Lee, T.-H.: Iterative learning control design based on composite energy function with input saturation. *Automatica* 40, 1371 (2004)
10. Xu, H., Huang, Y.: Optimal Iterative Learning Control for Nonlinear Discrete-time Systems. *CCIE* 1, 190 (2010)
11. Hwang, D.H., Bien, Z.: Iterative learning control method for discrete time dynamics systems. *IEE2D* 138, 139 (1991)

The Research and Design of Realizing Interconnection and Intercommunication Local Public Emergency Platform

Gu Yan

Guangzhou University
510006 Guangzhou, P.R.China
gy_guyan@yahoo.com.cn

Abstract. Local public emergency platform is a public safety platform under the framework of national emergency platform, with features like monitoring, dynamic prediction, rapid warning and sensitive reaction, decision-making and command, multi-function, advancement and reliability, efficient disposal, helps realize the inter-connection and inter-communication in national public security emergencies. To achieve this goal, we must integrate modern information technology, integrate resources and information of disaster site, on-site command center and the rear command center, constructing multi-system integrated structure of emergency platform system; integrate existing information resources to unify and fuse emergency data, at the same time, use hierarchical classification management technology to construct the comprehensive emergency data center; study the key technology of platform implementation to provide technical support to build local public emergency platform; On this basis, form highly networking and intelligent local emergency command scheduling system, and provide information and communication platform for local management scientific decision in public safety problems.

Keywords: Public security, Emergency platform, Data center, Semantic web, Dynamic preplan library.

1 Preface

Public safety problems mainly involve areas such as natural calamities, accidents and disasters, public health, social security, and major urban infrastructure and key sites security. Public safety is a cornerstone of local security and stability. China is one of the few countries in which many types of emergency coexist, frequently happen and cause great damages. As the urbanization process is speeding up in China, China has entered a high risk of public safety problems, and the frequency of these problems will greatly increase. In recent years, there were more than 200,000 people died and more than 2,000,000 people injured because of public safety problems, and the economic loss reached 6% of the GDP. In China, the SARS in 2004, Wenchuan earthquake in 2008, N1H1 in 2009, Quzhou mud-rock flow in 2010, various frequent disasters and production safety accidents, and the earthquake just happened in Japan in 2011, all shows that public security problem has become one of the important factors which affect human survival and social stability. In this grim situation, how to reduce casualties and economic loss caused by disasters and breakthrough key technologies which restrict local public safety has

become an urgent problem in local areas as well as nationwide. Therefore, relying on Comprehensive integration of high-tech, and under the framework of the general platform for the integrated national public security emergency, it is extremely significant to construct the key technology for the decision-making and command platform for local public safety, and realize integration, real-time change, precision and quick response in public emergency, providing reliable technical support and assurance for local implementation of emergency disasters and accidents and public emergencies ability, and upgrading the ability of local government to response to disasters and public emergencies.

2 General Ideas for Emergency Platform Design

Local public emergency platform is a platform under the framework of national public emergency management with monitor, dynamic prediction, rapid warning and sensitive

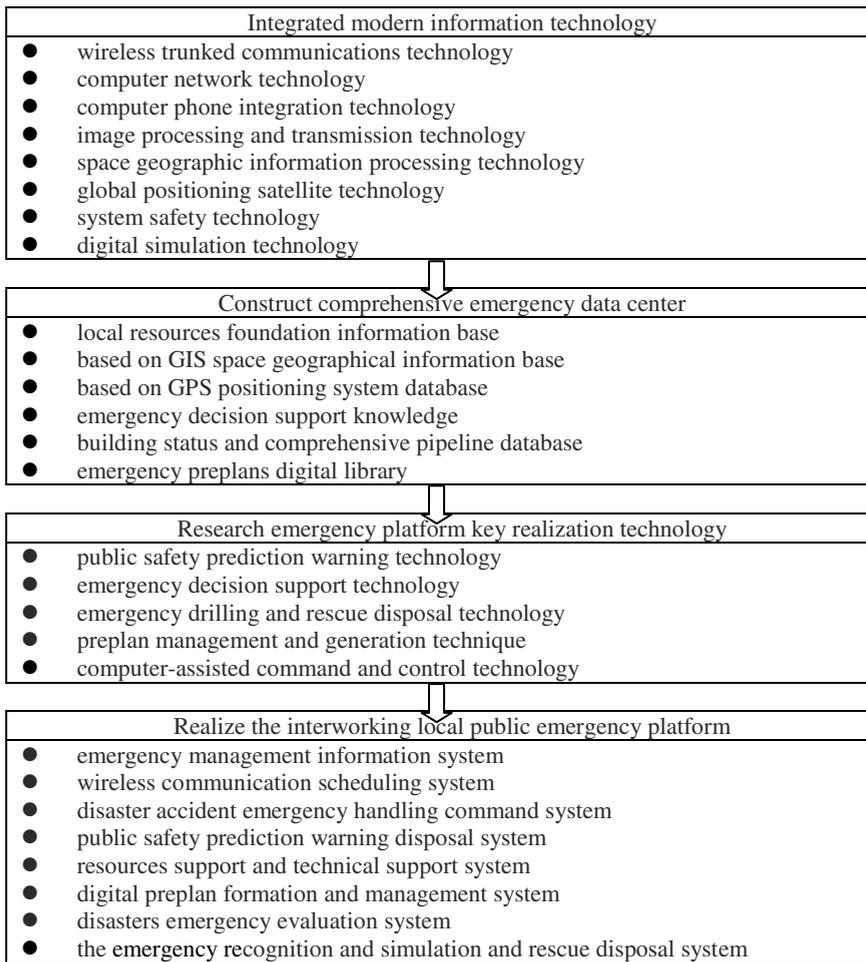


Fig. 1. General Design Ideas for Emergency Platform

reaction, decision-making and command, multi-function, reliability and efficiency as its characteristics, and it can realize inter-communication with the general national public emergency platform. To achieve this goal, we must integrate modern information technology, integrate resources and information of disaster site, on-site command center and the rear command center, constructing multi-system integrated structure of emergency platform system; integrate existing information resources to unify and fuse emergency data, at the same time, use hierarchical classification management technology to construct the comprehensive emergency data center; study the key technology of platform implementation to provide technical support to build local public emergency platform; On this basis, form highly networking and intelligent local emergency command scheduling system, and provide information and communication platform for local management scientific decision in public safety problems. The general ideas for the emergency platform design are shown in figure 1.

3 The Design of Emergency Platform Prototype

The main function of local public emergency platform is to provide services for core operation including early warning, monitoring, dynamic decision-making, emergency interaction, comprehensive coordination, information management and business processing. System design must attach importance to the system's safety, reliability and operability, openness, practicality and flexibility. The system uses modular, multilayered structure design. Emergency platform is structured by two parts including hardware and software based support system and integrated application system. The support system includes: wireless trunked communications system, computer network system, image processing and transmission system, various video conference system, space geographic information system (GIS), global positioning system (GPS), security support systems, disaster tolerance system etc. Application system includes: emergency management information system, wireless communications scheduling system, disaster accident emergency handling commanding system, digital preplan formation and management system, forecasting and warning disposal system, resources support and technical guarantee system, disasters identification and emergency evaluation system, simulation emergency drilling and rescue disposal system, etc. The entire system takes the imported heterogeneous data as the main sources, the emergency command as data terminal, the metadata management, data mining strategy management, data cleaning rules management as the major data processing rule, and it takes data mining, decision support, preplan generation as the core of data processing to complete system overall objectives. The whole system model framework is illustrated in figure 2.

4 Digital Dynamic Preplan Library Established

Emergency preplan is an important basis for emergency command and disposal in case of emergent events. Once an emergent event occurs, the emergency system, judging the category and degree of the event, can quickly generate an emergency preplan and

feasible operation plan by the emergency preplan generation system, thus, it achieves the integration of multi-departments, multi-disciplinary, multi-level resources, realizing emergency command and emergency rescue organization.

The digital dynamic semantics preplan library is established based on the semantic web emergency response structured technology. Essentially, semantic preplan library is a semantic database in which the data is stored in RDF file form, focusing on establishing a repository on public emergency preplan with data sources from unstructured content of all kinds of preplan which can be submitted through bookmarks, mail or in manual way. In the representation of a preplan, RDF can be applied to mark the data structure, and the original pure text descriptive preplan document can be processed into query-able and combinable information with computer semantics, i.e., a preplan is constituted by a group of data with certain characteristics such as preplan title, preplan type, formulate time, preplan content and so on, each of which may be identified by computer. Using RDF to describe preplan data constitution can be expressed as follows:

```
<? xml version="1.0"?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:contact="http://www.w3.org/2000/10/swap/pim/contact#"
xmlns:cd="http://www.counterpreplan.security/cd#">
  <rdf:Description
rdf:about="http://www. counterpreplan.security/cd/Disaster ">
    <cd: preplan title>Urban earthquake emergency preplan</cd: preplan title>
    <cd: preplan type>Earthquake</cd: preplan type>
    <cd: formulate time>2007-6-1</cd: formulate time>
    <cd: preplan content 1> Emergency organization command system </cd:
preplan content 1>
    <cd: preplan content 2> Monitoring and warning </cd: preplan content 2>
    <cd: preplan content 3> Emergency standard operating procedures </cd:
preplan content 3>
    <cd: preplan content 4> Standard operating procedures later </cd: preplan
content 4>
    <cd: preplan content 5> Emergency security support system </cd: preplan
content 5>
    <cd: preplan content 6> Contingency preplans to drill and training </cd: preplan
content 6>
    <cd: preplan content 7> Accessories </cd: preplan content 7>
  </rdf:Description>
</rdf:RDF>
```

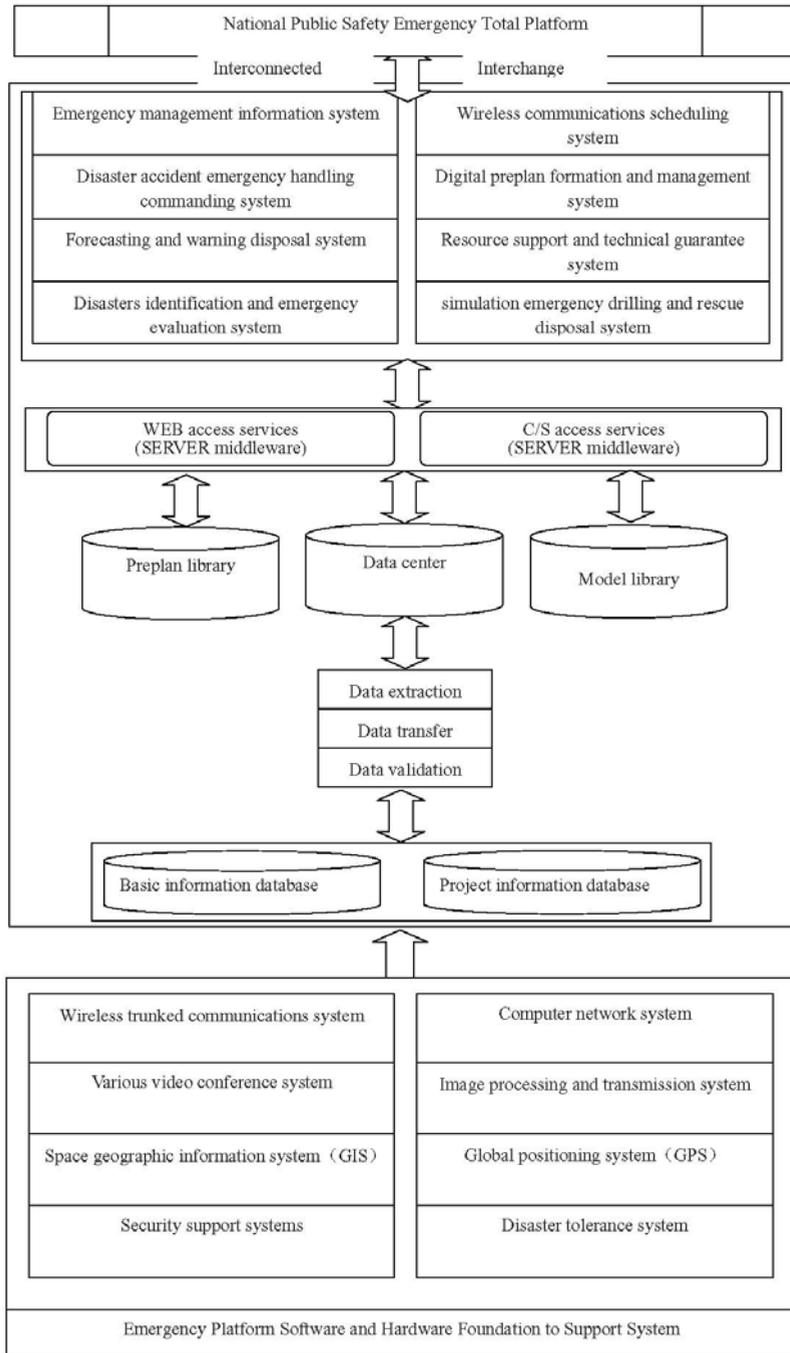


Fig. 2. The Whole System Model Framework

5 Local Public Safety Data Center Design

The public safety data center which takes data storage and usage as its core integrated by storage equipment, network equipment, application data and application programs, with functions such as integrated based data storage, information query, data management and data maintenance and so on. It should have a high degree of flexibility and adaptability to make a rapid change in external demand. Therefore, data center needs to adopt emergency platform "level + hybrid" method for data center system structure design according to the need of the emergency system. From the system structure, data center can be divided into several levels: the source data layer, storage equipment layer, core data layer, security component layer, data access layer. From the data structure, data center can be divided into three administrative levels including foundation database layer, public security project database layer and business database layer. Through methods of concentration, fusion, sharing and exchange, data center is completed in content, interconnected in physics and integrated in logic. Meanwhile, on the core layer in the data center, the data "hybrids" together in a certain structure. In a storage device layer, large capacity high-performance storage solutions is used via disk array center for storage center, storage area network (SAN) for storage structure, with tape library for backup center. Public safety data center architecture is shown in figure 3.

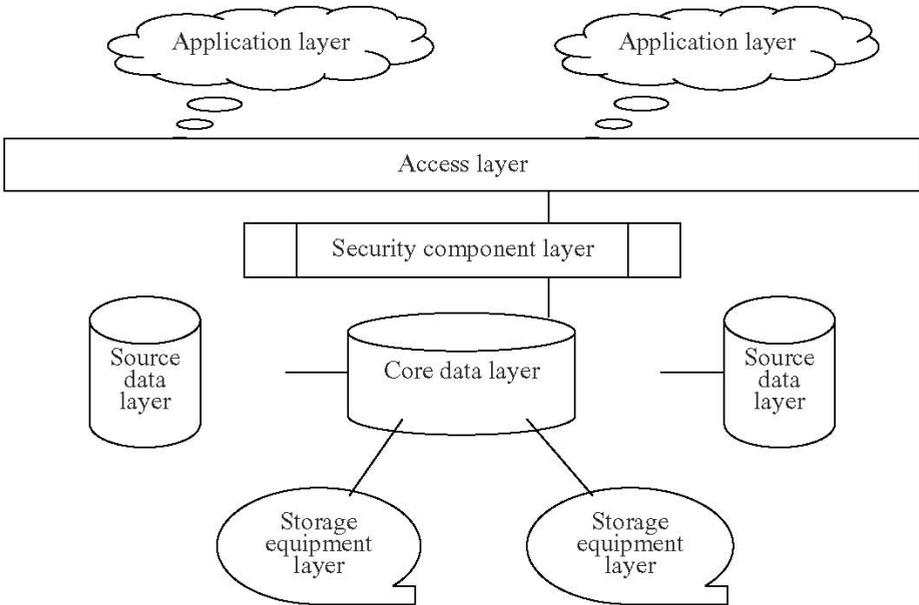


Fig. 3. Public Safety Data Center Architecture



6 Summary

Local emergency platform is an emergency platform under the framework of the national public emergency platform. It is also an organic component of the local e-government systems. When the platform is completed, it will become an important platform for the local emergency information management, monitoring, dynamic prediction, rapid warning, decision-making and command and efficient disposal, making the local emergency information resources integrated and strengthened; thus, it forms a highly intelligent local emergency command and scheduling system and provides an information and communication platform for scientific decision making in local management and public safety.

References

1. Liu, T., Li, H., Deng, Y.: Public Incident Emergency Management Information System for Both Peacetime and Wartime. *Journal of China Occupational Safety and Health Management System Certification* 5 (2005)
2. Mao, Z., Zhu, B.: Study on the Frame of Decision Support System for Emergency Counter-preplan of City. *Journal of Geomatics & Spatial Information Technology* 30 (2007)
3. Li, S., Si, S., Zhang, X.: The Application of City Emergency Management Information System Function and GIS Function. *Journal of Urban Management Science & Technology* 20 (2007)
4. Tian, Y.: Modeling of Urban Community Information Management System for Public Security Emergencies. *Journal of Wuhan University of Technology* 29 (2007)
5. Liu, H., Liu, Y.: Study on Social Accident Information System for Meeting an Urgent need. *Journal of Information Science* 23 (2005)
6. Liao, G., Weng, T.: A Study on Assistant Decision Support System for Urban Emergency Response. *Journal of Engineering Science* 7 (2005)

The Normal Range of House Price Rising Rate of China

Zhang Hongli

Xi'an University of Science and Technology Shaanxi Province, China
honglizhang@sina.com

Abstract. The development of real estate industry is significant to the national economy and the people's livelihood, and the regulation of real estate is an important content of macroeconomic regulation and is controversial invariably. The key of real estate regulation is the regulation of house price, i.e. to control the rising extent of house price within a normal range. The article calculates the target range of house price regulation in 3σ Method. According to the conclusions, should control the house price rising rate in first-tier cities within $[-10.84\%, 10.56\%]$, and that in second-tier cities within $[-5.13\%, 15.99\%]$. The conclusions can provide a theoretical support for the stipulation of real estate regulation policies and the evaluation of the effects of real estate regulation policies.

Keywords: Real Estate, house price, regulation, normal range.

1 Introduction

From 2003, the nation has made a series of real estate regulation policies for alleviating the economic and social contradictions resulted from the too fast rising of house price. From the second half year of 2009, the Chinese real estate market walked out the adjustment situation during the international financial crisis, with both volume and price rising, and the government is facing the difficulty of real estate regulation again. The regulation of house price is the key of real estate regulation. But the target range of house price regulation, i.e. whether the rising rate of house price is normal and requires the regulation of the government, and whether the regulation measures have achieved the predicted target, hasn't a clear definition until now. This is the purpose of the article.

2 Research Ideas

The target range of house price regulation is the reasonable range of house price fluctuation, and means the rising or declining extent of house price is coordinative with the situations of economic growth and does not need the government to intervene. When the rising extent of house price year on year is within this range, it is considered as normal and does not need to be regulated and controlled. When the rising extent of house price exceeds the upper limit of normal fluctuation range, it may be considered as "overheated" of the real estate market, and needs contractionary

regulation; vice versa, when the declining extent of house price exceeds the upper limit of normal fluctuation range, it may be considered as “overcoldded” of the real estate market, and needs expansionary regulation. According to this, the running states of real estate market can be divided into three types: normal, basically-normal, abnormal (overcoldded or overheated). Currently, the main target of Chinese house price regulation is to prevent too large house price rising extent and overheated real estate market.

For convenient to study the target range of house price regulation, the article chooses two important indexes, house price rising rate and house price rising rate/urban GDP growth rate, and calculates the normal range of these two indexes in Composite Simulation Method. The Composite Simulation Method was established on the basis of Japanese experience, determined the critical values of the indexes according to the statistical law, and so to judge the running state of real estate market.

The traditional method to determine critical values is Empirical Data Method. Due to its poor accuracy, two new methods to determine critical values, Neural Network Method and 3σ Method based on normal distribution assumption, have grown up recently. Currently, the application of Neural Network Method is still at the start, not mature and perfect enough. Therefore, the article determines critical values in the modified 3σ Method.

3 Theory of 3σ

Normal distribution is the most common type of probability distribution type. 3σ Method assumes that the index variables are distributed normally, and accordingly studies the change law of indexes[]. The probability function of normal distribution is:

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \tag{1}$$

Of which, μ is the mean value of overall samples, and σ is the standard deviation of overall samples.

According to the normal distribution law and formula can calculate the following probability:

$$\Phi(x) = p(\mu - \sigma < x < \mu + \sigma) = p[-\sigma < (x - \mu) < \sigma] = p\{-1 < \frac{x - \mu}{\sigma} < 1\} \tag{2}$$

So, $F(x) = F(1) - F(-1) = 0.8413 - 0.1587 = 0.6826$ (3)

Similarly, can obtain:

$$p(\mu - 2\sigma < x < \mu + 2\sigma) = \Phi(2) - \Phi(-2) = 0.9546 \tag{4}$$



$$p(\mu - 3\sigma < x < \mu + 3\sigma) = \Phi(3) - \Phi(-3) = 0.9973 \quad (5)$$

$$p(\mu - 4\sigma < x < \mu + 4\sigma) = \Phi(4) - \Phi(-4) = 0.999937 \quad (6)$$

Accordingly, can obtain the following conclusion:

Firstly, assuming that all the characteristic values of indexes are subject to normal distribution, then in the total data of characteristic values near 70% are within $\pm\sigma$, 95.46% within $\pm 2\sigma$, and only less than 5% exceed $\pm 2\sigma$. It shows that most of the characteristic values are distributed within $\pm 2\sigma$; at the same time, the indexes within $\pm 2\sigma$ can represent the random distribution law of indexes in real estate industry.

Secondly, the principle of 3σ is not affected by the value of σ . No matter what size the value of σ is, 68.3% characteristic values are contained within $\pm\sigma$. 95.46% and 99.73% characteristic values are contained respectively within $\pm 2\sigma$ and within $\pm 3\sigma$.

For more accurately defining the normal range of house price fluctuation, it is necessary to calculate those of first-tier cities and those of second-tier cities respectively. Currently, the house price in first-tier cities is on the high side, with large investment risk, weakening attraction to investment capital, and has exceeded the house purchasing ability of residents. The range of one time of standard deviation can be chosen as the upper limit of the normal range of rising extent fluctuation of house price, and the range of three times of standard deviation can be chosen as the lower limit of the normal range of declining extent fluctuation of house price. For the second-tier cities with lower house price, one time of standard deviation is too conservative, thus, two times of standard deviation is chosen as the upper limit of the normal range of house price rising extent, and the range of one to two times of standard deviation is considered as the normal range of house price fluctuation.

Taking the mean value of each index as the central value of its fluctuation, firstly calculate its central value. Then determine the normal growth range. According to the above analysis, for the first-tier cities, can regard the range of one time of standard deviation from the central value as the upper limit of normal range, and the range of three times of standard deviation as the lower limit of the normal range, i.e. $[\mu - 3\sigma, \mu + \sigma]$. for the second-tier cities, can regard the range of one time of standard deviation from the central value as the normal range, i.e. $[\mu - \sigma, \mu + \sigma]$. regard the range of one to two times of standard deviation from the central value as the basic normal range, i.e. $[\mu - \sigma, \mu - 2\sigma]$ and $[\mu + \sigma, \mu + 2\sigma]$. Regard the range of over two times of standard deviation from the central value as the abnormal (overcoldded or overheated) range, i.e. $[-\infty, \mu - 2\sigma]$ and $[\mu + 2\sigma, +\infty]$. In this way, the normal range of indexes in the second-tier cities is $[\mu - 2\sigma, \mu + 2\sigma]$, and its boundary is the critical value for judging whether regulation is needed.

Obviously, the analysis on the regulation standards of real estate market is taking a specific economic development situation as premise. The above calculation is obtained from the statistical data of 2003~2008, therefore, its analysis conclusion is suitable for the similar macroeconomic situations. In this period, the yearly average

growth rate of GDP in Chinese cities reached 10.68%, the financial institutions' loan rate of five years above was 6.12~7.20%, yearly average inflation rate was 3.2%, and the cities were in the fast urbanization development stage with yearly average growth rate of urban population of about one percent.

According to the above principles, we can calculate the normal range of house price rising rate and house price rising rate/urban GDP growth rate.

4 Normal Range of House Price Rising Rate

The indexes of house price are the most important indicators for reflecting the trend of house price variation and the situation of real estate market. According to the China Statistical Yearbook, can calculated the house price rising rate in these 35 cities in 2003~2008 (Table 1).

Table 1. House Price Rising Rate in 35 Large and Middle Cities (%)

| | 2000 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | | 2000 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
|-----------|------|------|------|-------|-------|-------|------|-----------|------|------|------|-------|------|------|-------|
| Total | 1.1 | 4.8 | 9.7 | 7.60 | 5.51 | 7.60 | 6.46 | Qingdao | 2.3 | 14.6 | 15.3 | 10.93 | 6.92 | 6.50 | 5.09 |
| Beijing | -0.5 | 0.3 | 3.7 | 6.68 | 8.75 | 11.40 | 9.45 | Zhengzhou | -0.5 | 2.0 | 4.0 | 6.97 | 5.69 | 6.30 | 3.32 |
| Tianjin | 0.0 | 4.1 | 13.5 | 5.97 | 6.70 | 6.90 | 5.84 | Wuhan | 1.5 | 3.8 | 8.4 | 6.83 | 3.03 | 5.20 | 4.91 |
| Shijiazhe | 1.8 | 0.3 | 3.6 | 5.55 | 4.34 | 7.60 | 5.75 | Changsha | -0.4 | 0.5 | 3.3 | 2.83 | 5.30 | 8.40 | 6.66 |
| Taiyuan | 1.1 | 2.8 | 6.4 | 5.60 | 3.90 | 4.40 | 5.70 | Guangzhou | -2.7 | -0.7 | 2.7 | 4.65 | 6.22 | 6.60 | -0.16 |
| Huhehaote | 2.0 | 0.7 | 5.2 | 11.78 | 9.46 | 4.40 | 1.23 | Shenzhen | -0.8 | 2.2 | 4.6 | 7.22 | 12.3 | 16.3 | -1.86 |
| Shenyang | 3.0 | 7.6 | 15.9 | 7.50 | 6.58 | 6.10 | 4.57 | Nanjing | -0.7 | 2.1 | 5.7 | 4.93 | 4.12 | 7.60 | 8.16 |
| Dalian | 0.2 | 0.7 | 4.6 | 9.15 | 10.87 | 7.20 | 4.82 | Haikou | -0.6 | 2.7 | 5.9 | 2.53 | 2.80 | 6.60 | 10.4 |
| Changchun | 6.6 | 0.2 | 0.2 | 1.93 | 1.60 | 6.40 | 6.97 | Chengdu | 1.3 | 2.9 | 7.9 | 7.20 | 3.00 | 6.90 | 6.29 |
| Ha'erbin | 1.8 | 0.2 | 4.7 | 4.60 | 3.27 | 6.80 | 6.36 | Guiyang | 3.9 | 1.3 | 2.6 | 9.78 | 7.12 | 7.60 | 3.39 |
| Shanghai | -1.4 | 20.1 | 15.9 | 9.72 | -1.33 | 3.40 | 5.91 | Kunming | 0.2 | -0.9 | 2.3 | 2.60 | 4.41 | 6.90 | 6.60 |
| Nanjing | 1.6 | 9.8 | 15.3 | 8.05 | 4.33 | 6.60 | 2.75 | Chongqing | 1.8 | 6.1 | 13.9 | 2.90 | 1.33 | 3.50 | 3.22 |
| Hangzhou | 4.9 | 6.1 | 11.7 | 9.72 | 2.63 | 7.30 | 8.58 | Xi'an | 1.3 | 1.4 | 5.0 | 4.28 | 3.59 | 6.40 | 8.06 |
| Ningbo | 5.5 | 16.6 | 13.9 | 6.35 | 2.16 | 8.60 | 9.16 | Lanzhou | 0.5 | 1.8 | 8.7 | 5.55 | 4.72 | 6.00 | 9.77 |
| Hefei | 0.0 | 4.1 | 5.6 | 6.15 | 1.25 | 1.80 | 8.35 | Xining | 1.1 | 1.9 | 4.0 | 3.35 | 2.79 | 3.80 | 7.51 |
| Fuzhou | 0.3 | 1.1 | 3.6 | 4.40 | 6.67 | 6.80 | 3.85 | Yinchuan | 2.2 | 2.1 | 4.4 | 2.65 | 2.33 | 3.90 | 11.7 |
| Xiamen | 0.1 | 2.8 | 7.3 | 8.00 | 6.95 | 7.00 | 2.71 | Wulumuqi | 2.4 | -0.1 | 0.7 | 0.92 | 1.19 | 9.00 | 15.5 |
| Nanchang | 3.2 | 4.8 | 7.3 | 8.30 | 6.23 | 6.80 | 4.15 | Qingdao | 2.3 | 14.6 | 15.3 | 10.9 | 6.92 | 6.50 | 5.09 |
| Jinan | 2.8 | 3.1 | 10.3 | 7.55 | 4.27 | 5.20 | 7.22 | Zhengzhou | -0.5 | 2.0 | 4.0 | 6.97 | 5.69 | 6.30 | 3.32 |
| Wuhan | 1.5 | 3.8 | 8.4 | 6.83 | 3.03 | 5.20 | 4.91 | Changsha | -0.4 | 0.5 | 3.3 | 2.83 | 5.30 | 8.40 | 6.66 |

Then calculate the mean value and mean square deviation in these 35 cities, and determine the critical value of house price rising rate when the real estate market is under normal state in the method of 3σ .

According to this can obtain the critical value of these 35 large and middle cities under abnormal state (overcoldded or overheated). When the house price rising rate exceeds the critical value in a city, it can be considered that the real estate market is overcoldded or overheated in the city.

Table 2. Critical Value of House Price Rising Rate in 35 Large and Middle Cities (%)

| | σ | $\mu+\sigma$ | $\mu-\sigma$ | $\mu-2\sigma$ | $\mu+2\sigma$ | $\mu+3\sigma$ | $\mu-3\sigma$ | | σ | $\mu+\sigma$ | $\mu-\sigma$ | $\mu-2\sigma$ | $\mu+2\sigma$ | $\mu+3\sigma$ | $\mu-3\sigma$ |
|-----------|----------|--------------|--------------|---------------|---------------|---------------|---------------|-----------|----------|--------------|--------------|---------------|---------------|---------------|---------------|
| Total | 2.53 | 8.6 | 3.59 | 3.59 | 11.16 | 13.69 | -1.46 | Zhengzhou | 2.44 | 6.4 | 1.53 | 1.53 | 8.85 | 11.29 | -3.35 |
| Beijing | 4.28 | 10.0 | 1.40 | 1.40 | 14.25 | 18.53 | -7.17 | Wuhan | 2.15 | 7.0 | 2.66 | 2.66 | 9.10 | 11.25 | -1.63 |
| Tianjin | 3.72 | 9.9 | 2.42 | 2.42 | 13.59 | 17.31 | -5.03 | Changsha | 2.96 | 6.8 | 0.84 | 0.84 | 9.72 | 12.68 | -5.08 |
| Shijiazhe | 2.30 | 6.4 | 1.83 | 1.83 | 8.74 | 11.04 | -2.77 | Guangzhou | 3.37 | 5.7 | 1.00 | -1.00 | 9.11 | 12.49 | -7.74 |
| Taiyuan | 1.72 | 6.0 | 2.55 | 2.55 | 7.71 | 9.43 | -0.89 | Shenzhen | 6.22 | 11.9 | -0.51 | -0.51 | 18.15 | 24.37 | -12.96 |
| Huhehaote | 3.93 | 8.9 | 1.04 | 1.04 | 12.82 | 16.74 | -6.81 | Nanjing | 2.87 | 7.4 | 1.69 | 1.69 | 10.29 | 13.16 | -4.05 |
| Shenyang | 3.82 | 11.1 | 3.51 | 3.51 | 14.96 | 18.77 | -4.13 | Haikou | 3.33 | 7.7 | 1.01 | 1.01 | 10.99 | 14.31 | -5.64 |
| Dalian | 3.73 | 9.1 | 1.63 | 1.63 | 12.82 | 16.55 | -5.82 | Chengdu | 2.41 | 7.5 | 2.66 | 2.66 | 9.88 | 12.29 | -2.15 |
| Changchun | 2.88 | 6.3 | 0.54 | 0.54 | 9.16 | 12.04 | -5.21 | Guiyang | 2.86 | 8.0 | 2.24 | 2.24 | 10.82 | 13.67 | -3.48 |
| Ha'erbin | 2.21 | 6.2 | 1.75 | 1.75 | 8.37 | 10.58 | -2.66 | Kunming | 2.77 | 5.9 | 0.39 | 0.39 | 8.70 | 11.47 | -5.15 |
| Shanghai | 7.66 | 15.1 | -0.19 | -0.19 | 22.79 | 30.45 | -15.51 | Chongqing | 4.02 | 8.7 | 0.65 | 0.65 | 12.72 | 16.75 | -7.39 |
| Nanjing | 4.35 | 11.3 | 2.57 | 2.57 | 15.61 | 19.96 | -6.12 | Xi'an | 2.30 | 6.6 | 1.99 | 1.99 | 8.88 | 11.18 | -2.60 |
| Hangzhou | 2.83 | 10.1 | 4.45 | 4.45 | 12.93 | 15.75 | -1.20 | Lanzhou | 3.11 | 8.4 | 2.18 | 2.18 | 11.51 | 14.62 | -4.04 |
| Ningbo | 4.60 | 13.5 | 4.30 | 4.30 | 18.09 | 22.69 | -4.90 | Xining | 1.90 | 5.4 | 1.59 | 1.59 | 7.29 | 9.19 | -2.21 |
| Hefei | 2.79 | 6.7 | 1.10 | 1.10 | 9.47 | 12.26 | -4.47 | Yinchuan | 3.20 | 7.4 | 0.99 | 0.99 | 10.59 | 13.79 | -5.40 |
| Fuzhou | 2.30 | 6.1 | 1.51 | 1.51 | 8.42 | 10.73 | -3.09 | Wulumuqi | 5.40 | 9.6 | 1.17 | -1.17 | 15.04 | 20.44 | -11.98 |
| Xiamen | 2.83 | 7.8 | 2.15 | 2.15 | 10.65 | 13.48 | -3.52 | Zhengzhou | 2.44 | 6.4 | 1.53 | 1.53 | 8.85 | 11.29 | -3.35 |
| Nanchang | 1.70 | 7.5 | 4.13 | 4.13 | 9.22 | 10.92 | 0.73 | Wuhan | 2.15 | 7.0 | 2.66 | 2.66 | 9.10 | 11.25 | -1.63 |
| Jinan | 2.52 | 8.3 | 3.26 | 3.26 | 10.81 | 13.33 | -1.77 | Changsha | 2.96 | 6.8 | 0.84 | 0.84 | 9.72 | 12.68 | -5.08 |
| Shenzhen | 6.22 | 11.9 | 0.51 | -0.51 | 18.15 | 24.37 | -12.96 | Guangzhou | 3.37 | 5.7 | 1.00 | -1.00 | 9.11 | 12.49 | -7.74 |

Finally, the 35 cities can be divided into first- and second-tier cities by experience. Generally, first-tier cities mean Beijing, Shanghai, Hangzhou, Shenzhen, and Guangzhou, and all the others are second-tier cities. Here taking eight typical cities, Nanjing, Tianjin, Shenyang, Wuhan, Chengdu, Chongqing, Xian, and Urumchi, to calculate their critical values of house price rising rate (table 3-6).

Table 3. Upper Limit of House Price Rising Rate in First-tier Cities (%)

| City | Beijing | Shanghai | Hangzhou | Shenzhen | Guangzhou | Average |
|-------------------------------|---------|----------|----------|----------|-----------|---------|
| Critical Value for Overheated | 10.0 | 15.1 | 10.1 | 11.9 | 5.7 | 10.56 |

Table 4. Upper Limit of House Price Rising Rate in Second-Tier Cities (%)

| City | shenyang | chengdu | Xi'an | nanjing | wuhan | chongqing | average |
|-------------------------------|----------|---------|-------|---------|-------|-----------|---------|
| Critical Value for Overheated | 14.96 | 14.99 | 8.88 | 15.61 | 9.10 | 12.72 | 14.47 |

Table 5. Lower Limit of House Price Rising Rate in First-tier Cities (%)

| city | beijing | shanghai | wuhan | gunagzhou | average |
|-----------------------------|---------|----------|--------|-----------|---------|
| Critical Value for Overcold | -7.17 | -15.51 | -12.96 | -7.74 | -10.84 |

Table 6. Lower Limit of House Price Rising Rate in Second-tier Cities (%)

| City | shenyang | chengdu | Xi'an | nanjing | wuhan | chongqing | average |
|--------------------------------|----------|---------|-------|---------|-------|-----------|---------|
| Critical Value for Overcoldded | -4.13 | -2.15 | -2.60 | -5.03 | -1.63 | -7.39 | -5.13 |

It can be seen from the calculated results that, the upper limit of normal rising rate is 10.56%, lower limit is -10.84%, and the target range of regulation is [-10.84%, 10.56%] in first-tier cities; the upper limit of normal rising rate is 12.47%, lower limit is -5.13%, and the target range of regulation is [-5.13%, 12.47%] in second-tier cities.

5 Conclusion

The real estate market is a typical regional market. Due to the unbalanced economic development in Chinese regions and the large difference of running states of real estate market in cities, therefore, should select the regulation target and corresponding policies and tools via classification.

House price regulation is the most important content of real estate regulation. For the comparison of house price fluctuation range, should inspect whether the growth rate of house price in the same period is within the normal range by monthly data. That is to say, comparing the mean values of commercial residential building of all months to those in the same period of last year, if the growth rate in the same period exceeds the normal range determined in this article, the central and local governments should consider publishing regulation measures for the city or region. According to the practical situations in China, currently the main target of house price regulation is to prevent fast rising of house price, i.e. prevent overheating of real estate market.

According to the calculations, should control the house price rising rate in first-tier cities within [-10.84%, 10.56%], and that in second-tier cities within [-5.13%, 15.99%]. Looking from the practical situations of real estate market in recent years, the normal range of house price fluctuation determined in the article is reliable. In some years real estate market was overheated, looking from the monthly comparative indexes, the house price rising extent exceeded the upper limit of normal range and needed the government to regulate and control in Beijing, Shenzhen, Shanghai, Hangzhou, and Haikou, etc.

References

1. Huang, Z.: Several Theoretical Problems on Bubble Economy and Its Measurement. Financial Research 06 (2002)
2. Duan, Y., Tian, P., Zhang, X.: Relative efficiency control fuzzy system. In: Proceeding of the 46th ISSS Annual Conference, Shanghai, China (2002)
3. He, Y., Zhang, B.: Research on influence of macro-regulation on real estate industry in China. Price Theory and Practice (12), 45–46 (2006) (in Chinese)
4. Yin, H., Tan, S.: Policy orientation of macroeconomic regulation on Chinese Real estate market. Statistics and Decision (10), 100–102 (2006) (in Chinese)

A Bottom-Up Incremental Algorithm of Building Concept Lattice

LingLing Lv¹, Lei Zhang², PeiYan Jia², and FuNa Zhou²

¹ College of electric power, North China Institute of Water Conservancy and Hydroelectric Power, Zhengzhou, 450011, P.R. China

² Institute of Data and Knowledge Engineering, Henan University, Kaifeng, 475001, China
lingling_lv@163.com, {zhanglei, jpy}@henu.edu.cn

Abstract. The construction algorithm of concept lattice is an important topic in formal concept analysis. So far, a series of algorithms have been proposed, which can be divided into two categories: Batch and incremental algorithms. Because it is easy to maintain, the incremental algorithms are very promising algorithms. Based on analysis to the process of incremental construction of concept lattice, a bottom-up incremental algorithm of building concept lattice has been proposed. The tests show that this algorithm outperforms classical Godin algorithm for most types of contexts.

Keywords: Concept lattice, Construction algorithm, Bottom-up.

1 Introduction

Concept lattice is the core data structure of Formal Concept Analysis (FCA) [1] and it demonstrates the knowledge view of data. As a powerful mathematical tool of knowledge discovery and data analysis, concept lattice has attracted more and more attention from the researchers and found its important application in machine learning[2], data mining[3], social network analysis[4] and so on.

Concept lattice formation process is a concept clustering process. Since concept lattice has been proposed, building concept lattice from a formal context is always the emphasis in concept analysis studies. Because of completeness of concept lattice, construction efficiency of concept lattice is always a major factor affecting application of concept lattice. So far, a series of algorithms have been proposed for improving efficiency of building concept lattice. The existing concept lattice constructing algorithms can be substantially divided into incremental constructing algorithm and batch constructing algorithm. The idea of batch algorithm is generating all the concepts firstly, and then generating edges based on their direct predecessor - successor relationship and realizing the construction of concept lattice finally, e.g., Bordat algorithm [5], Nourine algorithm [6], and so on. The incremental constructing algorithm is based on iterative, which is carried as follows: one object is added from formal context to concept lattice once and concept generating and relation updating are realized simultaneously, then another object is continue to added to the new concept lattice. Because it is easy to maintain, the incremental constructing algorithm is always the hot topic of research, among which a classical algorithms are named Godin

algorithm [7]. In the recent years, many researchers [8, 9, and 10] have dedicated to improve the above algorithms to further enhance the efficiency of the algorithms.

Through the research on the characteristics of concept lattice structure, a simple and fast incremental construction algorithm is proposed in this paper. The main idea of this algorithm is as follow: algorithm accesses each concept of lattice in accordance with the breadth-first. For each concept, algorithm judges whether the concept is a generator or not. If the concept is a generator, algorithm generates a new concept, then adjusts child relation of new concept and passes new concept as candidate child of parent of the generator. The algorithm generates all new concepts through searching for generator and adjusting Hasse diagram. Compared with the existing algorithm, it has higher efficiency.

2 The Basis of Concept Lattice

This section mainly deals with some basic definitions of concept lattice [1].

Definition 1. *Formal context is a triple $K = (G, M, I)$, where G is an object set, M is an attribute set and I is the binary relation between G and M . For an object $g \in G$ and an attribute $m \in M$, then gIm denotes that object g has attribute m .*

Definition 2. *Given object set $O \in P(G)$ and attribute set $D \in P(M)$ of a formal context, define the following two mappings:*

$$f(O) = \{m \in M \mid \forall g \in O, gIm\};$$

$$g(D) = \{g \in G \mid \forall m \in D, gIm\}.$$

Definition 3. *In a formal context, if a binary group $C = (O, D)$ satisfies $O = g(D)$, $D = f(O)$, then we call the binary group (O, D) as a formal concept (referred as concept in case of no confusion), where O is the element of object power set $P(G)$ and is called as extension of concept C , D is the element of attribute power set $P(M)$ and is called as intension of concept C .*

Definition 4. *Given concepts $C_1 = (O_1, D_1)$ and $C_2 = (O_2, D_2)$, if $O_1 \subseteq O_2$ (equivalent to $D_2 \subseteq D_1$), then we call $(O_1, D_1) \leq (O_2, D_2)$; furthermore, if there doesn't exist $C_3 = (O_3, D_3)$ such that $(O_1, D_1) \leq (O_3, D_3) \leq (O_2, D_2)$, then we call $(O_1, D_1) < (O_2, D_2)$, where (O_1, D_1) is called as child concept and (O_2, D_2) is called as parent concept. For arbitrary two concepts C_1 and C_2 , their upper bound is denoted by $C_1 \vee C_2$ and their lower bound is denoted by $C_1 \wedge C_2$.*

Definition 5. *The lattice generated by partial order relation \leq is called as the concept lattice of formal context K , which is denoted by $L(K)$.*

According to the principle of parent concepts up and child concepts down, connecting them by lines, we can obtain the Hasse diagram of concept lattice.

In the process of incremental building concept lattice, according to relation between concept and attributes set $f(\{x^*\})$ of new object x^* , we can defined types of concepts as follows[7].

Definition 6. For a concept $C = (O, D)$, if $D \subseteq f(x^*)$, then C is referred to as an modified concept. If C is an modified concept, then C will be modified to be $(O \cup \{x^*\}, D)$.

Definition 7. If $C_1 = (O_1, D_1)$ is a concept, and there does not exist any concept $C_2 = (O_2, D_2)$ such that $D_2 = f(x^*) \cap D_1$, then C_2 is called as a new concept. If for any child concept $C_3 = (O_3, D_3)$ of C_1 , relation $D_1 \cap D_3 = f(x^*) \cap D_1$ does not hold, then C_1 is called as the canonical generator of new concept C_2 .

Definition 8. For a concept $C = (O, D)$, if $D \not\subseteq f(x^*)$ does not hold, then C is called as an old concept.

3 Idea of Bottom-Up Incremental Algorithm

The main idea of the incremental constructing algorithm is as follow: Let $G_i = \{x_1, \dots, x_i\} \subseteq G$, $I_i = I \cap (G_i \times M)$, $G_{i+1} = G_i \cup \{x^*\}$, $I_{i+1} = I \cap (G_{i+1} \times M)$, where x^* is called as adding object. Given formal context $K_i = (G_i, M, I_i)$ and the corresponding concept lattice $L(K_i)$, for formal context $K = (G, M, I)$, the incremental constructing algorithm is dedicated to searching for concept lattice $L(K_{i+1})$ corresponding to $K_{i+1} = (G_{i+1}, M, I_{i+1})$. In the process of adding object x^* , the incremental algorithm has two main problems to be solved: changes of concepts set and changes of Hasse diagram.

Concepts of lattice $L(K_{i+1})$ after adding object x^* , according to definition 6-8, can be divided into three types: new concept, modified concept and old concept. The conditions of finding modified concept and the method of updating a concept are given in definition 6. According to definition 7, there is a one-to-one correspondence between new concepts and generator concepts. So we can generate a new concept through finding a generator concept according to condition 2 of definition 7.

The changes of Hasse diagram include two type changes: one type is building new links between new concepts and its parents or children; the other type is removing links between parent and child which are both linked to a same concept.

After finishing changes of concepts set and changes of Hasse diagram, we can get lattice $L(K_{i+1})$ from lattice $L(K_i)$. The key problem is how to fast change concept set and Hasse diagram. The incremental algorithm proposed in this paper is based on Proposition as follow:

Proposition 1. After adding object x^* into lattice $L(K)$, if there exists a concept $C = (O, D) \in L(K^*)$ satisfying $D \subseteq f(\{x^*\})$, then concept C must be a new concept or modified concept.

Proof: If concept C exists before adding x^* into $L(K)$, by definition 6, concept C must be an modified concept. If concept C does not exist before adding x^* into $L(K)$, by definition 7, concept C must be a new concept. Thus, the proof is finished.

Proposition 2. Given two concepts $C = (O, D)$, $C_1 = (O_1, D_1) \in L(K)$ satisfying $C \leq C_1$, $D \cap f(\{x^*\}) = \emptyset$, of $D_1 = \emptyset$, C_1 must be an modified concept. Otherwise, C_1 must be an old concept.

Proof: If $D_1 = \emptyset$, then $D_1 \subseteq f(\{x^*\})$ holds, by definition 6, C_1 must be an modified concept.

If $D_1 \neq \emptyset$, by $C \leq C_1$, and definition 4, we have $D_1 \subseteq D$. Furthermore, since $D \cap f(\{x^*\}) = \emptyset$, we have $D \not\subseteq f(\{x^*\})$. According to definition 8, C_1 must be an old concept. Thus, the proof is finished.

Proposition 2 tells us that, if intersection of intention of new object and intention of a concept of lattice is empty, then all parents of the concept must be modified concepts or old concepts.

Proposition 3. Given concepts $C_1 = (O_1, D_1)$, $C_3 = (O_3, D_3)$ and $C_2 = (O_2, D_2) \in L(K)$ satisfying $D_1 \subset D_2 \subset D_3$, if $f(\{x^*\}) \cap D_3 = f(\{x^*\}) \cap D_1$, then C_2 is neither a modified concept nor a generator concept.

Proof: Since $D_1 \subset D_2 \subset D_3$, $f(\{x^*\}) \cap D_3 = f(\{x^*\}) \cap D_1$, then $f(\{x^*\}) \cap D_3 = f(\{x^*\}) \cap D_2 = f(\{x^*\}) \cap D_1$ holds. By definition 7, we can see C_2 is not a generator concept.

By $f(\{x^*\}) \cap D_2 = f(\{x^*\}) \cap D_1$, $D_1 \subset D_2$, relation $D_2 \not\subseteq f(\{x^*\})$ holds. And by definition 6, we have C_2 is not an modified concept. Thus, the proof is finished.

Proposition 4. Given concepts $C_1 = (O_1, D_1)$, $C_3 = (O_3, D_3)$, $C_2 = (O_2, D_2) \in L(K)$, satisfying $D_2 \not\subseteq D_1$, $D_1 \not\subseteq D_2$, $D_2 \subset D_3$, $D_1 \subset D_3$, if $f(\{x^*\}) \cap D_3 = f(\{x^*\}) \cap D_1$, then C_2 is neither an modified concept nor a generator concept.

Proof: Since $D_2 \subset D_3$, $f(\{x^*\}) \cap D_3 = f(\{x^*\}) \cap D_1$, we have $f(\{x^*\}) \cap D_3 \cap D_2 = f(\{x^*\}) \cap D_2 = f(\{x^*\}) \cap D_1 \cap D_2$. Since $D_2 \not\subseteq D_1$, $D_1 \not\subseteq D_2$. By definition 3, there exists a concept whose intention is $D_1 \cap D_2$. By definition 7, we can see C_2 is not a generator concept.

The following proof of $D_2 \not\subseteq f(\{x^*\})$ is by contradiction. Assume $D_2 \subseteq f(\{x^*\})$. By $f(\{x^*\}) \cap D_3 = f(\{x^*\}) \cap D_1$, relation $D_2 \cap D_3 = D_2 \cap D_1$ holds. By $D_2 \subset D_3$, relation $D_2 = D_2 \cap D_1$ holds. Furthermore, $D_2 \subseteq D_1$ holds. This is a contradiction from assumption $D_2 \subseteq f(\{x^*\})$. So we have $D_2 \not\subseteq f(\{x^*\})$. By definition 6, C_2 is not an modified concept.

Thus, the proof is finished.

Obviously, By Proposition 3 and Proposition 4 the following two Corollaries can be deduced:

Corollary 1. If $C_1 = (O_1, D_1)$ is a generator of a new concept $C = (O, D)$ and $C_3 = (O_3, D_3)$ is a concept satisfying $D_1 \subset D_3$, then any concept (O_2, D_2) such that $D_2 \subset D_3$ and $D_2 \not\subseteq D_1$ is neither modified nor is a canonical generator of any new concept.

Corollary 2. If (O_3, D_3) is an old concept and $D_3 \cap f(\{x^*\}) = D_1$ (in this case, $(O_1, D_1) \in L(K^*)$ is modified) then any concept (O, D) such that $D \subset D_3$ and $D \not\subseteq D_1$ is neither modified nor is a canonical generator of any new concept.

Based on propositions and Corollaries above, a bottom-up incremental algorithm of building concept lattice proposed in this paper is described as follows:

01: Function AddObject(x^* : new object, $f(\{x^*\})$: attributes which x^* has, L:concept lattice):Concept

02: CardinalSet [i]:= \emptyset ; TagSet:= \emptyset ; C:=L.InfConcept; FindMax:=true; C.InSec:= C.Intent \cap $f(\{x^*\})$;

03: While FindMax do

04: FindMax:=false; Parents := GetParents(C, L);

05: For each PC in Parents

```

06:   PC.InSec:= PC.Intent∩f({x*}); Add PC to TagSet;
07:   If (PC.InSec=PC.Intent) C:= PC; FindMax:=true; Exit For; End If;
08:   End For;
09: End While;
10: If (C.Intent=f({x*})) RootConcept:=C;
11: Else
12:   Add C to CardinalSet[|C.Intent|];
13:   For j:=Size(CardinalSet) to 0 do
14:     For each C in CardinalSet[|j|]
15:       Remove C from CardinalSet[|j|];
16:       If (C.Intent=C.InSec )
17:         For each TC in C. CSet
18:           If (∄Intent of Childs of TC⊆TC.Intent) Add edge: C→TC; End If;
19:           Remove TC from C. CSet ;
20:         End For; Continue;
21:       End If
22:     ParentSet:= ∅; FindGenerator:=true;
23:     For each Parent PC of C
24:       If (PC.InSec=∅) PC.InSec:= PC.Intent∩f({x*});Add PC to TagSet; End If;
25:       Add PC to ParentSet;
26:       If (PC. InSec= C. InSec) NextParent:= PC; FindGenerator:=false; break; End
If;
27:     End For;
28:     If (FindGenerator)
29:       New NC(C.Extent, C.InSec) to L; Add Edges: each TC in C.CSet ; C.CSet :=
∅;
30:       If (RootConcept= ∅) RootConcept:= NC; End If; LinkSup:=true;
31:       For each PC in ParentSet
32:         If (PC.Intent= PC.InSec) OR (PC.InSec<>∅)
33:           If (PC.Intent= PC.InSec) Remove edge: PC→C; End If;
34:           If ( ∃ TC in PC.CSet such that NC.Intent ⊆ TC.Intent) TC:=NC;
35:           Else Add NC to PC. CSet ; End If;
36:           Add PC to CardinalSet[|PC .Intent|]; LinkSup:=false;
37:           End If
38:         End For
39:         If (LinkSup) Add edge: L.supConcept→NC; End If; Add edge: NC→C;
40:         Else
41:           For each TC in C. CSet
42:             If ( ∃ PC in NextParent. CSet such that TC.Intent ⊆ C. Intent ) PC:= TC;
43:             Else if (∄PC in NextParent.CSet such that TC.Intent ⊇ C.Intent) Add TC to
NextParent. CSet
44:             End If
45:           EndFor
46:         End If; C. CSet := ∅;
47:       End For
48:     End For

```

49: End If
 50: Reset *InSec* of all Concepts of *TagSet*
 51: Return *RootConcept*

Function *AddObject* is used to add object x^* to lattice L , and return the greatest lower bound of all modified and new concepts. Two fields, *InSec* and *CSet*, are added to type of concepts. The *InSec* field is used to store intersection of concept intention and $f(\{x^*\})$. The *CSet* field is used to store pointers of new concept. Utilizing *CSet* field, function *AddObject* will pass new concept to its parent and set parent- child link of them. Set *TagList* in function, is used to store concepts whose *InSec* field is changed. At the end of this algorithm, the *InSec* field of concepts in set *TagList* will be reset. The function *Parents*: =*GetParents(Concept, L)* in this algorithm is used to get all parents of *Concept*.

By Proposition 3, if the intersection of parent intention and $f(\{x^*\})$ is equal to the intersection of child intention and $f(\{x^*\})$, then child must not be a generator or modified concept. Furthermore, by Proposition 4, other parents of child must not be generator or modified concepts. Lines 03-09 of algorithm are used to skip and search for concept upward, until find a generator or modified concept.

Lines 03-09 of algorithm are used to search for the lowest modified concept. By definition 6, there does exists modified concept lower than the lowest concept whose intention is included in $f(\{x^*\})$. If there exists a lowest modified concept, algorithm exit and return the lowest modified concept. Otherwise, there must be a new concept. Line 12 is used to put a concept into *CardinalSet*, and Lines 13-47 are used to search for generator and modified concepts in descending order by intention of concept. By definition 6, Lines 16-20 are used to judge whether the concept is a modified concept or not. If it is a modified concept, then the algorithm judge whether the link between modified concept and new concept stored in *CSet* is set or not.

By definition 7, Lines 23-27 are used to judge whether the considered concept is a generator concept or not. If it is a generator concept, then the algorithm generates a new concept in Lines 28-40. Otherwise, the algorithm pass new concept in *CSet* of *Concept* to *CSet* of parent of *Concept*, for the convenience of judgment in future loops.

The main work of algorithm is finished in function *AddObject*. Procedure *BottomUpIncremental* described as follow is used to build concept lattice from formal context by calling function *AddObject*. Lines 02 of procedure *BottomUpIncremental* is used to initialize a greatest lower bound *BottomConcept*, and put *BottomConcept* into original lattice L . Lines 04-07 are used to get object in context one by one, and add each object into lattice L by calling function *AddObject*. Since extensions of concepts are not modified in function *AddObject*, Line 06 updates extensions of all upper concepts of return concept of function *AddObject*.

01: Procedure *BottomUpIncremental* (G, M, I)
 02: *BottomConcept* := (\emptyset , M)
 03: $L := \{\text{BottomConcept}\}$
 04: For each g in G
 05: *ObjectConcept* = *AddObject* ($g, f(\{g\}), L$)
 06: Add g to the extent of *ObjectConcept* and all concepts above
 07: End For

4 Experiment

In order to verify the efficiency of the algorithm, we realize the proposed BottomUpIncremental algorithm and original Godin algorithm [7] by Delphi7. We carried out three experiments as follows on a computer whose CPU is at 2.30 G, memory is 3G, and operating system is Windows XP.

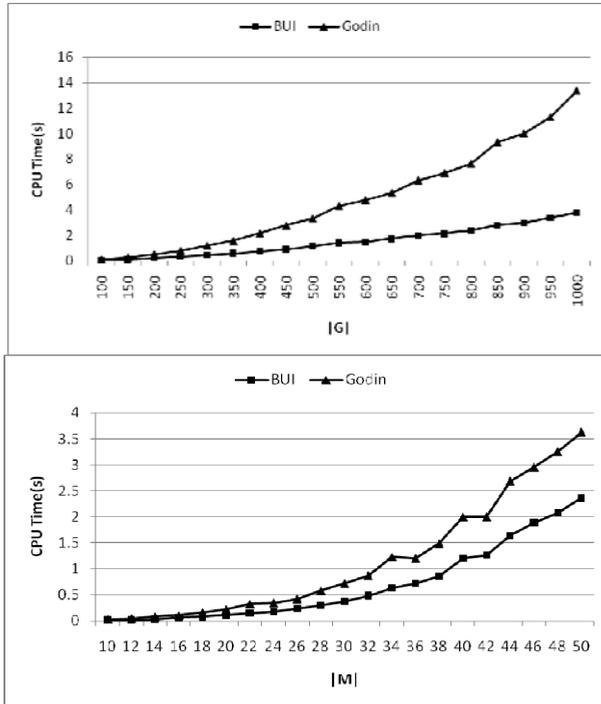


Fig. 1. Time performance comparison

Experiment I, experimental data include 19 randomly generated formal contexts, in which the total number of attributes is 20, the percent of density is 20% and the total number of objects varies from 100 to 1000, with the incremental step being 50 objects. The experiment results are shown in Fig.1. From Fig.1, we can see that in the formal contexts with above experiment data, the proposed algorithm has a slight time advantage than Godin algorithm.

Experiment II, experimental data include 21 randomly generated formal contexts, in which the total number of objects is 100, the percent of density is 25% and the total number of attributes varies from 10 to 50, with the incremental step being 2 objects. The experiment results are shown in Fig.1. From Fig.1, we can see that in the formal contexts with above experiment data, the proposed algorithm has a slight time advantage than Godin algorithm.

5 Conclusion

By analysis to incremental algorithm, some propositions about adding object incremental into lattice are proved according to structure of concept lattice. Based on this, a bottom-up incremental algorithm of building concept lattice has been presented. Because of taking full advantage of parent-child relation between concepts, the proposed algorithm can sharply reduce number of judgments of generator and modified concept. The experiments indicate that the proposed algorithm obtains better efficiency than Godin algorithm, and the time performance in lattice building has been improved markedly with the presented algorithm.

Acknowledgment. This work was supported by the National Natural Science Foundation of China (No. 60804026) and International Cooperation Programs of Henan Province (No. 094300510043).

References

1. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer, Berlin (1999)
2. Maddouri, M.: Towards a machine learning approach based on incremental concept formation. *Intelligent Data Analysis* 8, 267–280 (2004)
3. Shyng, J.Y., Shieh, H.M., Tzeng, G.H.: An integration method combining Rough Set Theory with formal concept analysis for personal investment portfolios. *Knowledge-Based System* 23, 586–597 (2010)
4. Snasel, V., Horak, Z., Kocibova, J., et al.: A analyzing social networks using FCA: complexity aspects. In: *International Joint Conferences On Web Intelligence And Intelligent Agent Technologies*, pp. 38–41. IEEE, Milan (2009)
5. Bordat, J.: Calcul pratique du treillis de Galois d'une correspondance. *Math. Sci. Hum.* 96, 31–47 (1986)
6. Nourine, L., Raynaud, O.: A Fast Algorithm for Building Lattices. *Information Processing Letters* 71, 199–204 (1999)
7. Godin, R., Missaoui, T., Aloui, H.: An incremental Concept Formation Algorithm based on Galois (Concept) Lattices. *Computational Intelligence* 11, 246–267 (1995)
8. Shi, B., Liu, Z., Yu, H.: Algorithm of Building Concept Lattice Based on Its Structure. *Computer Engineering* 33, 12–14 (2007)
9. Shen, X., Han, D., Liu, Z., et al.: Improvement on Constructing Algorithm of Concept Lattices. *Computer Engineering and Applications* 40, 100–103 (2004)
10. Kuznetsov, S.O., Obiedkov, S.A.: Algorithms for the Construction of Concept Lattices and Their Diagram Graphs. In: Siebes, A., De Raedt, L. (eds.) *PKDD 2001. LNCS (LNAI)*, vol. 2168, pp. 289–300. Springer, Heidelberg (2001)

Implementing an Active Database for Maintaining Asset Data

Rozana Kasbon, Sofia Shaharom, Eliza Mazmee Mazlan,
and Saipunidzam Mahamad

Department of Computer and Information Sciences, Universiti Teknologi Petronas
Bandar Seri Iskandar, 31750 Tronoh, Perak, Malaysia
{rozank, elizamazmee, saipunidzam_mahamad}@petronas.com.my,
sophisticated19@gmail.com

Abstract. This paper presents an approach of using an active database to store asset data. The asset here is referring to motors that are being used for the operation of oil and gas electrical equipments. An Electrical Engineering Department of an Oil and Gas company is now working on a move to restore work process improvement in order to resolve the current problem faced, which is the scattered asset database. With such problem in place, it is important to have an integrated database which allows them to track and record numerous motor histories as well as able to produce constructive information with regards to the overhauling of motors. This project aims to develop an active database whereby it can act as a tool to indicate the progress of work (motor overhaul) by using a traffic light system; e.g. green, yellow and red color-coding based on pre-specified criteria.

Keywords: Active Database, Event Condition Action (ECA) rule, Asset Database.

1 Introduction

Keeping scattered data into a common pool like a database has been a common practice in most organizations today. With a proper structuring of the database, information can be easily retrieved and produced in the form of reports. Nowadays, almost all companies rely heavily upon the usage of a DBMS. This work proposed the used of an active database instead of a conventional database system that is considered as passive database management system to store asset information. Asset here refers to motors that are being used in electrical equipments of an Electrical Engineering department of an Oil and Gas Company.

The motors have to undergo servicing activities as part of its maintenance. It includes activities such as bearings changing, cleaning and test run; to name a few. Since these motors are utilized for various applications such as for the cooling system, the pump and to operate compressors as well as auxiliary motors for power generators and conditioning system, it is thus extremely important for the engineers to have a

database system to track the servicing activities and history i.e.; frequency of failure, model type and failure data of these motors.

Currently, the standard practice being used is that, all recording and documenting of work is being done manually in individual excel files. Whenever there is a need for machine / equipment overhaul, request is being done manually by keying in information into separate excel files. This imposes several difficulties especially to the personnel and causes ineffectiveness in operation. Firstly, it has been very difficult and tedious for the personnel to extract important data and figures. For example, whenever there is a need to identify the overhaul history of an individual motor, they will have to find the relevant files and locate it manually. Secondly, personnel will also have to constantly and manually monitor daily progress of motor overhauling. Whenever a motor is being repaired, the personnel will have to manually and constantly check on the system just to ensure that it is within time frame allocated.

This paper presents an initiative in developing an integrated motor active database that not only allows the personnel to extract vital information regarding motor overhauling (i.e.; frequency of motor overhaul, overhaul history of individual motor, quantity of low voltage, LV and high voltage, HV being overhaul quarterly) but also acts as a traffic light tool to identify and indicate the status of the motor work progress by fully exploiting the concepts of an active database.

2 Literature Review

Drag and drop has become a major feature of today's GUI development. The basic sequence is holding the object with mouse button, drag it to the targeted location and drop it by releasing the button. The purpose of drag and drop could invoke various types of actions for creating multiplicity type of association in between objects. It includes dragging a file into different directory to change its position, dropping a data file onto application program for viewing purpose, rearranging the icon in the folder and etc.

Number of errors could be minimized with the present of several methods such as reminder or feedback that reveals the current situation or action. The present of feedback modality [5, 11, 15] has established a performance baseline that provides the ability to correct the error at the same time increase exploration of speed and accuracy.

A database management system (DBMS) is a collection of interrelated data and a set of programs to access those data. The primary goal of a DBMS is to provide a way to store and retrieve database information that is both convenient and efficient. Nowadays, almost all companies rely heavily upon the usage of a DBMS. The vast usage and implementation of DBMS precedes the traditional file-processing system whereby crucial information is stored in various different file [1]. However within a DBMS as well, there are two significant types which are; conventional (traditional) DBMS and active DBMS. Conventional database management systems are passive, in the sense that they only manipulate data in response to explicit requests from applications [2]. On the other hand, unlike any conventional DBMS, an active DBMS

allows users to specify actions to be taken automatically, without user intervention, when certain conditions arise [3].

Active database can also be better defined as a database with the event monitoring scheme for detecting when certain data is INSERTED, DELETED, UPDATED, or SELECTED, and automatically executing the actions in response when certain events happen and particular conditions are met [4]. Traditional database management systems (DBMSs) are passive in the sense that commands are executed by the database (e.g., query, update, delete) as and when requested by the user or application program [2]. However, some situations cannot be effectively modeled by this pattern.

Over the past decade, several research groups have built prototypes that try to incorporate such independence into databases whereby the database will automatically perform an action whenever an event is triggered. Examples include SAMOS [5], ACOOD [6] and HiPAC [7].

A database is used in conjunction with monitoring devices to record and respond to situations outside the database. For example, in an aircraft monitoring database, the following rule adapted from Naqvi and Ibrahim [8] could inform a controller when two aircraft are approaching each other:

```

on update to pos of aircraft
  if exists
    (select *
     from aircraft Other
     where distance (Other.pos,new.pos) < 5000  and
           distance (Other.pos,old.pos) > 5000);
do (send message to controller)

```

In this example, the event being monitored is the position of an aircraft communicated to the database from an external device, and the action taken is a change to a display that the air traffic controller is monitoring. Both the new value and the old value for the pos affected by the event are accessed from within the condition.

There are also systems that are using conventional database to store profile information but rely on Multi Agent System (MAS) to provide an adaptive system such as one that is discussed in [9].

As previously mentioned, among the problems faced by the department personnel is that there are no means to indicate or notify them on the status of the motors. Because an active database is capable to trigger an action with result to an event [10], this unique feature can be used to build the tool for the personnel to evaluate on the status of the motors and act as an indicator should the overhauling motor reach the deadline. The tool in this context would be the traffic light tool explained earlier. The Event-Condition-Action (ECA) feature of an active database can be applied and further enhance to ensure that the tool acts as an indicator / reminder / to the personnel on the status of motor overhaul.

3 System Design

The system is developed with four main functions which are; recording motor details function, report generation function, search engine and finally the traffic light tool. Fig. 1 describes the architecture of the active database used in the traffic light tool. The event in this case refers to when the system retrieves the current date. The current date will then be compared to the expected completion date and the condition of the motor overhaul progress will be determined; whether the motor is within, approaching or have exceed the deadline. Based on the condition, the database can then proceed to perform the action whereby to display it as Green, Yellow or Red depending on the motor work progress. Green means that the motors overhaul is within the specified time frame (i.e.; 10 days); the yellow indicates that the motor overhaul is approaching deadline (i.e.; 1 – 2 days) and finally the red means that the motor has passed the allocated time.

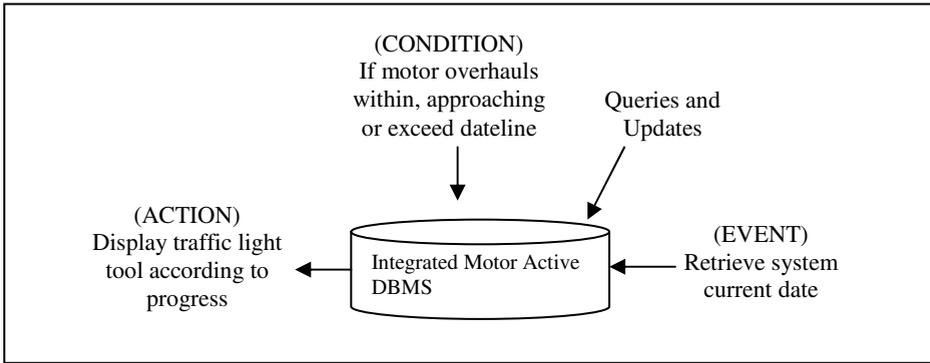


Fig. 1. Active Database Architecture

Fig. 2 depicts how the system implements the active features as well as the various conditions used to categorize the motor according to the motor work progress.

The flowchart shows the condition that specifies what conditions has to be met in order to set the status of the motor work progress. Relating back to the ECA rule of the active database concept, the event that occurs is when the traffic light tool is loaded. The system will then compare the motors' expected completion date to the current date. Depending on the condition, the set status action will be executed. For example, when the motors' expected completion date is less than or before the current date, therefore the system will automatically update the motor's status as 'exceed' and will then display motor in the 'red' category. On the other hand, if the period of days between the motor's expected completion date and the current date is less than 2, therefore the motor's status will be set to approaching and the system will display it as 'yellow'. Finally the motor status will be set to within the deadline and displayed as 'green' if the expected completion date is just greater than the current date. The form is shown in Fig. 3.

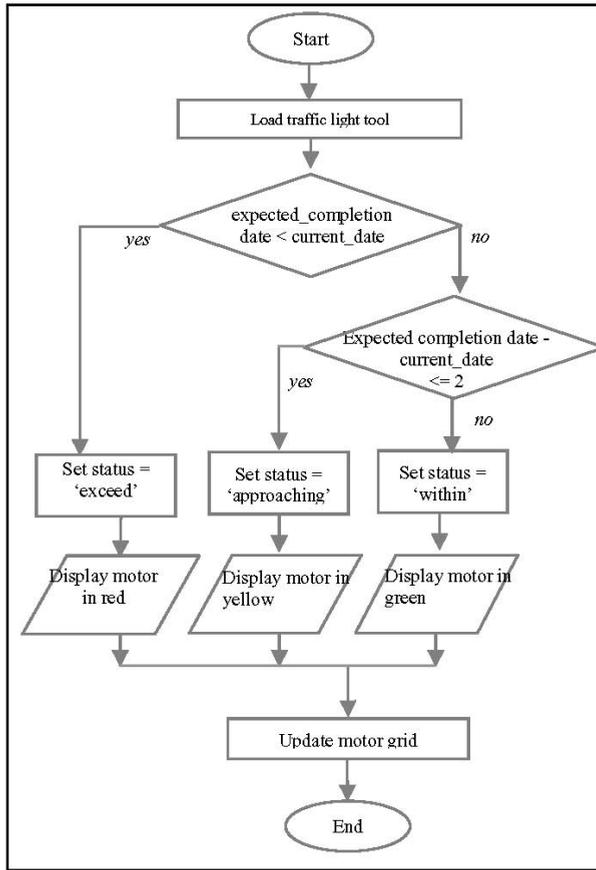


Fig. 2. Flowchart to determine Motor Work Progress Status

For the first function, prior to recording all motor work details, users must first select whether it is a new or an existing motor. If it is a new motor, the system will create a new motor record in the database whereas if it is an existing motor, the updated details will be recorded for the selected motor. Fig. 4 is the form used to gather information for new motor;

The second function which is the report generation function allows the users to extract information based on three factors; quantity of motor overhauled, overhaul history of individual motors and the quantity of low voltage and high voltage motors overhauled quarterly. Fig. 5 (a), Fig. 5 (b) and Fig. 5 (c) depict the forms used for every factor respectively.

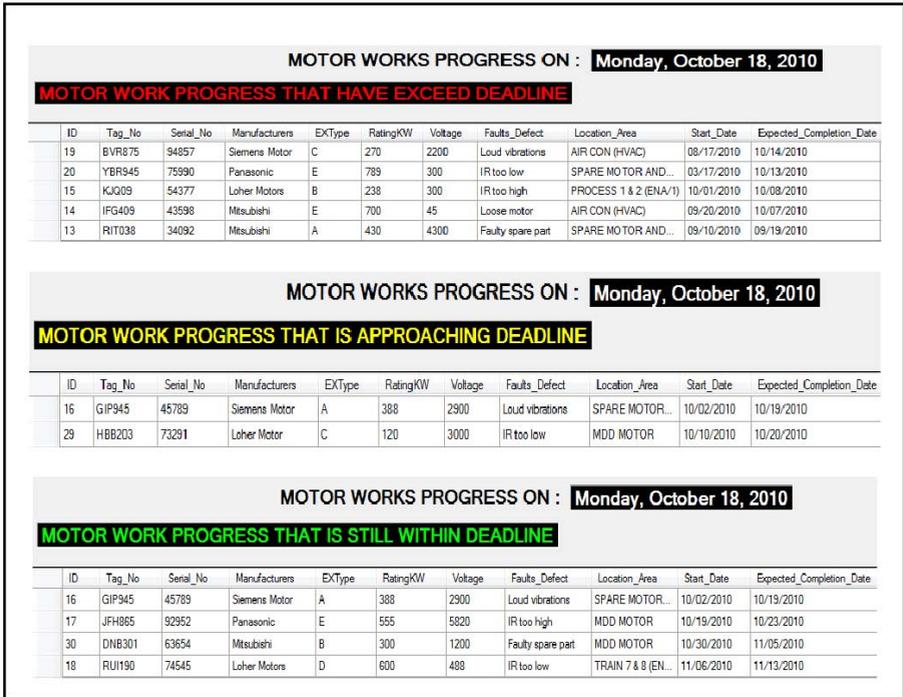


Fig. 3. Motors – Exceed, Within and Approaching Deadline

ELECTRICAL WORKSHOP MOTOR, TRANSFORMERS & AC/DC SOLINOID COILS OVERHAUL/REWINDING TECHNICIAN

NEW MOTOR OVERHAUL REPAIR OR REWINDING REQUEST

Equipment Tag : C11456

Equipment Serial No : 5734

Make / Manufacturer : Loher Motors

Equipment EX Type : 0

Equipment Rating (kW) : 947

Equipment Voltage (V) : 1500

Area : SPARE MOTOR AND STATOR

Faults / Defects : Faulty spare parts

Start Date : 10/19/2010 Expected Completion Date : 10/22/2010

Type of Work to be Performed : Overhaul

CONDITION BEFORE MOTOR OVERHAUL

Insulation Resistance (IR) : 200 Col Resistance (CR) : 150

CONDITION AFTER MOTOR OVERHAUL

Insulation Resistance (IR) : 180 Col Resistance (CR) : 130

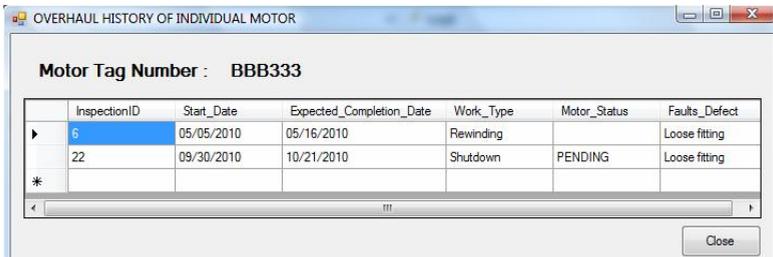
Vibration Level : 0.07

Continue Clear Cancel

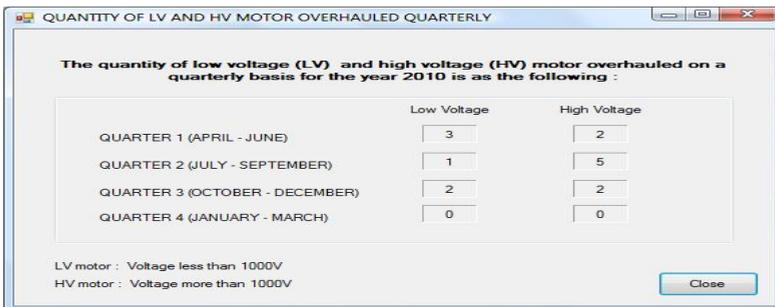
Fig. 4. Form to Add New Motor



(a) Motor overhauled monthly



(b) Overhaul history of individual motors



(c) Quantity of Low Voltage and High Voltage Motors Overhauled

Fig. 5. Report generation function

Apart from that, users can also search the important details associated to every motor such as the motor serial number, name of manufacturer, motor EX type, motor rating as well as motor voltage. By using the search engine function, users must only select the motor tag number and all the information mentioned earlier will be displayed as shown in Fig. 6.



Fig. 6. Motor Search Engine Function

4 Conclusion

This paper highlights the design of an active database is used to develop an Integrated Motor Active Database for maintaining motor asset for the Electrical Department of a company. This project combines the research on the active database pedagogy and prototype building aspect of an Integrated Motor Active Database which goal is to allow the department personnel to extract vital information regarding motor overhauling and to have a tool to monitor motor work progress.

References

1. Navathe, S., Tanaka, A., Madhavan, R., Gan, Y.H.: A Methodology for Application Design Using Active Database Technology. Technical Report RL-TR-95-41, Georgia Institute of Technology (1995)
2. McCarthy, D., Dayal, U.: The Architecture of an Active Database Management System. In: Proceedings of the 1989 ACM SIGMOD International Conference on Management of Data, pp. 215–224 (1998)
3. Comai, S., Tanca, L.: Termination and Confluence by Rule Prioritization. IEEE Transactions on Knowledge and Data Engineering 15(2), 257–270 (2003)
4. Tan, C.W., Goh, A.: Implementing ECA Rules in an Active Database. Knowledge-Based Systems 12(4), 137–144 (2001)
5. Geppert, A., Gatzju, S., Dittrich, K.: Architecture and implementation of the active object-oriented database management system SAMOS. Computer Science Department, University of Zurich (1995)
6. Berndtsson, M.: ACOOD: An Approach To an Active Object-Oriented DBMS. Department of Computer Science. University of Skovde, Sweden (1991)
7. Dayal, U., et al.: The HiPAC project: combining active databases and timing constraints. ACM-SIGMOD Record 17(1), 51–70 (1998)

8. Naqvi, W., Ibrahim, M.: Rule and knowledge management in an active database system. In: Proceedings of the First International Workshop on Rules in Database Systems, pp. 58–69 (1994)
9. La Gennusa, M., et al.: Multi-Agent Systems as Effective Tools for the User-Based Thermal Comfort: an Introduction. World Applied Sciences Journal 10(2), 179–195 (2010)
10. Paton, N.W., Diaz, O.: Active Database Systems. ACM Computing Surveys 31(1), 1–3 (1999)

The Design and Implementation of a Gesture-Driven System for Intelligent Wheelchairs Based on the Orientation Histogram Method

Yuan Luo and Yu Xie

School of Optoelectronic Engineering Chongqing University of Posts and Telecommunications
Chongqing China

luoyuan@cqupt.edu.cn, cxieyuq@163.com

Abstract. An approach of hand gesture recognition, setting the orientation histogram of the picture as the characteristic vector of hand gesture, is discussed in this paper. It can decrease the influence of light changes during the process of recognition effectively. A gesture-Driven system for intelligent wheelchairs is also introduced in the paper. Experimental results show that the method is robust and accurate.

Keywords: Hand gesture recognition, orientation histogram, intelligent wheelchair.

1 Introduction

The interaction based on hand gesture is an important aspect in human-computer interaction. At present, hand gesture recognition is mainly divided into dynamic trajectories recognition and static gestures recognition. For the dynamic trajectories recognition, it is mainly based on Adaboost algorithm [1], Hidden Markov Model (HMM) and Dynamic Time Warping [2]. For the static hand gesture recognition, it is mainly based on neural network algorithm and elastic chart matching method. But this paper adopted the orientation histogram, which is robust of light to recognize hand gesture, for dynamic and static recognition methods are all sensitive to light changes. Firstly, because of operating in the local grid cell of the image, this method has a good invariance for its geometric and optical image deformations, which only occur in larger areas of space [3]. Secondly, on the condition of the rough spatial sampling, fine directional sampling and a strong local optical normalized, some slight body movements are available. Because they can be ignored without affecting the test results as long as hands keep a stable position relatively.

2 Hand Gesture Recognition Based on the Orientation Histogram

Orientation histogram is a kind of feature descriptor which is used to detect objects in computer vision and image processing. This method uses the gradient direction characteristics of the image, as the same as the edge orientation histogram method,

SIFT descriptors, and context shape method. Its characters are calculating in a grid unit with dense grid and uniform size, and using the overlapping local contrast normalized method to improve accuracy [4]. Navneet Dalal and Bill Triggs, from the French National Institute of Computer Technology and Control, published the paper about the histograms of oriented gradient first on CVPR in 2005.

Set the Video Streaming acquired by the USB camera as input to obtain the value of each pixel. Calculate its direction according to the gradient operator, and express it by the local orientation angle θ , which is defined as the pixels intensity difference of image in the horizontal and vertical direction. It is a function about the position (x, y) [5].

$$\theta(x, y) = \arctan[I(x, y) - I(x - 1, y), I(x, y) - I(x, y - 1)] \tag{1}$$

$I(x, y)$ is the gray value of the image point (x, y) .

Taking the gradient amplitude of each pixel as weight, count and get the histograms of oriented gradient of each unit. So we can get vector Φ with N variables, $N_i \in (360^\circ(i-1/2)/N, 360^\circ(i-1/2)/N)$.

$$\Phi(i) = \sum_{x,y} \begin{cases} 1 & \text{if } |\theta(x, y) - \frac{360^\circ}{N} i| < \frac{360^\circ}{N} \\ 0 & \text{else} \end{cases} \tag{2}$$

The value of N is another problem needed to be considered. Different experiments with different N , 20, 24, 36, 40 respectively, were done in this paper. To get enough information and reduce the amount of calculation, we take $N = 36$ finally.

Normalize each unit of the histogram of oriented gradient to reduce the influence of illumination. And we can get one thirty-sixth of the histograms of oriented gradient. Constitute the characteristic vector of gestures by fitting all the histograms of oriented gradient in detecting window together. Then match it with that of the training sample. Φ_1 and Φ_2 represent the characteristic vector of present gestures image and one gesture in sample library, respectively. The smaller Euclidean Distance $(\Phi_1 - \Phi_2)^2$ is, the effect of matching is better.

This method is robust to illumination, as shown in Fig. 1. Set (a) as the training sample, and (b) is the hand gesture needed to be recognized. (c) and (d) are their corresponding orientation histograms. We can see the orientation histograms in different illumination are very similar.

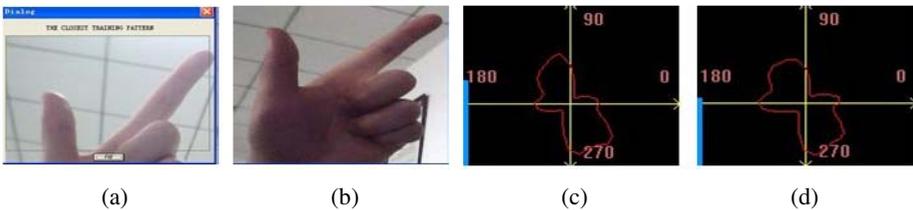


Fig. 1. Recognition effect under different illumination and their histogram

3 Human-Computer Interaction System Based on Hand Gesture Recognition

System flow chart for realizing human-computer interaction system based on hand gesture recognition is shown in Fig. 2. Specific as follows:

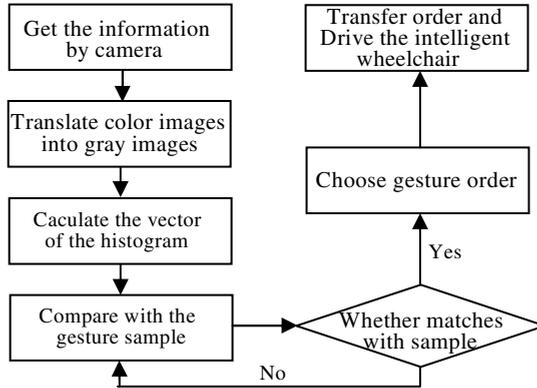


Fig. 2. System flow chart

1. Set a sample library of hand gestures. Parts of samples are shown in Fig. 3. In this paper, general USB cameras are used to acquire decades of images, with 640*480 in size. The sample library contains 5 basic commands that control the wheelchair. They are forward, backward, turn left, turn right and stop, respectively.

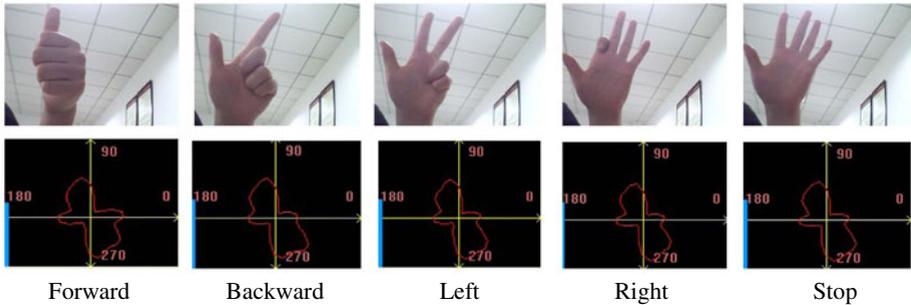


Fig. 3. Gesture samples

2. Translate the input image into a gray image, and calculate its vector of histogram. Then count its oriented histogram.

3. Match the input image with the characteristic vector of the gesture in sample library. Set the Euclidean Distance as measurement scale, and the match effect will be much better when the value of Euclidean Distance is smaller.

4. Transfer the information of hand gestures into drive signal through wireless network connected to the intelligent wheelchair, so as to control the intelligent wheelchair. Implementation of methods mentioned above is shown in Fig. 4.



Fig. 4. Controllment the intelligent wheelchair

4 The System Realized in Intelligence Wheelchair

We use the normal camera, Intel E5200 1G memory laptop. Intelligent wheelchair communicates with the PC machine through wireless network. In particular, control signals of PC machine are sent to the upper system based on ARM9 processor of intelligent wheelchair. The ARM9 processor gets the contact with DSP driver modules by RS232 Serial Bus, so as to complete the controlling of the wheelchair's movement.

Table 1. Recognition rate of hand gesture

| Basic instructions | Test number | Correct number | Recognition rate (%) |
|--------------------|-------------|----------------|----------------------|
| Forward | 30 | 27 | 90 |
| Draw back | 30 | 28 | 93.33 |
| Turn left | 30 | 28 | 93.33 |
| Turn right | 30 | 24 | 80 |
| Stop | 30 | 29 | 96.67 |

Experiment shows that this method by the histogram of oriented gradient can effectively recognize hand gestures, and can satisfy the requirement of real-time. Its recognition rates are shown in Table 1.

5 Conclusion and Prospect

To control the movement of the intelligent wheelchair, an approach of hand gesture recognition based on the histogram of oriented gradient is discussed in this paper. The method is broadly divided into three steps: Firstly, create a sample library of gestures, and definite each gesture. Secondly, start the wheelchair, and establish a connection to PC machine by wireless network. Finally, compare the hand gestures with the sample library, so as to control the wheelchair.

Because this method calculates the number of pixels in each orientation of graphic image, it has little effect when illumination changes. Meanwhile, this method has the

advantages of less calculation and common in hardware, by which we can easily set up a real-time gesture control system.

Of course, there are still some shortcomings. For example, there will be some difficulty in recognizing of hand gestures, if there is an oblique angle between the hand posture and the camera. To overcome this problem, we introduce BP neural network method to train the hand gestures sample, which is our next research direction.

Acknowledgment. This paper is supported by International Science & Technology Cooperation Program of China 2010DFA1160.

This paper is also supported by the Foundation for Science and Technology Research Project of Chongqing (CSTC:2010AA2055).

References

1. Viola, P., Jones, M.: Robust Real-Time Face Detection. *International Journal of Computer Vision* 57, 137–154 (2004)
2. Moni, M.A., Ali, A.B.M.S.: HMM Based hand gesture recognition: A review on techniques and approaches. In: *IEEE International Conference on Computer Science and Information Technology*, Bangladesh, pp. 433–437 (2009)
3. Sun, L., Zhang, L.: Static Sign Language Recognition Based on Edge Gradient Direction Histogram. *Microelectronics & Computer* 27, 148–153 (2010)
4. Dalal, N., Triggs, B.: Histograms of Oriented Gradients for Human Detection. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, France, vol. 1, pp. 886–893 (2005)
5. Liu, X., Jiang, C., Feng, H.: Recognition of Chinese sign Language Based on Orientation Histogram. *Computer Emulation* 26, 244–247 (2009)
6. Geng, L., Wu, X., Peng, Z.: Gesture Recognition Based on OHV and Neural Net. *Signal Processing* 42, 339–342 (2005)

A Flash-Based Framework for Learning Thai Language as Second Language in Preschool Education

Saipunidzam Mahamad, Yossawadee Kaoemanee, Mohammad Noor Ibrahim,
and Rozana Kasbon

Department of Computer and Information Sciences,
Universiti Teknologi PETRONAS

Bandar Seri Iskandar, 31550 Tronoh, Perak, Malaysia

saipunidzam_mahamad@petronas.com.my, flumini48@gmail.com,
mnoor_ibrahim@petronas.com.my, rozank@petronas.com.my

Abstract. Learning a national language as second language has become an apprehension in everyday life for a certain group in the country. The emergence of multimedia technologies with appropriate learning approach could enhance the learning process of a particular language. This paper presents a framework for the design and implementation of an alternative web-based learning game that aims to encourage the kindergarten learner to learn Thai as their second language. It comes about to expose the young children to the technology while they learn from the lessons and play the games which is the mechanism to deliver the educational content. The concept play and learn is implemented in this system so that students are engaged both their physical and mental skills to solve problems through the interactive activities provided by the system. The system in its final state will be a flash-based educational game integrating Thai preschool syllabus.

Keywords: Flash-based Educational Game, preschool education, play and learn approach.

1 Introduction

Preschool education is considered as a foundation of all learning [1]. It is a vital period for a person to develop the cognitive, physical, emotional, and also the social skills. Most children start their preschool education at the age of 3 years old. At the preschool, the children will learn new things and gather new knowledge from the activities that prepared by the preschool teachers based on the certain approaches applied by the school. It has been realized the necessity to build a strong basic education at the preschool level to have more chance to be succeeded in the higher level of education [2].

In Thailand, the preschool education is considered as a non-formal education service. Most of the public kindergarten school has advocated the child centered approach and some have practiced learner-centered teaching. The use of integrated curriculum in teaching traditional courses occurs primarily through projects and learning centers, which reflect children's interests and suggestions. Teachers guide children's involvement in projects and enrich the learning experience by extending

children's ideas, responding to their questions, engaging them in conversation, and challenging their thinking. The children are required to have an intermediate communication skill especially to speak in Thai official language in order to proceed with the formal education at school. Otherwise, the children will face difficulty in learning other subjects and developing other skills.

People throughout Thailand speak in different dialects based on to the region where they live. Although there are many differences in terms of vocabularies, each dialect still has some similarities to the official language of Thailand. However, in southern part of Thailand i.e. Yala, Pattani, Naratiwat and Songkhla, there are over 80% of Malay ethnic that are using Thai language as their second language. An observation in [3] found the residents are using Malay language with Patani dialect in their daily communication. The different in daily language may barricade children in learning at preschool, which later may affect the formal education at the higher level. This group of children has to learn Thai as their second language; consequently, faced difficulty in communicating and representing their experiences, feeling and emotion to the teachers and friends when they are in classroom. Furthermore, this also will affect their performances as compared to other students who their native language is Thai.

The challenge is not only arise in children but also with the preschool teachers who have difficulties in interacting with and understanding the children. In daily school activities, a teacher has to spend more time to interact and assist this particular group of children in learning Thai language. The teachers has to repeat the vocabularies in order to help them memorizing the vocabularies which cause the other group of children tend to loss attention after a while because the teacher cannot interact with all of them at once.

Preschool education plays a very important role in developing knowledge-based society since the early ages. Educators have to put in efforts to ensure the delivery of quality education to the society. Technologies such web and multimedia has proven to have a bright future and has shown significant improvement in sustaining student's interest [4, 5]. These technologies have been widely used to improve the efficiency and effectiveness of the teaching and learning process. It will spur major changes in today's education. Web and multimedia allows student to go beyond the limit of traditional education approaches. It creates an active learning model enriching the learning process. Teacher will become the facilitator; facilitate the student leading to the learning path rather than traditional approach of being the main sources of information and understanding. Integrating the current multimedia technology in education such as text, graphic, audio, video, and animation has increased the interest of learning and help learners to retain the knowledge learned.

This paper proposes an interactive multimedia system for learning Thai language as the secondary language. It integrates the Thai preschool syllabus with flash-based approach and game applications to attract and sustain the interest of the preschool students. The aim of this research is to encourage students to learn essential Thai vocabularies, which are corresponding to the curriculum. This is very important to prepare them for their education at the higher level. However, the proposed system does not meant to replace the traditional way of teaching and learning in classroom but as a supporting learning tool that creates more fun and excitement for students to learn.

In this paper, an overview of the design of the system is presented. The next section, this paper reviews the related works regarding the studies of children development, educational games and Thai preschool approach, in particular, the effects of current

technologies to the environment of learning. The methodology and conceptual framework of the project are discussed along with the result, discussion and conclusion in the last section.

2 Related Work

It is commonly accepted that the early years of life is the important period for learning. Children at the ages of 3 to 6 years have rapid growth in every aspect and able to learn at an exceptionally fast rate [2, 4]. The way children learn is always referring to 'natural learner'; learn from the surrounding environment i.e. family, teachers, preschool peers and everything they see. This fast progress seems keep surprising the parents and teachers all the time. However, with the limited attention span, preschoolers are easily to loss concentration [4]. No doubt, to attract their attentions and to make them concentrate to whatever planned activities are not an easy task for the teachers.

The nature of young children is keen on playing and learning new things. The importance of playing in this period of life has been recognized by the educators. Thus, the concept of 'play and learn' is often used as a medium to promote children's learning. While playing, young children in a way will acquire various knowledge and skill [6]. On the other hand, if the activities are fun, they will give active participation and full concentration.

Different multimedia elements provided by today's technology are also one of the new inspiring things that all of the children are trapped into it. Through the use of multimedia, animations in particular, it is easy to engage children's motivation and attention [5].

For decades the advancement of the technology entails a vast change in many aspects of human life. The Internet connection and information and communication technology have facilitated not just the communication but also education. Today's children are connected to a world large library which they can get access to it by using only their fingertips. They are born in a computer environment and explored the computer at a very young age [7].

In teaching young children, conveying the learning contents in an enjoyable activity is needed in creating the atmosphere of learning. Computer and Online games are among the popular tools which are being used as educational games meaning the combination of entertainment and education. However, a full advantage will occur only when the educational content is well integrated into the structure of the game [8]. This technology, in addition, can transfer an extensive educational content to the students too. The study revealed that educational games can contribute to the positive and better academic achievement of children [9].

Furthermore, with the help of today's interactive technology, educational games are far more powerful. The environment which encourages student to learn by doing and visualize the difficult to understand concept is also simplified, accordingly enhances student participation and motivation [10]. The study has proved that the use of computer technology help in developing the language and literacy development by engaging the children with longer and more complex speech. The study also shows that with computer, the social interaction can be elicited more compare to the use of some traditional class room games e.g. block building and puzzle. This will support the

development of proposed system which the target audience of the games is the young children who use Thai as their second language.

Interactivity is another important element in learning that can draw learner's attention toward what they are being taught. It is hardly to achieve through the use of normal text or exercise books since the books cannot have two-way communication with the learners. Research indicates that people are able to memorize 20 percent of the content that they see, increase to 40 percent if they see and hear but if they can see, hear and do simultaneously the amount of content that people are capable to memorize increases up to 75 percent [11]. This research also supported by another study which reveals that the students can gain better academic achievement when they learn from visual and auditory information compare to those receiving only pictures and text from the normal text books [12,13].

From all the premises above, flash software then become a first choice for computer-based educational game developers as it excels in all the aspects needed to produce animation, multimedia and games [14]. Furthermore by allowing the combination of animation, speech, sound and games in one, the flash lessons and games will be able to reach more and variety learning style of children[15] .

In the past, teacher-centered approach was the teaching style whereby the teachers are the center of attention. The students only learn passively to memorize but not thinking. As a result, most of the students end up with lacking the critical thinking skills. Thus to improve the quality of education, Ministry of Education of Thailand decided to reform the system by changing from teacher-centered to child-centered or learner-centered approach. However, this is still being a big challenge in order to fully practice and implement the approach [1, 16].

Child-centered or learner-centered approach focuses on the children's cognitive and emotional development. It has an environment that encourages children to initiate their own learning which they can choose what to learn according to their preferences and continue learning on their own pace [1]. Further more in such environment, the children needs, and opinions are taken into consideration. The role of teacher in this approach is as a facilitator who guides children and helping them in learning what is best for them [16].

The proper integration of technology such as computer and Internet can be use effectively to create the learning environment for the child-centered approach. With technology, students are engaged both their physical and mental skills to solve problem through the interactive activity. They can have more control over their learning content, learn by doing, receive feedback and refine the understanding when they want. [16].

3 System Framework

The conceptual framework comprises two main stages is shown in Fig. 1. The first stage is the early analysis, which consists of problem identification and objective determination. The second stage is the application development. The main objective of this proposed system is to encourage students to learn and memorize necessary vocabularies that will be using quite often in the classroom. Thus, the main language of the alternative learning game is Thai language. The approaches and lessons are based on the Thai preschool syllabus. Thus, the content of the lessons and games are age-appropriate.

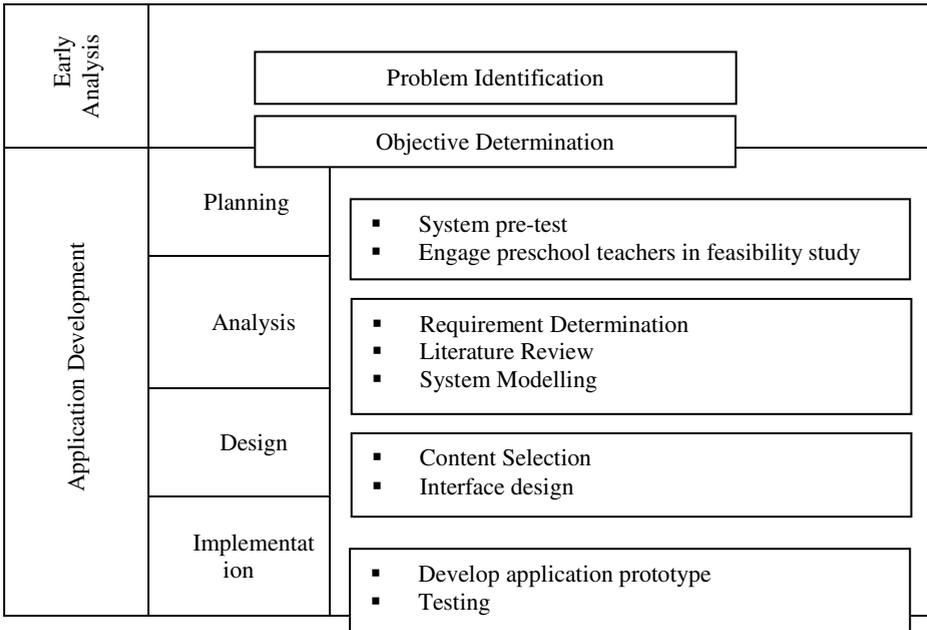


Fig. 1. System conceptual Framework

Fig. 2 shows the UML diagram the interaction of students and teacher with the proposed system. The student has an access to the five module developed in alternative learning game. Each of the module, the system will capture the progress and the performance of the student and allow the teacher to monitor and view the result of each learning activities.

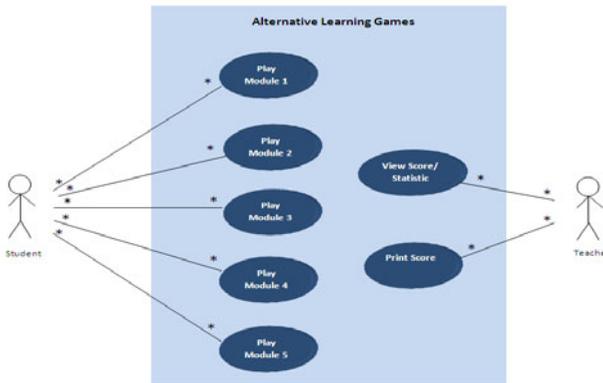


Fig. 2. Use Case Diagram of Alternative Learning Game

The design of the system architecture for Alternative Learning game is shown in Fig. 3. The users i.e. the teachers or preschoolers start the communication with the system by accessing the flash application through the web-browser. Flash application will then sends a request or data to the database by the assistance of PHP which is a server side script language.

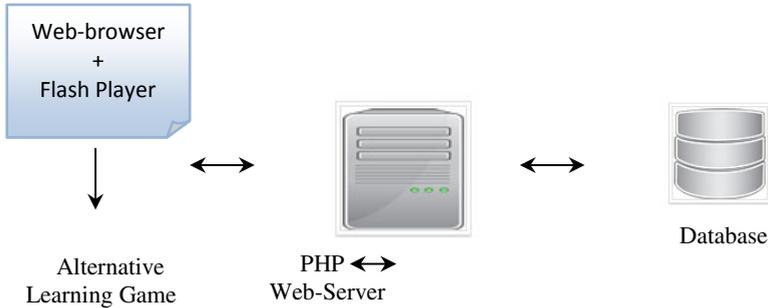


Fig. 3. Alternative Learning Game Architecture

4 Discussion

An observation and interview have been conducted to discuss further problems regarding teaching and gaining feedback from students whom cannot speak Thai fluently. It is aim to gather information on approach being use by both teachers in handling these problems in order to develop the proposed system. The Baan Najuak School in Songkhla Province was selected which is one of the part that faced the problem of learning Thai language as the second language. The session was conducted with two kindergarten teachers as the interviewees. The first teacher was Mrs. Narin the kindergarten teacher with 30 years teaching experience in the kindergarten at many schools of Songkhla Province. The second was the assistant teacher, Ms. Pausiyah with five years experience in assisting Mrs. Narin at Baan Najuak School.

Based on the interview session, the problems of students who cannot speak Thai fluently will cause a delay on the student learning and complete most of the assigned activity slower than others. The students tend to see what the other students do and will start to follow. The students again face difficulties in expressing their opinion, feeling and experiences when the class has storytelling activity.

The teachers start introducing the Thai language by showing some element to children and then translating the student dialect into Thai language. This situation somehow has caused a teacher has to learn the dialect in order to be at par with the student and has found with knowing and understanding the dialect makes learning process easier on interacting with students in introducing the Thai language. Besides, the teacher will need is to use and repeat the similar sentence in asking and

communicating with the student. With this method, it allows the student to remember the simple pattern and a ways to begin the communication with others. Apart from that, the games, songs and also fairytales also being use to gain students attention in learning. The students will know new vocabularies from those activities without realizing that they are learning. On the other part, the teacher will use computer in teaching the movement and rhythmic sometime.

5 Conclusion

The research reveals that the education can be benefited from the extensive multimedia technology. The technology is not only support but also increase the opportunity of learning. Alternative learning games proposed a new way of learning Thai language using the flash-based educational game as a medium to expose children to the emerging technology in learning while playing. Pedagogy concept currently adopted in the Thai preschools was taken into consideration during the design and development of proposed system. As a result, the lessons and game is age-appropriate and match the abilities of the children.

References

1. Israsena, V.: Thai Teachers' Beliefs about Learner-Centered Education: Implications for Success for Life Thailand. University of North Texas (2007)
2. Holt, J.: How Children Learn. Pitman Publishing Company, New York (1995)
3. Isaya, K.: The Usage of the Jawi Alphabet System in Patani Malay: A Challenge to Malay Society in Endangered Languages. Songklanakarin Journal of Social Sciences and Humanities 6, 99–136 (2010)
4. Karuovic, D., Radosav, D.: Educational Software for Preschool Children, pp. 291–295. IEEE (2009)
5. Pohradsky, P., et al.: Application of ICT in Pre-school Education. In: 52nd International Symposium ELMAR 2010, pp. 159–162 (2010)
6. Kotani, T., et al.: An Analysis on the Process from Play to Learning, pp. 231–234. IEEE (2004)
7. Verenikina, I., et al.: The affordances of computer play in young children: A preliminary study. University of Wollongong (2008)
8. Salom, M.F.: Making Educational Game Educational. MediaKidz Research & Consulting (2005)
9. Rosas, R., et al.: Beyond Nintendo: design and assessment of educational video games for first and second grade students. Computer Education 40, 71–94 (2003)
10. Young, M.E.: Technology and Early Child Development, pp. 13–14. Knowledge Enterprise, Inc., TechKnowLogia (2001)
11. Neo, M., Neo, K.T.K.: Innovative teaching: Using Multimedia in a problem-based learning environment. Educational Technology & Society 4 (2001)
12. Haughey, M., Muirhead, B.: The pedagogical and multimedia designs of learning objects for schools. Australian Journal of Educational Technology 24(1), 470–490 (2005)
13. Mayer, R.E., Moreno, R.: Animation as an Aid to Multimedia Learning. Educational Psychology Review 14(1), 87–99 (2002)

14. Zhang, H.Y.: An Educational Flash Game for Preschool children. In: The 1st International Conference on Information Science and Engineering (ICISE 2009), pp. 3207–3209 (2009)
15. Lamb, A., Larry, J.: Flash: engaging learners through animation, interaction and multimedia. *Teacher Librarian* 33 4, 54–56 (2006)
16. Thamaraksa, C.: Student-Centered Learning: Demystifying the Myth. *SLLT*, vol. 12, pp. 59–71. Mahidol University (2003)

Emergency Personnel Evacuation Model and Its Application Based on Cellular Automaton

Demin Wu^{1,2}, Wenyong Dong¹, and Jiangpeng Wang³

¹ School of Computer, Wuhan University, Wuhan, China

² The College of Continuing Education, Wuhan University, Wuhan, China

³ Institute of Electronic Information, Wuhan University, Wuhan, China

{woodmean, Hubei_001}@163.com

Abstract. In this article, cellular automaton model will be extended and applied to the evacuation of large public venues. First, we analyze the affecting factors in macro approach, such as physical distribution, psychological preferences, bypass behaviors and so on. Then, we apply the model in the process of evacuation in China Pavilion of World Expo after we have established an extended cellular automata model and proved its reliability. Simulation results show that people who do not consider bypass apparently stand in queues during evacuation, and the herd mentality, in general, is to hinder evacuation process to a certain extent, and it is worth noting that a certain percentage of the herd mentality could shorten the evacuation time, and additionally, family behaviors will extend the evacuation time. Finally, according to experimental results, we put forward rational proposals for the personnel evacuation in large public venues and optimize the model.

Keywords: cellular automata, personnel emergency evacuation model, public safety.

1 Introduction

In recent years, accidents in public places caused by fire, earthquake, terrorism and other unexpected factors are not rare to see, and emergency evacuation of staff performance has become a critical factor in the safety design of large public venues. Currently, the macro and micro approaches both exist in analyzing the general evacuation model. The macro approach takes the overall movement of the crowd as the object, ignoring individual tendencies and identities. That is, personnel are regarded as a continuous flow of media, and the principle of fluid mechanics is applied to establish personnel evacuation model [1]. The micro approach considers thoroughly individual tendencies and identities and, accordingly, constructs evacuation models [2], as is typically shown in the social force models proposed by Helbing and his colleagues [3].

Following the evacuation process in large public venues projected by the extended cellular automata model, this paper simulates possible unwanted situations against the background of Shanghai World Expo and puts forward rationalized proposals after those simulation results have been tested and verified in the emergency evacuation performance at the China Pavilion.

Cellular automata is a distributed and space scalable system consisting of bulk simple individuals through local contact. It is a dynamic system derived from the finite state variables (or cell) in the uniform grid [4]. Its strong points lie in the revelation of the macroscopic behavior occurred naturally through simple, local microscopic principles, and it describes many physical phenomena of nonlinear differential equations without the need to create complex [5]. In this paper, cellular automata model is used to simulate an emergency evacuation process in large public venues, and, the established two-dimensional stochastic cellular automata model is used in the study of special psychology and behavioral simulation in the state of evacuation.

2 The Establishment and Extension of Cellular Automaton Model

Cellular Automaton Model. The basic components of cellular automaton include basic cell, lattice, neighbor and rule. Cellular automata can be considered as a lattice and the transformed function defined in the space can be expressed by a four-tuple [6]:

$$C = (D_n, S, N, f) \quad (1)$$

In this formula, D_n is a positive integer, standing for cellular automata dimension, namely n Dimensional type space; S is a finite set of states, and the state of cell at the moment of t corresponding to Lattice site r on can be expressed as:

$$S(r, t) = \{S_1(r, t), S_2(r, t), \dots, S_k(r, t)\} \quad (2)$$

Here N is the neighbor of the cell with r as the centre and a subset of the finite sequence D_n :

$$N = \{N_1, N_2, \dots, N_q\} \quad (3)$$

N_q stands for the position of q relative to r ; f is the conversion rule $S(r, t) \rightarrow S(r, t + 1)$:

$$f = \{f_1, f_2, \dots, f_m\} \quad (4)$$

The largest area occupied by a human being depends on his or her physical size in all directions, especially on the thickness of the shoulder and body. In order to simplify calculation, a normal practice is to reduce an abstract individual or a rectangular area into an oval. This is shown in Fig 1.

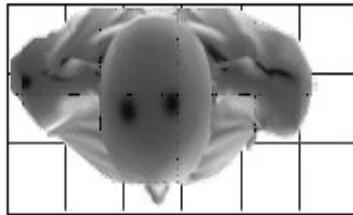


Fig. 1. The rectangular model of human physiological

At this point the individual area $S_p(m^2)$ can be expressed as:

$$S_p = b_p d_q \tag{5}$$

According to the measurement data of the body thickness and the shoulder width of common Chinese, the average projection area is $0.197m^2$ calculated by the rectangular body [7]. This article sets the mesh size to accommodate a person, that is $0.4m \times 0.4m$.

When the cellular automata model is applied to evacuation modeling, its dimension is 2, which is composed of two-dimensional cellular model of the grid. The basic framework of the model is the result of the even division of the building's horizontal area, where every grid is taken by its neighbors, other obstacles or a person, or it remains empty. Thus every cell appears in three states: 0 stands for emptiness, 1 stands for occupying by one person, 2 stands for wall or other obstacles. All the cells (except for those taken by other buildings) change continuously, and cell states at the time of $t + 1$ and t are associated, and so are the states of the neighboring cells. In the following Von Neumann model [8] in Fig 2, the dark area in the heart represents the center cell, while the light-colored areas around stands for other areas.

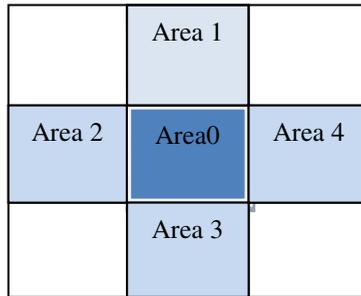


Fig. 2. Von Neumann Model

After the building being divided in multiple grids, every grid can only contain one person. The states of the grids appear in two situations: occupied by pedestrians or obstacles, or empty, so here the two-dimensional array $A(i \times j)$ refers to the information whether the space is occupied or not. That is, 1 means the space is taken, 0 means the space is not taken. Then, in two-dimensional plane, from upper left corner point, a two-dimensional plane can be create for the origin of coordinates x - y which positions itself within the field of the two-dimensional array, following the column by the same direction, the direction of the grid in the ranks of the order. Thus cell $A(i, j)$ center coordinates (x, y) can be established as: $x(i, j) = 0.15i - 0.125$; $y(i, j) = 0.15j - 0.125$. In this way the principle of the shortest path can be used to calculate the distance between the exit and the cellular and therefore helps to pin down a proper choice among available cellular selections.

Extended Cellular Automata Model

Physiological Factors Analysis

According to statistics, people can be grouped into young people by 85%, the elderly by 9%, children by 4%, the disabled by 2%, where different groups of people maintain different velocities. A velocity factor can be obtained from their group characters. To



simplify the processing, the paper divides the people into two groups: the young, the old and children [9]. The young have a proportion of 85%, and the rest are the old and children. A normal pedestrian's speed is 1.0m / s, and the size of a grid in the above figure is 0.4m × 0.4m, then a time step is 0.4s. Obviously in an emergency evacuation case, human walking speed will increase significantly. According to the classification of groups, the evacuation speed of young people is 1.2m / s, while for the older and children, the speed is 0.6m / s [10].

Table 1. Distribution of People's Physiological Factor

| Groups | Speed of evacuation | weight | Amount of grids |
|----------------------|---------------------|--------|-----------------|
| The young | 1.2m/s | 0.85 | 3 |
| The old and children | 0.6m/s | 0.15 | 1.5 |

Analysis of Psycho Preferences

In addition to distinguishing physiological factors, the psychological factors are also noteworthy when irrational behavior appears under non-normal conditions in the process of evacuation in emergency situations. According to Sime JD's relevant studies on emergency evacuation [11], some typical special psychological preferences can be decided.

Group psychology: the so-called herd mentality refer to the individual's pursuing the most others' practice by escaping from the large crowd, and giving up his/her own assertive and views. In the evacuation process, herd mentality is a common phenomenon: people tend to follow the majority or the others right in front of them because they are in panic and unease and try to gain security close to others, which turns out that their escaping behavior differs a lot from their everyday movement habit.

It is generally believed that herd mentality can be expressed as the direction appeal. Following is the specific formula

$$P_d(i, j) = \frac{N_{ij}}{\sum_{k=1}^m N_k} \quad (6)$$

In the formula, $pd(i, j)$ represents the cell direction appeal probability in the position of candidate A (i, j) N_{ij} represents the amount of people passing the candidate cell A (i, j) at the present moment, m represents the number of cells in the centre cellular area, $\sum_{k=1}^m N_k$ is the sum-up number of people passing all the candidate cells at the present moment. Clearly, the appeal direction formula shows that the more people pass the lattice, the greater the direction appeal becomes.

Family psychology: In the evacuation, there are some inevitable occasions where families and friends are together. In such emergency situations, the individual will consider the safety of their relatives and friends. One of the examples is that parents seldom abandon their children. So, we need to consider family psychology in the evacuation, in which case the individual chooses to assemble with their family members and friends before they escape, not escaping himself/herself from the exit without his/her family members and friends.

3 Model Simulations and the Result Analysis

Simulation Analysis of a Single Space. In order to prove the established evacuation model of the extended cellular automata, the model is operated in simulation in a rectangular stadium. The basic parameters of the stadium are $31.5\text{m} \times 18\text{m}$, with two doors at the top, one door beneath, and all the doors being 2.0m in length. 635 individuals are randomly generated within the stadium, where the proportion of the elderly and children is 20%.

Without considering the herd mentality, bypass behavior, family behavior and the point where the urgent events (such as the fire position) occur, the simulation results can be shown in Fig 3. It's obvious that the crowd choose to escape from the nearest exit(s), so people in all the three areas choose to escape from the most advantageous exits. Based on the extended cellular automaton model, we can use the directional appeal factor $P_d(i, j)$ to indicate a cellular direction. In the process of simulating, every step of the mobile individual is the result of his/her decision to choose a best route based on all the people's movement in his/her field of vision. Therefore, the larger the directional appeal factor is, the stronger the herd mentality grows.



Fig. 3. A Single Space

In the process of emergency evacuation, the individual is vulnerable to the impact of the crowd behaviors surrounding, people's choice of direction and route in front provides reference for the individuals behind them. Let's go back to the above-mentioned stadium ($31.5\text{m} \times 18\text{m}$), when the directional attractive factor is set to 0.6, then you can see conspicuous herd mentality in the crowd. In Fig 4, people are supposed to escape from the door beneath, yet the result goes to the contrast that almost no people choose the door beneath because the majority rush to the upper doors. At the same time, people who are not affected by herd mentality choose the upper doors and the door beneath to escape effectively. The conclusion can be made that herd mentality is not conducive to crowd evacuation—to escape under the direction of the working staff is the most effective method.

Model Analysis. When considering the composition and distribution of people, we mention previously that 85% of the general population is young people, while 15% of the population is the elderly and children. Now we will analyze the influence under different population composition situations. In the figure, the vertical axis represents the proportion of the elderly and children, the horizontal axis represents the evacuation time (in seconds), and the simulation is still in the previously mentioned stadium, 635

individuals. In the process, the evacuation speed for young people and the elderly (children) are set at 1.2m / s, 0.6m / s respectively. The average simulation results in 20 experiments are shown in Fig 5.

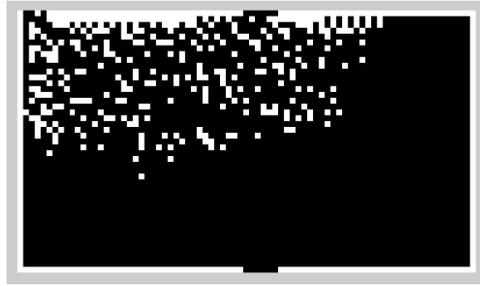


Fig. 4. Affected by herd mentality

Obviously, it is clear that as the proportion of the elderly and children grows bigger, the evacuation time becomes longer. The reason lies in the fact that young people walk faster and thus gain a more favorable effectiveness to escape in a shorter time.

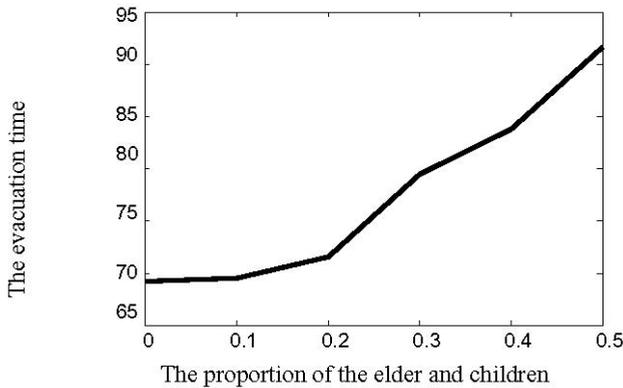


Fig. 5. Population composition situations and the evacuation

When considering the relationship between population density and evacuation time, it is clear that the bigger density makes the evacuation time longer. That is to say, the more people in the venue, the more difficult it will become to evacuate the personnel. Simulation results on the average of 20 experiments are shown in Fig 6, through which we find that when the crowd density is small, density $p = 0.05 \sim 0.15$, as the density increases, the evacuation time increases considerably. When the population density has reached a certain $p=0.4$ level, and the evacuation time increases smaller against the crowd density. This is because when the population density has reached a higher level, the evacuation inevitably falls into a crowded event. People’s walking speed will change in the crowded state, so the relative rate of change declines instead.

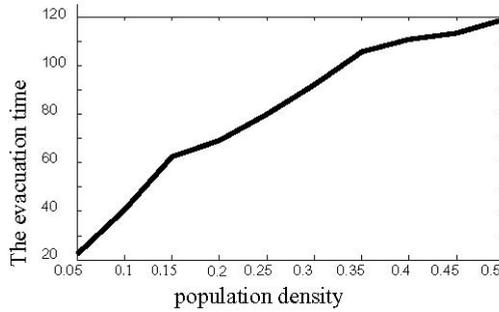


Fig. 6. Population density and evacuation time

According to the simulation results, we put forward proposals for the evacuation of large public venues as following:

1) Fire and personnel evacuation facilities should be taken into account in the very beginning of the construction of the venues. For example, simplified direction signs and words should be provided in the key channels.

2) In the event of an emergency, be sure to assign enough staff to divert the flow of people. Actually, in the evacuation process, the individuals may make non-rational behaviors due to their view restriction and affection from the others. Therefore, surveillance cameras can be built in various parts of the stadium so that crowd distribution can be effectively seized to reduce the time of personnel evacuation through scientific and effective method.

3) From psychological analysis of personnel actions we know that family behaviors generally affect evacuation, but herd mentality does not always affect the evacuation because, to some extent, individuals can be obtained from the preceding information about the evacuation. However, an excessive number of herd mentalities will significantly affect the evacuation.

4 Conclusions

The established model of cellular automata has a certain degree of generality through the establishment of extended cellular automaton model; real-life evacuation can be effectively simulated when we consider the personnel physiological factors and psychological factors. Through the application of the model in China's Pavilion of the World Expo, it is shown that the evacuation time is much less than the national standard for fire protection. The model still should be improved for its ignorance of the impact of the location where emergencies occur. For example, personnel evacuation in a fire needs to keep people away from the ignition point, and further researches should be made to study the psychological factors.

Foundation item: Support by the National Natural Science Foundation of China under Grant No. 60873114 and 60773009, Post Doctoral Funds number 20080440073 and Tianjin Natural Science Foundation (09JCYBJC27700).

References

1. Helbing, D., Farkas, I., Vicsek, T.: Simulating dynamical features of escape panic. *Nature* 407, 487–490 (2000)
2. Wolfman, D.: Cellular automata for traffic simulations. *Physical A* 263(1-4), 438–451 (1999)
3. Lu, W., Xiong, H., Li, X.: Evacuation Prediction in the Underground Shopping Street. *Fire Safety Science* 8(3), 72–79 (1999)
4. Liao, Y.-f., Ma, X.-q.: Simulation of Pedestrian Dynamics under Subway Fire Using Extended Cellular Automaton. *Journal of System Simulation* 20(24) (2008)
5. Jin, X.-g.: The simulation of Cellular Automata based on Matlab. *Computer Simulation* 19(4), 27–30 (2002)
6. Zhang, Z.-g., Li, Q.: Study on people's evacuation with cellular automata. *Computer Engineering and Design* 30(8), 1991–1993 (2009)
7. Sime, J.D.: Crowd Psychology and engineering. *Safety Science* 21(1), 1–14 (1995)
8. Helbing, D., Keltsch, J., Molnar, P.: Modelling the evolution of human trail systems. *Nature* 388, 47–50 (1997)
9. Helbing, D., Molnar, P.: Social force model for pedestrian dynamics. *Phys. Rev. E* 51, 4282–4286 (1995)
10. Muramatsu, M., Nagatani, T.: Jamming transition in two-dimensional pedestrian traffic. *Physica A* 275, 281–291 (2000)
11. Tajima, Y., Nagatani, T.: Scaling behavior of crowd flow outside a hall. *Physica A* 292, 545–554 (2001)

Software Protection with Encryption and Verification

Wentao Liu

Department of Computer and Information Engineering, Wuhan Polytechnic University,
Wuhan Hubei Province 430023, China
uddisoap@gmail.com

Abstract. The software protection is the key technology for the shareware and it can delay and prevent the software reverse and cracking. There are many methods such as junk code, anti-debug, virtual machine, deformation, packers, encryption and verification. The software registration is the most important step of protection and it can be implemented by many forms such as network registration, license key or file, dongle. This paper introduces and compares these protection methods. Some architectures of protection based on encryption and verification are provided and they are realized by RSA encryption and many verification methods which include self verification and key data verification. The method use the exe file and DLL file to protect each other and use the license file to check and provide software function. The method uses the RSA in the license key file to register and it can make it difficult to create the key generator for the cracker.

Keywords: software protection, encryption, verification.

1 Introduction

The software piracy and cracking harmed the development of software especially the shareware. There are many strategies to eliminate the influence such as some laws, education, self discipline and software protection technology. The software protection is the basic link and the author of software can master it directly. In general, there are two types of software: freeware and shareware. The freeware can be got and used freely. The shareware uses the style of trial version for the user before buying it. The trial version limits some functions which can be opened for the registration user. They also can limit the time or frequency of use. The cracker can remove this limit when the software protection is not strong. Because of the crack many software can't be sold normally and it lead to the loss of the author. Further, the result is that the author will not upgrade the software.

The software system protection can be implemented based on hardware or software. The hardware protection includes dongles and CD-ROM or floppy disc protection. The software protection includes license key, license file and network verification. It involves many protection technologies such as encryption, SEH, anti-debug, junk code, virtual machine, self verification and software packers.

2 Software Protection Methods

Dongle

The dongle is running on the interface of computer and it can be implemented by USB machine or other special machine. It includes the key algorithm or verification modules of software. The software must read some data from the dongle when it is running. When it can't get the correct data, the software can't run correctly. Many business software use the dongle to protect the software and it can get high strength protection but it need more fees.

License Key

This is more popular method of software protection. The software uses the user name and the license key to check the correct user. If the license key doesn't match the username, the software can't get correct result. The software can use more encryption algorithm to make the license key. In order to make the unique license key for each customer, the software can use the machine key as the username. The machine key can be created by the CPU key, hard disc key, network card key and motherboards key.

License File

The key data or the registration information are stored in one file and the software get the data from the file when it running. It can contain more data than license key but it need give the user a file when the users buy the software so it is not convenient for the registration.

Network Verification

If the author use the machine key as the user name, so the software must be used in one computer. This problem can be resolved by network verification. The license key can be stored in one network service and the software must read the data from the service when it running. The data stored in the service can be more complex such as software key data which is encrypted in the service. When the username and license key is matched, the key data can be transported to the client.

Junk Code

The cracker decompiles the binary application to the ASM code and he can read it easily. So we can add some junk code in the binary application and make the cracker can't understand the real meaning even though he can decompiles it. The junk code is no useful in the binary application but it can confuse the cracker to decompile the binary application. It can use some jump code to implement and also need some stack operations.

Anti-Debugging

The cracker uses the tools to get the ASM code for the binary application. The tools include OllyDBG, WinDBG, SoftICE, IDA Pro. The tools provide more functions to debug the application so the cracker can find the principle of program to crack. It can remove a protection code or change some code from a trial version. Anti-Debugging is to prevent crackers from using these debugging tools. In win32 platform, there are many technologies to achieve the debugging information such as IsDebuggerPresent, CheckRemoteDebuggerPresent, NtQueryInformationProcess, NtGlobalFlag and so on [1].

Software Packers

The cracker can use the disassembler tools to study how the application works and it can modify the compiled application to change the process. The cracker can bypass the routine of checks and delete it from the application and make the trial version become registration version. So we can make some packers for the binary software and it will make more mechanism to protect the software. There are many tools such as Themida, Execryptor, Armadillo and so on.

Anti-Memory Dump

When the cracker is studying the binary code, they can make the code in memory to a local file. The binary program must be running in the memory finally and the cracker can copy the memory data using some debugging tool easily. The anti memory dumping is also a protection method.

Virtual Machine

The binary application must be running in the CPU environment. So the cracker can analysis these ASM code to get real meaning. If the binary application is running in a virtual machine which is not the real code in the binary application, the cracker can't get the real meaning of program directly. The cracker must master the virtual machine code firstly. If the virtual machine is very complex, it will prevent the cracker reading the code.

Software Watermarking

The software watermarking is made to protect the copy of software. The software piracy can't delete the software watermarking and it can give the piracy the proof. It can identify the author of software with uniqueness.

3 Encryption and Verification

RSA

The traditional cryptography uses one key which shared by both sender and receiver. The RSA [2-3] uses two keys: a public key and a private key and is an asymmetric method and it is a public key cryptosystem. The public encryption method relies on a public encryption algorithm, a public decryption algorithm and a public encryption key. It uses the large integers and security is due to cost of factoring large numbers. So the RSA can become very secure because it is very hard to factor N to find P and Q if N is sufficiently big. The prime factorization is computationally very hard so it can't be cracked easily. The public key is used to encrypt messages or verify signatures and it may be known by anybody. The private key is used to decrypt messages or create signatures. The body that encrypts message can't decrypt messages. The plaintext is encrypted by public key and become cipher text and it is can be transmitted in the network. The receiver can decrypt the cipher by the private key. It can be used for key distribution or digital signature. There are three approaches to attack RSA: The brute force key search, mathematical attacks and timing attacks. Because of the character, it can be used for the registration of software. The author can use the private key encrypt the username and create the key to the user. The software can decrypt the key by the public key to get the correct registration. Even if the cracker get the public key, it can't

get the private key to create the registration key. The drawback of RSA is slow especially for the application with quick running and the message must be broken into chunks and each block is encrypted separately.

Elliptic Curve Cryptography

Elliptic curve cryptography (ECC) is an asymmetric algorithm and it works on the basis of elliptic curves. The principle is that an elliptic curve is a simple function that is drawn as a gently looping curve on the X and Y plane. The two points can be added to get a third point on the curve.

Data Encryption Standard

The Data Encryption Standard (DES) [4] was developed in the 1970 by the National Bureau of Standard and its aim is to give a standard method for protecting key data and it became a federal standard in 1976. It is symmetric cryptography and same key is used for encryption and decryption. It has been widely used around the world. DES uses the confusion and diffusion techniques and the diffusion is got through numerous permutations and the confusion is made through the XOR operation. The system establishes a key and encrypt message symmetrically using key. It is father than asymmetric methods. The DES is easy to implement in hardware or software and is hard to attack. The algorithms use a few fast subroutines and decryption uses same routines. It uses the block cipher with 64 bits at a time and initial permutation rearranges 64bits for the plaintext and encoding is in 16 rounds. The DES performs two operations: bit shifting and bit substitution. The key can make the process of works and in finally the cipher text cannot be got from the original without the key. The key may be stored as a 64-bit number and every eight bit is a parity bit which is pitched during the algorithm. The way of using block cipher is called mode of use and there are four modes: Electronic Codebook Book (ECB), Cipher Block Chaining (CBC), Cipher Feed Back (CFB) and Output Feed Back (OFB). The triple des run DES three times with ECB mode.

Advanced Encryption Standard

The Advanced Encryption Standard (AES) [5] was published by National Institute of Standards and Technology in 2001. AES is similar to DES with block cipher and contains 128-bit blocks. The key may be 128-bit, 192-bit or 256-bit. The S-boxes is based on modular arithmetic with polynomials which is non-linear and easy to analyze. The AES is intended to replace DES and it is efficient than DES. AES is also a symmetric block cipher and to encrypt and decrypt message using the same key. The plain text and the cipher text have the same size. The AES uses four types of transformations to make security function which include substitution, permutation, mixing and key adding. The AES uses two invertible transformations and the first transformation is made at the encryption site. The second transformation is shifting and it permutes the bytes. The AES gives a better combination of security and speed than DES. It can give higher security against brute force attack. The AES is more secure than DES because of the larger size key and the statistical attacks can't decrease the security of the algorithm. The differential and linear attacks can't be implemented on the AES. So it is more secure and can replace the DES. The AES can be realized in software and hardware. The AES only use the table lookup, XOR and shift operations. The AES contains the character of simple design, high speed algorithm and low

memory costs. Because of the simple algorithm in the AES, it can be easily implemented with cheap processors and don't cost more money.

Blowfish

The Blowfish [6] is produced by Bruce Schneier in 1993 and it contains many characters such as high speed, easy implementation, compact and security. It uses a variable key from 32bit from 448bit and it produce sub key. Blowfish is made as a block-mode cipher and it uses 64-bit block cipher with variable length key. The operation only includes binary addition and XOR operation and it has 16 rounds. It can get more quickly rapid and it is free. It uses dynamic s-boxes and it dependent on the key. The Blowfish is made to run quickly on 32-bit microprocessors and it is optimized for such situations where there are few key changes. So far there is no attack in the full 16-round version.

Rivest Cipher

The Rivest Cipher (RC2) was designed to be a DES replacement. According to RSA, RC2 is up to three times faster than DES, but the different key lengths is the vulnerabilities in the algorithm. RC4 is made as a stream cipher but all the symmetric ciphers were block ciphers. The stream cipher works by enciphering the plaintext in a stream with the format of bit by bit. RC4 can use the key which length is from 8 to 2048 bits. The algorithm can give the rapid which is 10 times faster than DES. The vulnerable character of the encryption is the possibility of weak keys. RC5 [7] is made as a block cipher which contains multiple variable elements, many rounds, key sizes, and block sizes. RC5 can give adequate security for brute forcing technology. The RC6 [8] has a 128-bit block size and it is separated into four words of 32 bits each. The RC6 can run on 32-bit computers. The available key lengths can lead brute force attacks to slow.

Message Digest

The Message Digest algorithm is different form encryption algorithm. It can be sued to verify the correctness of message. It use hash functions to get the result data which contains the unique character. The hash is a special mathematical function that performs one way encryption procedure. Once the algorithm is made there is no method to transform the cipher text to the plaintext. And to the different plaintext the result must be different. That is to say to get two different plaintexts that have the same hash value is impossible.

Message Digest (MD) [9] is made to produce a message digest or hash from the original data using the algorithm. Message Digest 2 takes a data input with any length and creates a hash output with 128 bits. MD2 is optimized for 8-bit machines. The MD4 and MD5 are optimized for 32-bit machines so the MD5 is more popular. The MD4 is not secure because of basic flaws in the algorithm. Message Digest 5 is similar to the MD4 algorithm and it is slightly slower and more secure. MD5 produces a 128-bit hash of a message of any length.

Secure Hash Algorithm

Secure Hash Algorithm (SHA) [10] was produced in 1993 by the National Institute of Standards and Technology for secure hashing in the U.S. Digital Signature Standard (DSS). It uses block mode and get an input of up to 264 bits and compress it to 160 bits. When the hash result has been generated for all blocks the message is represented by a 160-bit string. SHA is one of the more secure hash functions because its output is 160-bit long than 128-bit in MD5.

4 Protection Architecture

In this section, some concrete method or architecture are provided using above technology. The protection can't be implemented by one technology and it must combine more technology into an integrated protection system.

The architecture of using license key or file is shown in the figure 1. The username can be telephone, email, company name and the machine code. The encryption method can be RSA, DES and AES. In order to make complex, the process can make a temporary key which is made from username by some hash algorithm such as MD5 and SHA. The author creates the license key and the consumer use the software and the software can decrypt the license key. If the temporary key from the license key is equal to the temporary key from the username using the MD5, the registration is successful. The process can be implemented in many places in the software and for many times. The username can be created by the machine code such as CPU code, hardware key and network MAC code and so on. So the license key must be used in only one machine. If the software must be used in more computers, the registration progress can be provided in network verification. When the consumer uses the software and it will connect the server which store many usernames information, the server can compare the username with the information stored in the server.

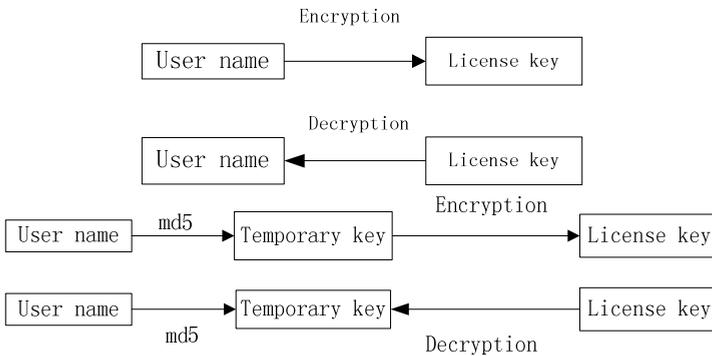


Fig. 1. The architecture of license key

The key data in the application must be encrypted and it can be read by the cracker easily using some decompile tools. The architecture of key data protection is shown in Figure 2. The key data is encrypted by RSA firstly and it be encrypted by DES again which use the key of hash of EXE or DLL file. Finally the key data can be stored in the key file or database. The key file and database also can be encrypted again. When the applications use the key data, it firstly makes computation of hash of the EXE file or DLL file. The hash result is the key of DES and it decrypts the cipher data. If the EXE file or DLL file is modified by the cracker, the key is not correct and the key data is can't be decrypted correctly. So the next operation will get the error results. The application will not make correct function. The RSA is an asymmetric algorithm and



the cracker can't get the private key so it can't make the key file. If the cracker wants to crack the software, he must read the key data and modify the all points the data used. If there are many key data in the software, the workload of cracking the software will be very big for the cracker.

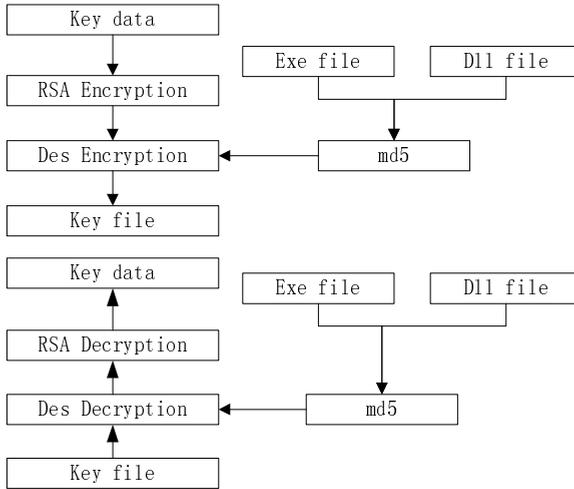


Fig. 2. The architecture of key data protection

The verification of architecture is shown in figure 3. The software can verify any things in any time. The object can be application file name, file creating time, file data, file size and the DLL file. The verification algorithm can use the MD5 or SHA and the application can verify these data in any stage. The verification times must be larger and it can increase the difficulty of crack. Finally, the information which include many license key and may verification information can be form into a key file. The key file is binary and it can be used for the software to register and verify the software. So if the cracker must crack the protection, he must know the stricter of key file and to create the every information modules. The author can make the key file more complex. The every module can be subdivided into many small pieces information and the random uselessness information can be inserted between the pieces information.

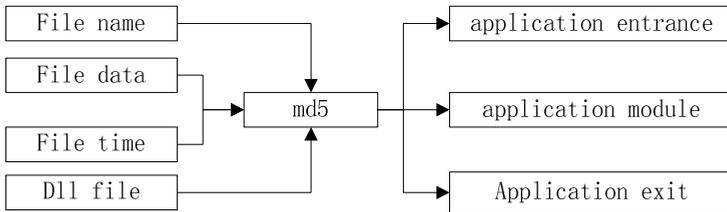


Fig. 3. The architecture of verification



5 Summary

The software protection is the key technology for the business software. Many shareware is cracked by the cracker and can't achieve the income. The crack behaviors prevent the software from developing normally. This paper introduces some software protection methods about some aspects such as dongle, license key, license file, network verification. And some protection technology is analyzed and it includes anti-debugging, anti memory dumping, software packers, junk code, virtual machine, and software watermarking. The encryption method and verification method used in the software protection are introduced. The encryption often used in the software protection includes RSA, DES, AES, RC, Blowfish. The verification method used in software protection is the MD5 and SHA. Finally, some architectures of protection are provided and they can increase the difficulty of crack. The license key is the basic module of shareware and it often be studied by the cracker firstly. The key data in the software can be encrypted by RSA and DES with the key which is the result of MD5 of EXE file or DLL file. So the cracker can't modify the EXE file or DLL file and it will lead the application can't produce correct result.

References

1. Jackson, J.: *Introduciton Into Windows Anti-Debugging* (2008), http://www.codeproject.com/KB/security/Intro_To_Win_Anti_Debug.aspx
2. Stallings, W.: *Cryptography and Network Security: Principles and Practice*, 3rd edn. Prentice Hall (2002)
3. RSA Cryptography Standard, <http://www.rsa.com/rsalabs/node.asp?id=2125>
4. Data Encryption Standard (DES) (1999), <http://csrc.nist.gov/publications/fips/fips46-3/fips46-3.pdf>
5. Daemen, J., Rijmen, V.: *The Design of Rijndael: AES - The Advanced Encryption Standard*. Springer, Heidelberg (2002)
6. Schneier, B.: *Description of a New Variable-Length Key, 64-Bit Block Cipher (Blowfish)*. In: Anderson, R. (ed.) *FSE 1993*. LNCS, vol. 809, pp. 191–204. Springer, Heidelberg (1994)
7. Rivest, R.L.: *The RC5 Encryption Algorithm*. In: Preneel, B. (ed.) *FSE 1994*. LNCS, vol. 1008, pp. 86–96. Springer, Heidelberg (1995)
8. Rivest, R.L., Robshaw, M.J.B., Sidney, R., Yin, Y.L.: *The RC6 Block Cipher, v1.1* (August 1998)
9. Rivest, R.: *The MD5 Message-Digest Algorithm* (1992), <http://tools.ietf.org/html/rfc1321>
10. *Secure Hash Standard* (2002), <http://csrc.nist.gov/publications/fips/fips180-2/fips180-2withchangenotice.pdf>

Improved Tri-training with Unlabeled Data

Tao Guo and Guiyang Li

College of Computer Science, Sichuan Normal University, Chengdu, China
{tguo35,guiyang.li}@gmail.com

Abstract. Semi-supervised learning, such as co-training and tri-training learning from a combination of both labeled and unlabeled data have become a topic of significant recent interesting. In tri-training algorithm, it generates three classifiers from the original labeled examples and then refined using unlabeled examples in the tri-training process. This algorithm neither needs sufficient and redundant view, nor does it require the use of different supervised learning algorithms whose hypothesis partitions the instance space into a set of equivalence classes. However, the performance of this algorithm goes degradation under certain condition. In this paper, we address three issues existed in tri-training, including unsuitable error estimation, excessively confined restriction and deficiency of weight for labeled example and unlabeled example. An improved tri-training algorithm (im-tri-training) is proposed to work on the three issues. Experiments on UCI data sets show that im-tri-training outperforms the tri-training on learning performance.

Keywords: Machine Learning, Data Mining, Learning from Unlabeled Data, Semi-supervised Learning, Co-training, Tri-training.

1 Introduction

In machine learning or data mining applications, such as web page classification, pattern recognition, a large amount of labeled data should be available for building a model with good performance. Many supervised learning algorithms (e.g. J4.8, Bays and SVM) have been developed and extensively studied use labeled data. Unfortunately, in many practices, it is often the case that there is a limited number of labeled data along with a large pool of unlabeled data [1]. Can unlabeled data be used to help to improve learning accuracy? It is noteworthy that a number of methods called semi-supervised learning have been developed for using unlabeled data to improve the accuracy of prediction [2]. Early methods in semi-supervised learning were using mixture models and extensions of the EM algorithm [3]. More recent approaches belong to one of the following categories: self-training, transductive SVMs, co-training, split learning, and graph-based methods [4]. Since the self-training process may erroneously label some unlabeled examples, sometimes the learned hypothesis does not perform well [2]. The transductive SVMs constructs a graph on the examples such that the minimum cut on the graph produces an optimal labeling of the unlabeled examples according to certain optimization function [5]. A prominent approach in semi-supervised learning is co-training proposed by Blum and Mitchel [2].

Co-training requires two sufficient and redundant views, such a requirement can hardly be met in most scenarios [5]. Goldman and Zhou proposed an improved co-training algorithm [6]. It employs time-consuming cross validation technique to determine how to label the unlabeled examples and how to produce the final hypothesis [6]. In 2005, Zhou and Li proposed a new co-training style algorithm named tri-training. It is easy to be applied to common data mining application. After repeat the experiment provided by Zhou, we find that the applicability of the tri-training is indeed broader than that of co-training in some instances. But the algorithm it-self is backwards in some circumstances and exist three issues: (1) estimation for classification error is unsuitable. (2) excessively confined restriction introduce more classification noise. (3) differentiation between initial labeled example and labeled unlabeled example is deficient. In this paper, an improved tri-training algorithm is proposed. Experiments on UCI data sets indicate that the improved tri-training can avoid above limitations and can effectively exploit unlabeled data to improve the accuracy of classifier.

The rest of the paper is structured as follows: Section 2 describes the issues existing in tri-training in details. Section 3 introduces the proposed algorithms. Section 4 shows experimental and comparative results in different UCI data sets. Section 5 concludes for current work.

2 Issues in Tri-Training

Zhou and Li proposed a tri-training algorithm for solving the problem of co-training in 2005 [5]. Tri-training is described in detail as follows: Let L denotes the labeled example set, h_1 , h_2 and h_3 denote initial classifiers, U denotes the unlabeled example set, x is an example in U . L_t and L_{t-1} denote the set of examples that are labeled for h_t in the t -th round and the $(t-1)$ -th round, respectively. Three classifiers are initially trained from labeled examples. Any two of three classifiers are used to label the unlabeled examples x independently, if two of them agree on the label under certain conditions; the labeled unlabeled example (named pre-label hereafter) will be used to teach the third classifier. It repeats this work until none of h_i ($i=1,2,3$) changes. The final prediction is produced with a variant of a majority vote among all the classifiers. The integrated algorithm for tri-training is shown in [5].

We run experiments with an array of algorithms that are all available in WEKA consisting of J4.8, BP neural network and Naïve Bayes. The experiments on UCI data sets show that the tri-training algorithm goes degradation in performance under certain condition. In detail, the issues existed in tri-training are described in following.

Estimation for Classification Error Is Unsuitable. The theoretical foundation of tri-training is entirely based on accuracy of estimation value for error rate e_i . In the above algorithm, e_i is the classification error rate of the hypothesis estimated by the function $MeasureError(h_j \& h_k)$. In the algorithm, the author described that “since it is difficult to estimate the classification error on the unlabeled examples, only the original labeled examples are used”. Also an assumption that the unlabeled examples hold the same distribution as that held by the labeled ones was made by author. It exists two problems: 1) Since labeled examples participate in training, it leads to a larger difference between estimation value of error rate made by labeled examples and that made by actual value. Usually, the error rate of estimation is much smaller than actual

error rate. 2) Overfitting exists in this algorithm. When error rate of estimation is reduced to a smaller value, the error rate will be increased significantly.

Table 1 shows a comparative result of error rate for estimation and actual value. For each data set, 25% are kept as test example while 20% rest of the data sets are used as labeled example. J4.8 is used in this comparison. The error in Table 1 can be computed

$$\text{as } \frac{\text{ActualError} - \text{EstimationError}}{\text{ActualError}} .$$

Table 1. Comparative result of error rate for estimation and actual value

| Data Set | Estimation Error | Actual Error | Error |
|---------------|------------------|--------------|--------|
| anneal | 0.0070 | 0.1005 | 93.03% |
| breast-cancer | 0.0634 | 0.3398 | 81.34% |
| colic | 0.0393 | 0.1927 | 79.61% |
| diabetes | 0.0451 | 0.2604 | 82.68% |
| glass | 0.0637 | 0.5104 | 87.52% |
| hypothyroid | 0.0023 | 0.0120 | 80.83% |
| ionosphere | 0.0101 | 0.1274 | 92.07% |
| iris | 0.0010 | 0.0690 | 98.55% |
| tic-tac-toe | 0.0419 | 0.2482 | 83.12% |
| vote | 0.0171 | 0.0432 | 60.42% |
| zoo | 0.0473 | 0.2226 | 78.75% |

Table 2. Comparative result for a learning process on colic data set

| Iteration | Estimation Error | Actual Error |
|-----------|------------------|--------------|
| 1 | 0.0188 | 0.1708 |
| 2 | 0.0185 | 0.1691 |
| 3 | 0.0181 | 0.1730 |
| 4 | 0.0181 | 0.1764 |

Table 3. Comparative result for a learning process on Ionosphere data set

| Iteration | Estimation Error | Actual Error |
|-----------|------------------|--------------|
| 1 | 0.0217 | 0.0786 |
| 2 | 0.0200 | 0.0862 |
| 3 | 0.0000 | 0.1083 |

From the comparative result, we may find that the estimation error rate on training examples has big differences with that on actual error rate. The biggest difference goes to 98.55% on *iris* data set.

Tables 2 to 3 are comparative results for a learning process on two different data sets. The iteration for each learning process keeps on 4 and 3 respectively. Tables 2 to 3 show that the estimation error goes down, even to zero for *ionosphere* on the third iteration. On the contrary, the actual error rate goes up. These experiment results demonstrate that overfitting do exist in the tri-training algorithm.

Excessively Confined Restriction. In tri-training algorithm, Eq. 1 from Anglimum and Laird [7] was introduced:

$$m = \frac{c}{\varepsilon^2 (1 - 2\eta)^2} \tag{1}$$

Where m is the sample size, ε is the hypothesis worst-case classification error rate, $\eta (< 0.5)$ is an upper bound on the classification noise rate and c is a constant. Eq.2 can be obtained from Eq.1

$$\frac{c}{\varepsilon^2} = m(1 - 2\eta)^2 \tag{2}$$

Assume t is the number of iteration. It can be derived that if $\frac{c}{\varepsilon_t^2} > \frac{c}{\varepsilon_{t-1}^2}$

then $\varepsilon_t < \varepsilon_{t-1}$. This relationship can be used to decide if the amount of additional data labeled is sufficient to compensate for the increase in the classification noise rate. The condition can be expressed as Eq.3:

$$m_t (1 - 2\eta_t)^2 > m_{t-1} (1 - 2\eta_{t-1})^2 \tag{3}$$

In Eq.3, $m_t = |L \cup L_t|$, $\eta_t = \frac{\check{\varepsilon}_t |L_t|}{|L \cup L_t|}$, $m_{t-1} = |L \cup L_{t-1}|$

As described in Anglimum and Laird [7], once the Eq.3 is satisfied, the newly added unlabeled examples could compensate the classification noise rate caused by themselves. Zhou and Li [5] separates the Eq.3 into $m_t > m_{t-1}$ and $\eta_t < \eta_{t-1}$, then derives $|L_t| > |L_{t-1}|$ and $\check{\varepsilon}_t |L_t| < \check{\varepsilon}_{t-1} |L_{t-1}|$. After analyzing the deduction, we found that the separation and deduction of Eq.3 is not correct completely because $|L_t| > |L_{t-1}|$ and $\check{\varepsilon}_t |L_t| < \check{\varepsilon}_{t-1} |L_{t-1}|$ is a sufficient condition for Eq.3, not necessary condition. When $|L_t| < |L_{t-1}|$ or $\check{\varepsilon}_t |L_t| > \check{\varepsilon}_{t-1} |L_{t-1}|$, Eq.3 is still established once the parameters are set suitable.

In tri-training algorithm, when $\check{\varepsilon}_t < \check{\varepsilon}_{t-1}$, $|L_t| > |L_{t-1}|$, and $\check{\varepsilon}_t |L_t| \geq \check{\varepsilon}_{t-1} |L_{t-1}|$ happens, L_t could be randomly subsampled, let integer s denotes the size of L_t after subsampled, then s can be computed as Eq.4.

$$s = \left\lceil \frac{\check{\varepsilon}_{t-1} |L_{t-1}|}{\check{\varepsilon}_t} - 1 \right\rceil \tag{4}$$

This kind of processing conceals a problem, that is, when randomly samples are fewer, the data distribution can be changed completely and will cause classification noise rate changed and make algorithm instability.

Table 4. Comparative result of error for initial and tri-training

| Data Set | Initial | Tri-training | Tri-training With Oracle |
|---------------|---------|--------------|--------------------------|
| anneal | 0.0409 | 0.0469 | 0.0469 |
| breast-cancer | 0.3491 | 0.3456 | 0.3505 |
| colic | 0.1918 | 0.1894 | 0.1968 |
| diabetes | 0.2865 | 0.2868 | 0.2941 |
| glass | 0.4491 | 0.4590 | 0.4622 |
| hypothyroid | 0.0130 | 0.0129 | 0.0130 |
| ionosphere | 0.1448 | 0.1541 | 0.1544 |
| iris | 0.1221 | 0.1121 | 0.1056 |
| tic-tac-toe | 0.2607 | 0.2703 | 0.2769 |
| vote | 0.0490 | 0.0518 | 0.0527 |
| zoo | 0.2640 | 0.2799 | 0.3040 |

Table 4 shows a comparative result of initial J4.8 and tri-training algorithm on some UCI data sets. In order to eliminate the influence of estimation error to the performance of the algorithm, Table 4 also gives a result of tri-training with Oracle. Where Oracle is a actual error rate which is generated on unlabeled data by $MeasureError(h_j \& h_k)$. Table 4 indicates that the error rate on tri-training is higher than that on initial J4.8 for most of data sets. When the Oracle is introduced to solve the first problem, the performance of the tri-training has no improvement except the *iris*. It is caused by excessively confined restriction.

Differentiation between Initial Labeled Example and Pre-label is Deficient. The performance of semi-supervised learning algorithms is usually not stable because it may erroneously label some unlabeled examples during the learning process [2]. Kurt Diressens [4] demonstrated that choosing appropriate weights for the pre-label example L_i , the classifier consistently improves on the original classifier. In [5], an unlabeled example can be labeled as long as the other two classifiers agree on the labeling of the example. In such a scheme, it treats all the pre-label examples which is in L_i exactly like the training examples and the weight of each example in L_i equals with initial labeled example. It is obvious that if the prediction of h_2 and h_3 on x is correct, then h_1 will receive a valid new example for further training; otherwise h_1 will get an example with noisy label.

3 Improved Tri-training

Inspired by Kure Diessens and Sally Goldman [4, 6], in this section, we proposed an improved tri-training. The algorithm will focus on the three issues existed in tri-training. The pseudo-code refined from tri-training is presented in Table 5.

Table 5. Pseudo -code describing the improved tri-training algorithm

```

im-tri-training( $L, U, Learn, Weight$ )
Input:  $L$ : Original labeled example set
       $U$ : Unlabeled example set
       $Learn$ : Learning algorithm
       $Weight$ : weight for unlabeled example
for  $i \in \{1..3\}$  do
   $S_i \leftarrow BootstrapSample(L)$ 
   $h_i \leftarrow Learn(S_i)$ 
   $z_i \leftarrow |S_i|$ 
end of for
repeat until none of  $h_i$  ( $i \in \{1..3\}$ ) changes
  for  $i \in \{1..3\}$  do
     $L_i \leftarrow \emptyset$ ;  $update_i \leftarrow FALSE$ 
     $e_i \leftarrow MeasureErrorCV(h_j \& h_k)$  ( $j, k \neq i$ )
    for every  $x \in U$  do
      if  $h_j(x) = h_k(x)$  ( $j, k \neq i$ )
      then  $L_i \leftarrow L_i \cup \{x, h_j(x)\}$ 
    end of for
     $z_i = |S_i \cup L_i| \left( 1 - 2 \frac{e_i |L_i|}{|S_i \cup L_i|} \right)^2$ 
    if ( $z_i > z_i'$ ) % otherwise Eq. 3 is violated
    then begin
      for every  $x \in L_i$  do
         $x.weight = weight \times |S_i| / |L_i|$ 
      end of for
       $h_i \leftarrow Learn(S_i \cup L_i)$ 
       $z_i' = z_i$ 
    end if
  end of for
end of repeat
Output:  $h(x) \leftarrow \arg \max_{y \in \text{label}_i: h_i(x)=y} 1$ 

```

In above algorithm, we use 10-fold cross validation on the original labeled example to determine the error rate. In order to make the newly added unlabeled examples to compensate the classification noise rate, Eq.3 is used directly to determine when an unlabeled example could be labeled for a classifier. For any classifier, if h_2 and h_3 agree on the labeling of an example x in U , then x can be labeled for h_1 . Since there is no-reliable unlabeled example, the weight is introduced to reduce the interference of noise data and improve the algorithm's learning ability. In this algorithm, S_i denotes labeled examples and L_i denotes unlabeled examples. We assign weights of 1.0 to every example in S_i and of $weight \times |S_i| / |L_i|$ to all unlabeled example in L_i .

4 Experiments and Analysis

In order to compare the experimental results equally with tri-training, eleven UCI Machine learning data sets are used in the experiments. For each data set, 25% data are

Table 6. Classification error rate of initial algorithm, tri-training and im-tri-training

| DataSet | J4.8 decision tree | | | |
|---------------|--------------------|--------------|--------------------------------|----------------------------------|
| | initial | tri-training | im-tri-training (weight =1) | im-tri-training (weight =0.1) |
| anneal | 0.0534 | 0.0553 | 0.0533 | <u>0.0502</u> |
| breast-cancer | <u>0.2788</u> | 0.2859 | 0.2830 | 0.2812 |
| colic | 0.1597 | 0.1565 | 0.1564 | <u>0.1554</u> |
| diabetes | 0.2885 | 0.2890 | 0.2864 | <u>0.2786</u> |
| glass | <u>0.4504</u> | 0.4987 | 0.4925 | 0.4722 |
| hypothyroid | 0.0137 | 0.0135 | <u>0.0129</u> | 0.0130 |
| ionosphere | 0.1689 | 0.1691 | 0.1750 | <u>0.1678</u> |
| iris | 0.0918 | 0.0918 | 0.0918 | <u>0.0837</u> |
| tic-tac-toe | 0.2866 | 0.2878 | 0.3037 | <u>0.2778</u> |
| vote | <u>0.0490</u> | 0.0518 | 0.0527 | <u>0.0490</u> |
| zoo | 0.2439 | 0.2639 | <u>0.2399</u> | 0.2520 |

| Data Set | BP Neural Network | | | |
|---------------|-------------------|---------------|--------------------------------|----------------------------------|
| | initial | tri-training | im-tri-training (weight =1) | im-tri-training (weight =0.1) |
| anneal | 0.0383 | 0.0365 | 0.0388 | <u>0.0360</u> |
| breast-cancer | 0.3239 | 0.3323 | <u>0.3197</u> | 0.3281 |
| colic | 0.2260 | 0.2347 | 0.2315 | <u>0.2163</u> |
| diabetes | 0.2661 | 0.2630 | 0.2651 | <u>0.2583</u> |
| glass | 0.6200 | 0.6223 | 0.6178 | <u>0.6176</u> |
| hypothyroid | <u>0.0747</u> | 0.0795 | 0.0793 | 0.0803 |
| ionosphere | 0.1850 | 0.1804 | 0.1781 | <u>0.1776</u> |
| iris | 0.3432 | <u>0.3351</u> | <u>0.3351</u> | 0.3621 |
| tic-tac-toe | 0.0543 | <u>0.0317</u> | 0.0292 | <u>0.0267</u> |
| vote | 0.0537 | 0.0525 | 0.0527 | <u>0.0518</u> |
| zoo | 0.3120 | 0.3399 | 0.3800 | 0.3120 |

| DataSet | Naïve bayes | | | |
|---------------|---------------|--------------|--------------------------------|----------------------------------|
| | initial | tri-training | im-tri-training (weight =1) | im-tri-training (weight =0.1) |
| anneal | 0.0840 | 0.0867 | 0.0859 | <u>0.0816</u> |
| breast-cancer | 0.2859 | 0.2830 | <u>0.2690</u> | 0.2816 |
| colic | 0.2380 | 0.2326 | 0.2389 | <u>0.2295</u> |
| diabetes | 0.2671 | 0.2645 | 0.2661 | <u>0.2630</u> |
| glass | <u>0.4932</u> | 0.5179 | <u>0.4932</u> | <u>0.4932</u> |
| hypothyroid | 0.0416 | 0.0439 | <u>0.0400</u> | 0.0405 |
| ionosphere | <u>0.1873</u> | 0.2103 | 0.2321 | 0.2103 |
| iris | 0.1162 | 0.0756 | <u>0.0729</u> | 0.0750 |
| tic-tac-toe | 0.3075 | 0.3154 | 0.3062 | <u>0.3037</u> |
| vote | <u>0.0879</u> | 0.0907 | 0.1018 | 0.0907 |
| zoo | <u>0.2680</u> | 0.2640 | <u>0.2520</u> | 0.2600 |

kept as test example while the rest are used as the set of training example. Three subsets will be generated using different randomly selection of training example. In each subset, L and U are partitioned with 80% unlabeled rate. J4.8 decision trees [8], BP neural networks, and Naïve Bayes algorithms are used as initial algorithm in this experiment. We collect results for different values of the weighting parameter *weight* ranging from 0.1 to 1.

Table 6 shows that im-tri-training can effectively improve the prediction with the classifiers under the all unlabeled rates. The biggest improvements achieved have been underlined in Table 6. In detail, with *weight*=1 and *weight* =0.1, when J4.8 decision trees are used as initial algorithm, im-tri-training has 2 and 7 wining data sets respectively. When BP neural networks are used as initial algorithm, im-tri-training has 2 and 7 wining data sets respectively. When Naïve Bayes are used as initial algorithm, im-tri-training both have 5 wining data sets.

Analyzing the experiment of im-tri-training, the results show that im-tri-training algorithm achieves a better performance on most of data sets. But on some data sets, the initial algorithm gets better classification accuracy with three different classifiers while tri-training and im-tri-training goes degradation. For example, *breast-cancer* and *glass* data sets get better performance with J4.8 than that with both tri-training and im-tri-training. The reason for this degradation comes from randomly subsample and deviation of estimation error with 10-cross validation. Therefore, the introduced noise will cause false positive and makes algorithm backwards.

5 Summary

Since the requirements for co-training are hard to meet, a co-training style semi-supervising algorithm called tri-training is proposed by Zhou and Li. Using multiple classifiers, tri-training composes a final detection with good efficiency and generalization ability without sufficient and redundant views. After analyzing the tri-training algorithm theoretically and performing experiments, this paper points out the three implicated issues, including error estimation, excessively confined restriction and deficiency of weight for labeled example and pre-label example. For refining the deficiencies existed in tri-training, an im-tri-training algorithm is proposed to select additional unlabeled example more reliably. In this algorithm, the excessively confined restriction is broadened properly to reduce classification noise. 10-cross validation technique is used to determine the error rate. And weight is also introduced to transform the unlabeled data into a weighted data that together with the original data. Experiments on UCI data sets show that im-tri-training is facilitated with good efficiency. However, extending im-tri-training is worth studying in determination of error rate in future work. Moreover, overcoming degradation when only fewer data sets is also an interesting study in the future.

References

1. Zhou, Z.-H.: Learning with unlabeled data and its application to image retrieval. In: Proceedings of the 9th Pacific Rim International Conference on Artificial Intelligence, pp. 5–10 (2006)
2. Blum, A., Mitchell, T.M.: Combining labeled and unlabeled data with co-training. In: Proc. 11th Annual Conference on Computational Learning Theory, pp. 92–100 (1998)
3. Li, J.: A Novel Semi-supervised SVM Based on Tri-training for Intrusion Detection. *Journal of Computers* 5(4) (April 2010)
4. Driessens, K., Reutemann, P.: Using Weighted Nearest Neighbor to Benefit from Unlabeled Data. *Encyclopedia of Machine Learning 2010*, 857–862 (2010)
5. Zhou, Z.H., Li, M.: Tri-training: Exploiting unlabeled data using three classifiers. *IEEE Transactions on Knowledge and Data Engineering* 17, 1529–1541 (2005)
6. Goldman, S., Zhou, Y.: Enhancing supervised learning with unlabeled data. In: Proceedings of the 17th ICML, pp. 327–334. Morgan Kaufmann, San Francisco (2000)
7. Angluin, D., Laird, P.: Learning from noisy examples. *Machine Learning* 2(4), 343–370 (1988)
8. Kotsiantis, S.B.: Supervised Machine Learning: A Review of Classification Techniques. *Informatica* 31(2007), 249–268 (2007)

Improved Particle Swarm Optimization Based on Genetic Algorithm

Chunhong Dou and Jinshan Lin

School of Information and Control Engineering, Weifang University,
Weifang 261061, P.R. China
chdoucon@hotmail.com

Abstract. The particle swarm optimization (PSO) algorithm is an intelligent optimization algorithm and has been successfully applied to many fields. However, it has the deficiencies of the poor search capability and the low calculation efficiency. To solve the problems, the genetic algorithm is introduced to improve the PSO algorithm in the paper. Then, three simulated examples were used to assess the ability of the improved PSO algorithm. The results showed that the improved PSO algorithm could jump out of the local minimum and achieve the better solution. Furthermore, the paper reveals that the improved PSO algorithm seems to be an improvement on the original PSO algorithm.

Keywords: Particle swarm optimization (PSO), improved particle swarm optimization (IPSO), genetic algorithm (GA).

1 Introduction

Particle Swarm Optimization (PSO) is an intelligent optimization algorithm and firstly inspired by social behavior of birds [1]. In recent years, the PSO algorithm is widely used for function optimization, network training, pattern classification, parameter optimization and signal processing [2]. Nonetheless, the PSO algorithm has still some shortcomings, such as the low convergence speed, inaccuracy solution and easily sinking into local optimal values. To relieve the problems existing in the PSO algorithm, the paper introduced the genetic algorithm into the PSO algorithm. The numerical experimental showed that the improved PSO algorithm in the paper could be an improvement on the original PSO algorithm in increasing convergent speed and searching accuracy.

2 Particle Swarm Optimization

To begin with, initialize a swarm of particles in solution space and each particle represents a potential optimal solution to the extreme optimization problem. Then, the characteristics of particles are described with position, velocity and fitness value; the fitness value can be calculated by the fitness function and decide whether or not the particles are superior [3].

Suppose that there is a population $X=(X_1, X_2, \dots, X_n)$ consisting of n particles in a search space of D -dimension, and the i th particle, expressed by the D -dimension vector $X_i=(x_{i1}, x_{i2}, \dots, x_{iD})^T$, means the location of i th particle as well as a potential solution to the problem. Thus, the velocity of i th particle is $V_i=(V_{i1}, V_{i2}, \dots, V_{iD})^T$, the individual extreme is $P_i=(P_{i1}, P_{i2}, \dots, P_{iD})^T$ and the global extreme of the population is $P_g=(P_{g1}, P_{g2}, \dots, P_{gD})^T$.

The particles update their velocity and location in terms of the individual extreme and the global extreme for every iteration and the formula is as follows:

$$V_{id}^{k+1} = \omega V_{id}^k + c_1 r_1 (P_{id}^k - X_{id}^k), \quad (1)$$

$$V_{id}^{k+1} = \omega V_{id}^k + c_2 r_2 (P_{id}^k - X_{id}^k), \quad (2)$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1}, \quad (3)$$

where ω is the inertia weight, $d=1,2,\dots,D$, $i=1,2,\dots,n$, k is the current iteration number, V_{id} is the velocity of particles, c_1 and c_2 are the nonnegative constant named acceleration factor, and r_1 and r_2 are the random number between 0 and 1. To avoid the blind searching of particles, the locations and the velocities of particles are suggested to be limited in $[-X_{\min}, X_{\max}]$ and $[-V_{\min}, V_{\max}]$, respectively.

3 Improved Particle Swarm Optimization Based on Genetic Algorithm

Some problems appearing in the PSO algorithm contain the poor search capability and the low calculation efficiency at the latter iteration stages. To alleviate the problems, the genetic algorithm is introduced to improve the PSO algorithm. More specifically, the mutation operation of the genetic algorithm is used to assist the PSO algorithm and then some initial variables are going to be reinitiated at a certain probability. The mutation operation extends the search space which is constantly reduced with the iterations, makes the particles get out of the optima location and search in the larger space, remains the diversity of the population and enhances the possibility of finding the better values. As a result, the initial particles are reinitialized at a certain probability after the particles have been renewed every time.

4 Numerical Evaluation

In this section, three examples were used to measure the ability of the improved PSO algorithm.

Example 1. A two-dimension Sphere function is written as follows:

$$f(x, y) = x^2 + y^2, (-10 \leq x, y \leq 10). \quad (4)$$

The Sphere function is unimodal and the optimal solution is located in the point $[0, 0]$. The fitness functions of the PSO algorithm and the improved PSO algorithm for the Sphere function are demonstrated in Fig.1. As shown in Fig. 1, the improved PSO algorithm looks superior to the PSO algorithm.

Example 2. A Rastrigrin function is expressed as following

$$f(x, y) = 20 + x^2 - 10 \cos(2\pi x) + y^2 - 10 \cos(2\pi y), (-10 \leq x, y \leq 10) \quad (5)$$

The Rastrigrin function has lots of local minimal points, but, among all the local minimal points, only the point $[0, 0]$ is the global optimal solution. The fitness functions of the PSO algorithm and the improved PSO algorithm for the Rastrigrin function are illustrated in Fig.2. As indicated in Fig. 2, the improved PSO algorithm seems to outperform the PSO algorithm.

Example 3. An Ackley function is stated as following

$$f(x, y) = 20 \exp[-0.2 \sqrt{\frac{1}{2} (x^2 + y^2)}] - \exp\left\{\frac{1}{2} [\cos(2\pi x) + \cos(2\pi y)]\right\} + 20 + e, (-10 \leq x, y \leq 10) \quad (6)$$

Similar to the second example, the Ackley function has also lots of local minima points, but, among all the local minimal points, only the point $[0, 0]$ is the global optimal solution. The iterative processes of the PSO algorithm and the improved PSO algorithm for the Ackley function are exhibited in Fig. 3. As evidenced in Fig. 3, the improved PSO algorithm shows superior to the PSO algorithm.

In general, the three functions have proved the superiority of the improved PSO algorithm over the PSO algorithm in increasing the convergence speed and reducing the iteration number.

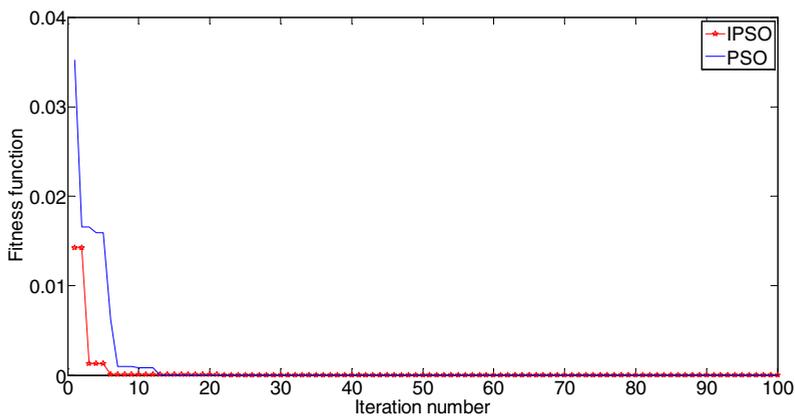


Fig. 1. The iterative processes of the Sphere function using the PSO algorithm and the improved PSO (IPSO) algorithm

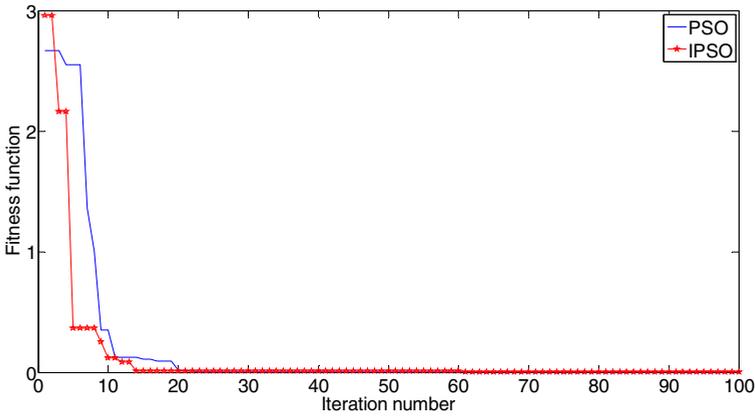


Fig. 2. The iterative processes of the Rastrigrin function using the PSO algorithm and the improved PSO (IPSO) algorithm

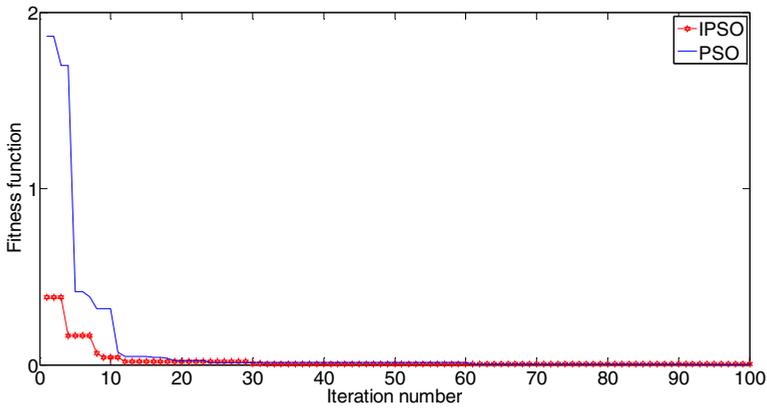


Fig. 3. The iterative processes of the Ackley function using the PSO algorithm and the improved PSO (IPSO) algorithm

5 Summary

Aimed at overcoming the drawbacks of the poor search capability and the low calculation efficiency, the improved PSO algorithm is presented in the paper. Subsequently, three functions were utilized to assess the performance of the improved PSO algorithm. The results showed that the improved PSO algorithm had some advantages over the PSO algorithm in increasing the convergence speed and reducing the iteration number. In addition, the paper points out that the improved PSO algorithm is seemingly an improvement on the PSO algorithm.



References

1. Chen, B., He, Z., Chen, X., Cao, H., Cai, G., Zi, Y.: A demodulating approach based on local mean decomposition and its applications in mechanical fault diagnosis. *Meas. Sci. Technol.* 22, 1–13 (2011)
2. Bhattacharya, S., Konar, A., Das, S., Han, S.Y.: A Lyapunov-Based Extension to Particle Swarm Dynamics for Continuous Function Optimization. *Sensors* 9, 9977–9997 (2009)
3. Rana, S., Jasola, S., Kumar, R.: A review on particle swarm optimization algorithms and their applications to data clustering. *Artif. Intell. Rev.* 35, 211–222 (2011)

A Scheme of LSB Steganography Based on Concept of Finding Optimization Pixels Selection*

Jianhong Sun^{1,**}, Yingjiang Li¹, Xiaohui Zhong², and Junsheng Li^{1,**}

¹ Engineering College of Honghe University, Yunnan Mengzi, 661100, China

² School of Power and Mechanical Engineering of WuHan University, Wuhan
430072, China

{sparkhonghe, hhljsh}@gmail.com

Abstract. Reduce the expected number of modifications per pixel (ENMPP) of cover image is a useful way to improve the imperceptibility in steganography. Many works related with this issue have been proposed in recent years. In this paper, a scheme of LSB steganography is proposed which based on concept of finding optimization pixels selection from cover image. The experimental results show that effect of reducing ENMPP is remarkable.

Keywords: Steganography, LSB, least significant bit, ENMPP.

1 Introduction

Steganography works as an effective method of hiding secret messages, thereby protecting the secret information from unauthorized or unwanted viewing. The advantage of steganography, over cryptography alone, is that secret messages do not attract attention to themselves. However, for against the deliberate cracker, a good steganography scheme needs more security to ensure that secret information can not be detected unless have attained authorized. Usually, the security of a steganography scheme depends on the imperceptibility to a certain degree.

In the past few decades, most of researches focused on transform domain technologies for their strong robustness property to against attacks; so long as many transform-based steganography technologies have been proposed, such like the Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT), Fractal Transform technologies. But the Spatial Domain technology which utilizes the least significant bits (LSBs) of images to embed large amount of data information has advantage of high payload, good visual imperceptibility, and extreme ease of implementation [1]. For get good imperceptibility in LSB steganography, the number of modifications pixels of cover image is expected to reduce during secret information embedding. Although, the number of modifications pixels is not only important factor influencing the security of the steganographic scheme. The choice of the cover object

* Pecuniary aid of the Education Department of Yunnan Province Research Fund(2010Y164)

Pecuniary aid of Honghe University Discipline Construction Fund (081203).

** Corresponding author.

and the character of modifications play an equally important role too. Nevertheless, it is true that for two steganographic schemes with the same embedding mechanism, the one that introduces fewer embedding changes will be more secure [2].

In our previous works, a Private Information Keeper (PIK) system based on LSBs technology was proposed in [3] [4]. The PIK system is simple and less security, only for the low level security applications such as keeping private photos and information for avoiding accidentally given away. In this paper, we discuss the method that how to reduce the number of modified bits of host image, resulting in improved imperceptibility.

2 LSB Steganography Review

The LSB steganography technologies are divided into two types: LSB replacement and LSB matching. In LSB replacement, the least significant bit of each selected pixel is replaced by a bit from the secret message [5]. Sometime the LSB is extended to LSBs (the lower 0~3bits of one byte). Obviously, the value of cover pixel will not be changed if the value of message bit m_i equals the value of cover bit c_i . According to statistical theory, the expected number of modifications per pixel (ENMPP) is 0.5. LSB matching is much complex more than LSB replacement. The secret information no longer just easy replaces the LSB of selected pixels. In order to get better imperceptibility and greater embedding capacity, many works based on LSB matching concept are constantly being proposed. In [6], matrix embedding was introduced by Ron Grandall and now it has become a general principle which can be applied to most steganographic schemes to improve embedding efficiency. Reference [7] present two approaches to matrix embedding for large payloads suitable for practical steganographic schemes-one based on random linear codes of small dimension and another based on Hamming codes. In [8], Mielikainen used a pair of pixels as a unit, embedded one message bit into the LSB of the first pixel, and a function of the two pixel values carries another bit of message, so that the ENMPP achieved 0.375. This method has been generalized by Xiaolong [1], and be named generalized least significant bit matching (G-LSB-M).

In this paper, we propose a steganography scheme based on optimization pixels selection (The concept of pixel selection was introduced in [6] to reduce change density of cover images.) to reduce the expected number of modifications per pixel of cover image in LSB replacement algorithm.

3 Proposed Scheme

General Embedding Process. The general embedding process of LSB replacement can be described as Fig.1. Where, $F([R,S],n,seed)$ is a function which random select n pixels from matrix $[R,S]$, $seed$ is a selected integer number which works as a seed of pseudo-random sequence generator and meanwhile it also is a secure key to the recipient for extracting secret information. R and S are the rows and columns of cover image. According to statistical theory, K times experiments with different pixels selection (selected by random function, $seed$ is one of variations) under the same experimental conditions, we have,

$$ENMP_{P_{mean}} = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K ENMP_P(seed_i) = 0.5$$

Obviously, the ENMPP will fluctuate around 0.5 when we choose a *seed* to generate a group of selected pixels from cover image. In general embedding process, the *seed* is come from the sender acting on a whim. However, the *seed* is an important parameter which resulting in the ENMPP is better/equal/worse than 0.5.

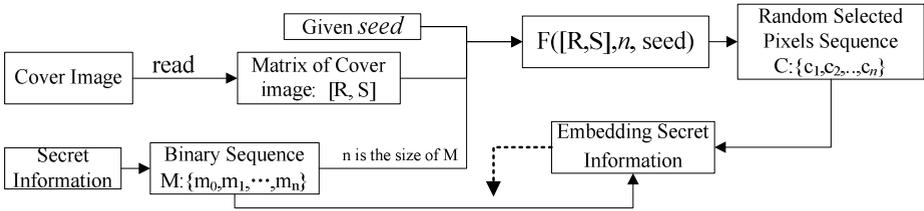


Fig. 1. General information embedding process

Proposed Scheme. For ensure the result is better than the mean value, 0.5. The information embedding process is modified as Fig.2 shown. Compare with general information embedding process, the proposed scheme requires offering a range [seed1, seed2] instead of parameter *seed*. It is worth noting that the binary sequence of secret information should do bitwise NOT operation then embedding $\sim M$ if the expression $|ENMP_{max}-0.5| \geq |ENMP_{min}-0.5|$ is true. In this way, we can ensure to get an optimization *seed* from [seed1, seed2] with relative-minimum ENMPP, and it satisfy the expression, $0.5 \leq ENMP \leq 1$.

A wide range of parameters [seed1, seed2] is good for getting a lower ENMPP, but meanwhile the computational complexity and executing time will increase.

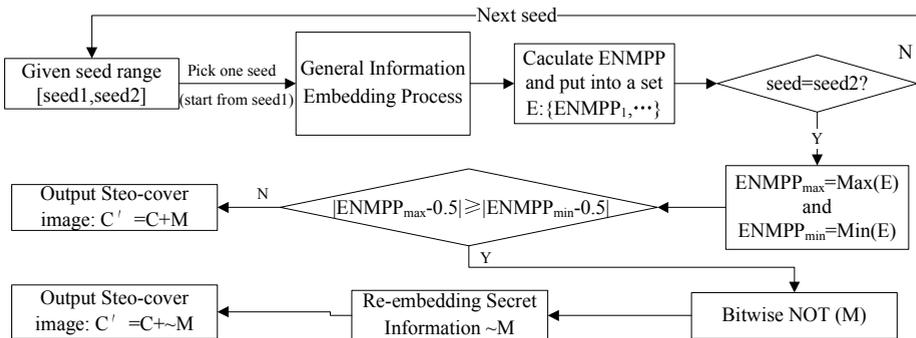


Fig. 2. The proposed information embedding process



4 Experimental Results and Valuations

In this section, we present the simulation results which simulated in MATLAB 2009a. The experiment examples are two 512×512 and two 256×256 standard gray images (lenna.bmp and baboon.bmp) work as cover images and 10 random given binary sequences with different length work as secret information candidates.

Experiment Results. The results are shown as Tab.1 and Tab.2. From Tab.1, as is illustrated by the length of secret information in the table the ENMPP has been on decrease, steadily decreasing from 0.4664 in 4848bits (lenna.bmp, embedding efficiency, $N/(R*S) = 1.85\%$) to 0.3077 in 104bits (lenna.bmp, $N/(R*S) = 0.04\%$). From the remarkable decline of ENMPP in Tab.1 and Tab.2, it goes without saying that small size secret information is easier to get better match, but ENMPP affected by embedding efficiency ($N/(R*S)$) is not much. Otherwise, big size secret information will result in time complexity increasing.

Table 1. Embedding experiment of proposed scheme (Seed range [100000, 110000])

| Secret Information | Cover Images | | | | | | Embedding Efficiency $N/(R*S)$ |
|--------------------|-----------------|--------|-------------|------------------|--------|-------------|--------------------------------|
| | Lenna (512×512) | | | Baboon (512×512) | | | |
| | Best Seed | ENMPP | Remark | Best Seed | ENMPP | Remark | |
| 4848 bits | 107660 | 0.4664 | Bitwise NOT | 101753 | 0.4717 | Bitwise NOT | 1.85% |
| 2784 bits | 108391 | 0.4659 | | 101661 | 0.4605 | | 1.06% |
| 1936 bits | 104888 | 0.4545 | | 105328 | 0.4582 | Bitwise NOT | 0.74% |
| 1192 bits | 100180 | 0.4430 | Bitwise NOT | 105154 | 0.4396 | Bitwise NOT | 0.45% |
| 896 bits | 103782 | 0.4397 | | 109088 | 0.4286 | | 0.34% |
| 776 bits | 107589 | 0.4343 | Bitwise NOT | 103087 | 0.4304 | | 0.30% |
| 752 bits | 108185 | 0.4136 | Bitwise NOT | 103485 | 0.4242 | Bitwise NOT | 0.29% |
| 688 bits | 101096 | 0.4172 | Bitwise NOT | 102442 | 0.4244 | Bitwise NOT | 0.26% |
| 408 bits | 100110 | 0.4044 | Bitwise NOT | 102816 | 0.3922 | Bitwise NOT | 0.16% |
| 104 bits | 109886 | 0.3077 | Bitwise NOT | 106314 | 0.3077 | Bitwise NOT | 0.04% |

Table 2. Embedding experiment of proposed scheme (Seed range [100000, 110000])

| Secret Information | Cover Images | | | | | | Embedding Efficiency $N/(R*S)$ |
|--------------------|-----------------|--------|-------------|------------------|--------|-------------|--------------------------------|
| | Lenna (256×256) | | | Baboon (256×256) | | | |
| | Best Seed | ENMPP | Remark | Best Seed | ENMPP | Remark | |
| 4848 bits | 109192 | 0.4703 | | 105125 | 0.4700 | Bitwise NOT | 7.40% |
| 2784 bits | 107658 | 0.4612 | Bitwise NOT | 102664 | 0.4619 | Bitwise NOT | 4.25% |

Table 2. (continued)

| | | | | | | | |
|-----------|--------|--------|-------------|--------|--------|-------------|-------|
| 1936 bits | 103315 | 0.4618 | Bitwise NOT | 102138 | 0.4525 | | 1.95% |
| 1192 bits | 108661 | 0.4463 | | 104599 | 0.4446 | Bitwise NOT | 1.82% |
| 896 bits | 102465 | 0.4342 | | 103953 | 0.4241 | | 1.37% |
| 776 bits | 106090 | 0.4304 | | 108537 | 0.4330 | | 1.18% |
| 752 bits | 109753 | 0.4309 | | 106645 | 0.4269 | | 1.14% |
| 688 bits | 108830 | 0.4259 | Bitwise NOT | 104364 | 0.4302 | Bitwise NOT | 1.05% |
| 408 bits | 109283 | 0.3995 | | 100381 | 0.3946 | | 0.62% |
| 104 bits | 105509 | 0.3173 | | 100192 | 0.3269 | | 0.16% |

Valuations. From the results we can see the proposed scheme is able to reduce the ENMPP through optimizing pixels selection. The imperceptibility is improved and the same concept also is able to be used in other algorithms for performance improvement. Furthermore, the *seed* as a secure key, it is calculated instead of chosen by the sender in the proposed scheme. The security is not reduced because the calculation process of *seed* concerning secret information, it is also an unknown parameter for cracker. Obviously, big *seed* is more secure, and does not affect the computation complexity of scheme. However, the proposed method is rarely contributing to the performance improved under big size of secret information embedded.

5 Conclusions

In this paper, we use the concept of finding the optimization pixels selection to embed secret information for reducing ENMPP, improving the imperceptibility of steganography. The experiment results shows that the ENMPP better than arithmetical mean is guaranteed in the proposed scheme. Be aimed at the big size secret information, we think it should be separated into several segment respectively embedding into different block of a cover image. In this way, it will generate multiple seeds and both the security and imperceptibility of steganography will be improved. This work has already on the list of our future works.

References

1. Li, X., Yang, B., Cheng, D., Zeng, T.: A Generalization of LSB Matching. IEEE Signal Processing Letters 16(2), 69–72 (2009)
2. Bierbrauer, J., Fridrich, J.: Constructing good covering codes for applications in steganography (2006), Information on <http://www.math.mtu.edu/~jbierbra/HOMEZEUGS/stegocovsurveyOct06.ps>
3. Sun, J., Yang, B., Gao, S., Li, J.: A Private Photos Keeper System Based on Spatial Domain Techniques. In: Proceeding of 2009 International Conference on Information Engineering and Computer Science, vol. 2, pp. 1222–1225 (2009)

4. Sun, J., Li, J., Li, Z.: An Improved Algorithm of Digital Watermarking Based on Wavelet Transform. In: WRI World Congress on Computer Science and Information Engineering, CSIE 2009, Los Angeles, California, USA, vol. 7, pp. 280–284 (2009)
5. YiYu, X., Wang, A.: Revisit LSB Matching. In: 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), pp. 410–413 (2010)
6. Crandall, R.: Some notes on steganography. Steganography Mailing List (1998), Information on: <http://os.inf.tu-dresden.de/westfeld/crandall.pdf>
7. Fridrich, J., Soukal, D.: Matrix Embedding for Large Payloads. IEEE Transactions On Information Forensics and Security 1(3), 390–395 (2006)
8. Mielikainen, J.: LSB matching revisited. IEEE Signal Process. Lett. 13(5), 285–287 (2006)

Research on Intelligent Generating Test Paper Based on Parallel Genetic Algorithm

Jianjun Li¹ and Meng Wang²

¹ Lushan College, GuangXi University of Technology, LiuZhou, China
31063@qq.com

² Department of Computer Engineering, GuangXi University of Technology,
LiuZhou, China
mwang007@163.com

Abstract. The aim of test paper composing is to compose an optimization test paper that satisfies the parameters which the user inputs, so the test paper composing problem is a classical multi-objective linear programming problem. This paper proposes an intelligent algorithm to generating test paper based on Parallel genetic algorithm, and provides a set of schemes of making papers of different degree of difficulties display in normal distribution. The algorithm adopts a new decimal system of subsection code, improves the traditional method of initializing the population and optimizes course of search. The experiment proves that this algorithm has better performance thus is more practical.

Keywords: Parallel genetic algorithm, artificial intelligence Intelligent Generating Paper.

1 Introduction

In the past decades, the rapid growth of the Internet has brought a great deal of changes in our educational environment. As one of the killer internet applications, the emerging e-learning is reshaping the instructional community and provides tremendous cost savings for both instructors and learners. In the e-learning context, one of the indispensable components is the e-examination system involving test generation, delivery, evaluation, and results publishing. In this paper, more emphasis was placed upon how to automatically generate tests meeting given constraint conditions specified by teachers.

The essentiality of automatic generating test paper is to generate a high quality paper according to consumer's requirements and generating algorithm. In the procedure of generating test paper, all the test questions are automatically chosen from item bank, and the generated paper is stochastic and reasonable.

Intelligent test paper auto-generating is an optimal combination problem with many constraints in math theory, the key of solving it is the design of effective algorithms. Some algorithms have been considered, such as select-random algorithm [1], backtrack algorithm, and intelligence algorithms [2]. The last type includes: test paper auto-generating based on IRT (Item Response Theory)[3], intelligent fixing test

paper based on granularity synthesis[4], fuzzy algorithm, target tree algorithm, differentiability algorithm, adaptive algorithm[5], particle swarm optimization algorithm, genetic algorithm(GA) etc.. In these algorithms, GA and improved algorithms more fit test paper auto-generating problem for its more extensive adaptability than other algorithms.

These objectives are incompatible, and cannot be satisfied at the same time. Most systems can only fulfill the objectives sequentially according to the prescriptive priority. So within relatively short time, it is difficult to solve the problem of generating a test paper. Therefore, under the circumstances that users require quite short time to generate test paper, we must reduce the multi-objective constraint and select some among the evaluation indexes of test paper. At present, many improvements have been made to the algorithms under the prerequisite of multi-objective.

This paper is organized as follows: Section 2 simply reviewed the Formalization of requirement of generating test paper. Section 3 presented using Genetic Algorithms to construct test paper. Section 4 proposed how to Simulation and analysis the system.

2 Formalization of Requirement of Generating Test Paper

Normally, algorithm of generating test paper is composed of a serial of regulations. And its requirement is generating a reasonable test paper, which satisfies special regulations.

For describing the requiremnet of generating test paper, we define the parameter target system of subject database: $T = \{t_1, t_2, \dots, t_n\}$, where $V(t_i) = \{v_1, v_2, \dots, v_n\} (t_i \in T)$ denotes the range of parameter t_i . Chosen target is given by $G = \{R_{t_1}, R_{t_2}, \dots, R_{t_n}\}$, where $R_{t_i} \subseteq V(t_i)$ is a set of parameter values, which expresses the range of a restricted subject on parameter t_i . During the process of generating test paper, subjects of the database can be seen as a set of parameter values. Accordingly, we can define the subject as follows:

$a = (id, vt_1, vt_2, \dots, vt_n), vt_i \in V(t_i)$, where id is the key of α in database and vt_i is the value of subject a in parameter t_i . $IB = \{a_1, a_2, \dots, a_n\}$ shows that the subject database is a finite set of subjects and if we give a choosing target G , then the subject set: $S(G, IB) = \{a | a \in IB, \forall R_{t_i} \in G, a.vt_i \in R_{t_i}\}$ can be obtained. In the following paper, $S(G, IB)$ is predigested by $S(G)$.

3 Parallel Genetic Algorithm

Parallel Genetic Algorithm (PGA) is a Genetic Algorithm adapt to parallel computers [4]–[6]. It is applicable to deal with complex optimization problem. There are three kinds of Parallel Genetic Algorithm: Master-slave model, thick granularity model and thin granularity model. Thick granularity model Parallel Genetic Algorithm is the most adaptable and widespread Genetic Algorithm, which can run on both

multiprocessor system and single-processor. In this paper, we choose thick granularity model Parallel Genetic Algorithm based on intermarriage strategy.

Premature convergence has been a main problem in Genetic Algorithms. In order to prevent premature convergence, the intermarriage strategy of human being is introduced into PGAs. Parallel Genetic Algorithm based on intermarriage strategy adopts multi-communities parallel evolution. In this method, the number of sub-communities is $M (M \geq 2)$. Elitists from different communities are intermarried when they satisfy with the intermarriage condition. And then the offspring of the elitists are copied to related communities. In order to preserve eminent genes, we adopt the strategy in which the eminent individuals are survived by competing the offspring and elitists. Then the eminent individuals are regarded as seed to participate in the next genetic process. Generating test paper based on parallel genetic algorithm can be represented as figure 1.

Subsection Code by The Decimal System. Suppose that the paper needs m questions and n test patterns altogether. Then each code is m bits and is divided to n segments. Each segment represents one test pattern. Suppose that the amounts of the number i test pattern is k_i , then code bit is $b(1)b(2) \dots b(k_1)b(k_1+1) \dots b(m)$. Among them : $\sum k_i = m$, $b(i)$ shows the serial numbers of examination question chosen in item pool.

The genetic generating test paper algorithm is designed by adopting the decimal subsection code. The method can overcome some obstacles, such as heavy calculation quantity for binary code, right accuracy restrained, etc. Furthermore, the method can

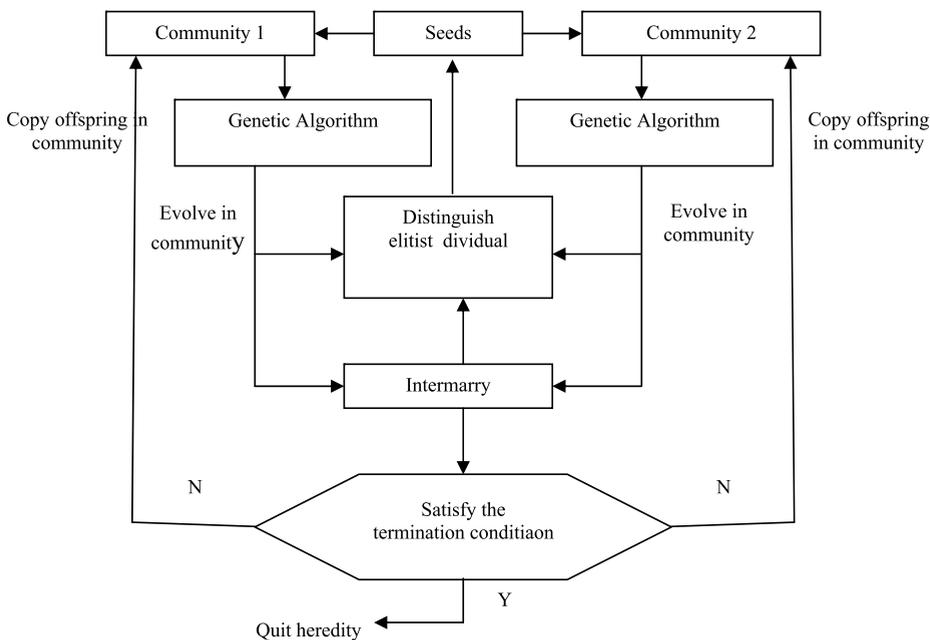


Fig. 1. Generating test paper based on parallel genetic algorithm

cels the course of coding and decoding. So it can relatively shorten the solving time and present good performance while evolving. Adopting subsection code also makes it easy to operate. Segment hybridization and segment variation also become easier. What's more, it helps the initialization of the population.

Confirming of The Individual Fitness Function $E(x)$. Based on “least-square methods” in curve fitting, fitness function is defined as follows:

$$E(x) = \sum_{i=1}^{10} (e_i)^2 \quad (1)$$

Among them: $e_i = B(x_i) - B'(x_i)$, awkward degree is below the mark of examination question which equals x_i and take the errors of the proportion and actual proportion. From adopting this method to define fitness function, the problem of generating test paper will be described as solving the minimum question.

In order to accelerate the optimizing speed and converge in a more excellent solving as soon as possible, penalty function (4) is set up. After one generation genetic finished, each chromosome's fitness function should be adjusted. The adjusted fitness function is:

$$E(x) = \exp(a * E(x)) * \text{namda} \quad (2)$$

If there is not an individual who meets the condition, $\text{namda} = 0.2$; If 80% of the individuals meet the condition, $\text{namda} = 0.8$; Otherwise, $\text{namda} = 1$. According to experiences, a fetches about 0.15.

Determine Initial Colony P0. In order to make initial population more diversified and further enhance the efficiency, we have improved the traditional initial algorithm for initializing every single segment of one chromosome. Improved algorithm is described as Fig.1.

This algorithm controls the uniqueness of visiting the serial number of examination question through generating examination question at random. According to producing times of random number, the traditional algorithm carries out random number for N times in best cases. But this possibility is as slim as 0. In general cases, the times of producing random number is greater than N . And this algorithm only produces N random number under any circumstance.

According to time of executing integer comparative operation, the times for traditional algorithm and this algorithm is the same under the best situation. it is N time. But in worst cases, the time of this algorithm is limited, and the time of traditional algorithm is infinite. Obviously, although this algorithm is the same as traditional algorithm in space complexity, it is superior to the traditional algorithm in time complexity.

Confirming Colony's Size M. Colony's size should be determined according to the need. In theory, M fetches a linear multiple of the length of a individual code, generally speaking, $m < M < 2 * m$, m is the quantity of examination questions included in the paper.

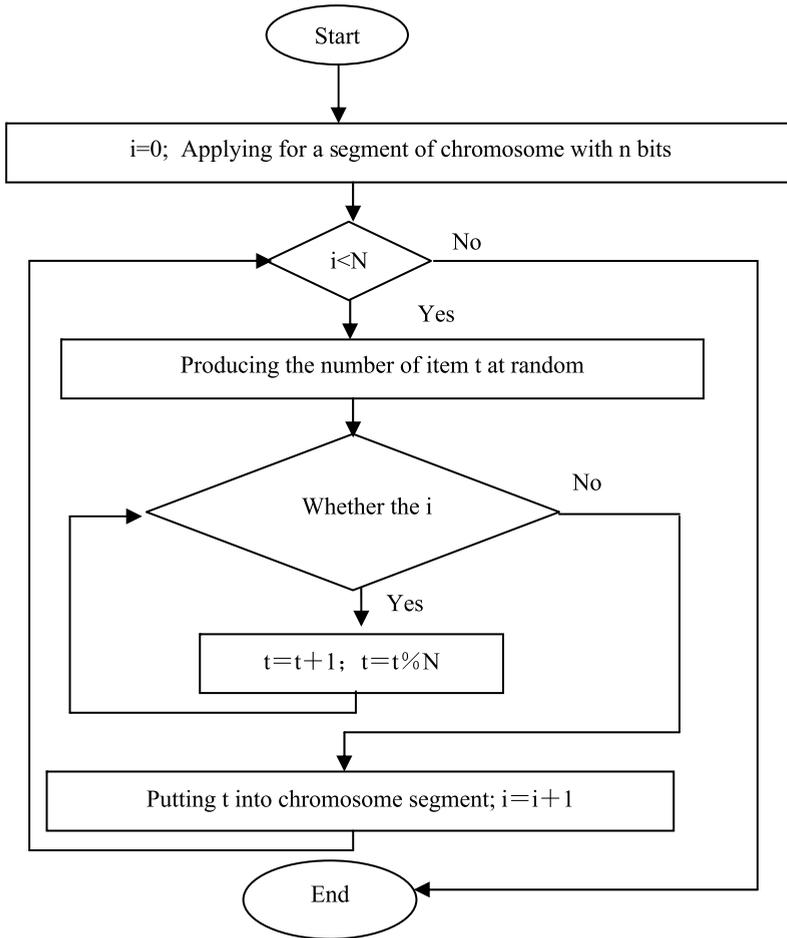


Fig. 2. The improved algorithm of generating test paper

Controlling Other Parameters and Variables of The Genetic Algorithm. In a simple genetic algorithm choosing probability P_r , hybridization probability P_c and variation probability P_m are secondary parameters.

According to choosing probability $P_r(x)$, population duplicate individual of colony to the new colony, that is determined by probability distribution corresponding to fitness function.

Genetic hybridization operators have many kinds; subsection code make dot-blot hybridization convenient. The hybridization course is carried on like this: Choosing two chromosomes in colony, with the i segment of two chromosome carries on hybridization sequentially according to probability P_c . New segment got this way may be illegal, since the number may be repeated, but even so it can only be repeated once at most, so these two new segments will need to be adjusted. As to repeated question number t that appeared, use new question No to replace t at position that it

appear for the second time, $t = (t+1)\%L$. Until new t is legal, L is the quantity of this test pattern in the exam pool. Probability P_c generally fetches about 0.5, we fetch $P_c = 0.6$.

In genetic algorithm, variation operator change some bit on chromosome bunches with a light probability P_m at random. The course is as follows: Investigate a certain bit $randn$ in this segment at random, make this question number as t , the question number t produced may be illegal to make a variation, so it need to be adjusted, and use new question No to replace t , $t = (t+1)\%L$ until new t is legal, L is the quantity of this test pattern in the exam pool.

4 Simulation and Analysis

The algorithm of this paper is realized with VC ++ language. Test pool is the test pool of "assembler language design program "which is applied for computer undergraduate course. There are 50 items in this test paper.

The size of colony adopted by generating test paper is 60. The biggest genetic algebra GNUM is established as 500 generations.

Table 1. Comparison sheet between the average of degree of difficulty of the paper and user's expectation

| Time of experiment | Expectation of degree of difficulty | Practical average degree of difficulty of generating test paper | Distribution function of generating test paper fitted |
|--------------------|-------------------------------------|---|---|
| 1 | 0.7 | 0.7161 | N(0.7029,0.0083) |
| 2 | 0.65 | 0.6288 | N(0.6502,0.0106) |
| 3 | 0.85 | 0.8219 | N(0.8478,0.0091) |

Through a lot of similar experiments, we find that the best average degree of difficulty of test paper generated through this algorithm is about the same as what users expect. And then we count corresponding P_i of 10 x_i . We utilize these 10 points (x_i, P_i) to fit the curve through least-square method. The fitting result is distribution function for degree of difficulty coefficient for the paper. From the experimental result, we can find that the paper generated can roughly display in a kind of normal distribution. And function expectations of normal distribution and users' expectation are pretty much the same. The result can be accepted. Therefore the calculating method of degree of difficulty distribution in this paper is feasible. It can guarantee that the degree of difficulty of the paper presents normal distribution.

The expectation of degree of difficulty is controlled as 0.75. We get three papers that the value of fitness function is below 0.05. Now we can get the advantage of this algorithm through comparing this algorithm with the algorithm taking out examination paper at random and the genetic algorithm that does not adopt penalty function.



Table 2. Comparison sheet of running time for three kinds of algorithms

| Time of experiment | Time of the algorithm taking out examination paper at random | Time of the genetic algorithm without penalty operator | Time of the algorithm in this paper |
|--------------------|--|--|-------------------------------------|
| 1 | 17.3s | 9.3s | 3.7s |
| 2 | 16.8s | 7.9s | 1.1s |
| 3 | 25.2s | 3.8s | 2.1s |
| 4 | 28.6s | 4.5s | 3.6s |
| 5 | 15.6s | 7.7s | 51s |

We can learn through the experiment that the system performance with the value of fitness function below 0.05 is very good. Because this algorithm introduces penalty operator, the speed of genetic and optimization can be accelerated accordingly. The result tends towards more excellent solution very fast. Therefore, the algorithm in this paper is an effective method to auto-generating test paper.

Acknowledgment. This paper is Supported by the Educational Commission of GuangXi Province of China (200808LX341, 200808LX338).

References

1. Wang, M., Jin, H., Wang, X.: Research on set at random algorithm in intelligent generating test paper. *Computer Engineering and Design* 27(19), 3583–3585 (2006)
2. Wang, Y., Zhang, Z., Cui, J.: Mathematical Model and Algorithm of Intelligent Test Paper Auto-generation System from Item Pool. *Systems Engineering —Theory & Practice* (9), 85–97 (2004)
3. Dong, M., Huo, J., Wang, X.: Model Management System for IRT-based Test Construction. *Journal of University of Science and Technology of China* 34(5), 612–617 (2004)
4. DeJong, K.A., Spears, W.M.: Using Genetic Algorithms to solve NP complete problems. In: *Processings of the Third International Conference on Genetic Algorithms*, pp. 124–132 (1989)
5. Wei, J.: Intelligent Test Paper Composition Systems Based on SOM. *Journal of Liaoning Normal University (Natural Science Edition)* 28(3), 283–284 (2005)
6. Luo, X., Sun, M., Tsou, B.K.: Covering Ambiguity Resolution in Chinese Word Segmentation Based on Contextual Information. In: *Proceedings of the 19th COLING*, pp. 598–604 (2002)
7. Li, J., Kwan, R.S.K.: A fuzzy genetic algorithm for driver scheduling. *European Journal of Operational Research* 147(2), 334–344 (2003)
8. Andre, J., Siarry, P., Dognon, T.: An improvement of the standard genetic algorithm fighting premature convergence in continuous optimization. *Advances in Engineering Software* 32, 49–60 (2001)

Design of Fault Diagnosis System for the Parameter-Depended Control System

Jing Zhou

School of Mathematics & Computer Science, JiangHan University, 430056,
Wuhan, China
zhou_8132@sina.com

Abstract. A robust parameter-depended reduced order(RPRO) fault detection filter(FDF) is designed. Contrary to the parameter-depended uncertainty system, the order of the linear matrix inequalities is reduced, then the RPRO fault detection and fault isolated filters are constructed. Then a RPRO fault isolation filter is designed for occurrence of both actuator fault and sensor fault in the aircraft's closed-loop control system, and fault diagnosis system is structured based on the fault isolation filters. Through the output of the fault diagnosis system, we can alarm the fault timely and the advantages of this approach are highlighted.

Keywords: fault detection and isolation, RPRO, parameter-depended system, bias fault, feature points.

1 Introduction

Due to the control system becomes more and more complicate, the system's reliability and security desires more attention. Especially to the aircraft's control system, the fault's occurrence will lead to inestimable loss. How to detect and isolate the fault effectively online, hence improve the reliability and security of the aircraft's control system is a very meaningful and valuable research topics.

A very interesting application of H_{∞} estimation algorithms to aircraft inflight ice detection was given in literature [1]. An approach in which the Luenenberger observer gains from a bank of dedicated Kalman filters are optimized for fault detection using H_{∞} optimization was proposed[2]. An application of the approach to a simplified longitudinal flight control system (i.e. short-period dynamics and LTI) with noise and disturbances resulting from Butterworth filters and considering system parameter fluctuations was given. And the H_{∞} -FDI Ricatti-based approach is used to design fault detection filters for an inverted pendulum. They use parametric linear fractional transformations (LFT) to represent uncertainty and then estimate the parameter uncertainty in order to declare the fault. Unfortunately the application was not fully carried out in this case (i.e. no simulations)[3].

2 Parameter-Depended Reduced Order Fault Detection Filter Design

RPRO fault detection filter. An aircraft's closed-loop system with the actuator fault or sensor fault is considered. The system shown as Eq. 1 is a LTI system which is constructed at several feature points, with the parameter-depended uncertainties. where $x(t) \in R^n$ means system state, $y(t) \in R^m$ describes measurement output, $f(t) \in R^{n_f}$ describes fault function, $d(t) \in R^{n_d}$ means bounded interference, $u(t) \in R^{n_u}$ means control input signal $B_f f(t)$ describes actuator fault and $D_f f(t)$ means sensor fault.

$$P: \begin{cases} \dot{x}(t) = A(\alpha)x(t) + B_u(\alpha)u(t) + B_d(\alpha)d(t) + B_f(\alpha)f(t) \\ y(t) = C(\alpha)x(t) + D_u(\alpha)u(t) + D_d(\alpha)d(t) + D_f(\alpha)f(t) \end{cases} \quad (1)$$

For the purpose of residual generation, the so-called fault detection filter in the Eq. 2 is used, and system residual $r(t)$ is defined as $r(t) \equiv z(t) - y(t)$.

$$\begin{cases} \dot{x}_F = A_F x_F + B_F y \\ z = C_F x_F + D_F y \end{cases} \quad (2)$$

Then adhibiting a standard error model $r'(s) = W_f(s)f(s)$ and the residual error is defined as $\tilde{r} = r' - r = Wf - r$. Thus, the block diagram for the residual error \tilde{r} with shaping filter W and FD filter F is shown in Fig.1. The two shaping filter W_d and W_F are used to regulate the system error which can refer to literature [4].

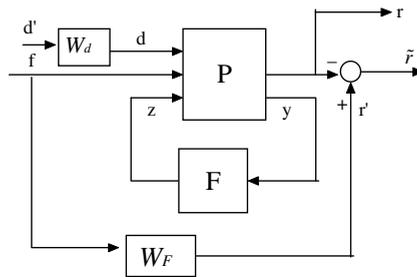


Fig. 1. The block diagram for the FD structure

Including the weighting functions W_d and W_F into the model (Eq. 1) leads to the standard linear-fractional form, then including the filter (Eq. 2) and using some linear fractional algebra, it can be verified that the losed-loop model $R = (\tilde{P} \times F)$ admits the state space equation described as literature[4].

Applying the BRL and the problem of robust fault detection filter for system (Eq. 1) transforms to existing decision variables K and P to satisfy the LMI inequalities. Based on the design method of the full order FD filter, we can decompose the LMI inequalities's decision variable K to $K_{11}(l_F \times l_F), K_{12}(l_F \times k), K_{21}(k \times l_F), K_{22}(k \times k)$, where l_F means full order filter's order, k means reduced order filter's order, \tilde{K}_{21} is a $k \times k$ dimension matrix ($l_F \geq k$) [5]. Considering the system (Eq. 1) which contains a uncertainty parameter α . According to affine secondary stability (AQS) theory [6], a parameter-depended Lyapunov matrix $P(\alpha)$ can be found and satisfies the Eq. 3 with all the P_i 's.

$$P(\alpha) = \sum_{i=1}^s \alpha_i P_i(\alpha), \alpha \in \Gamma \tag{3}$$

$$\begin{bmatrix} -(K_{11} + K_{11}^T) & -(\tilde{M}_1^T + M_2^T) & K_{11}^T \tilde{A} + \tilde{B}_F \tilde{C}_i + P_{1i} & \tilde{A}_F + \hat{P}_{3i}^T & K_{11}^T \tilde{B}_{d,i} + \tilde{B}_F \tilde{D}_{d,i} & 0 & K_{11}^T & \tilde{M}_1^T \\ * & -(M_1 + M_1^T) & M_2 \tilde{A} + \hat{B}_F \tilde{C}_i + \hat{P}_{2i} & \hat{A}_F + \hat{P}_{2i} & M_2 \tilde{B}_{d,i} + \hat{B}_F \tilde{D}_{d,i} & 0 & M_2 & M_1^T \\ * & * & -\lambda P_{1i} & -\lambda \hat{P}_{3i}^T & 0 & \tilde{C}_F^T & 0 & 0 \\ * & * & * & -\lambda \hat{P}_{2i} & 0 & -\hat{C}_F^T & 0 & 0 \\ * & * & * & * & -\gamma_d I & \tilde{D}_{d,i}^T (I - \hat{D}_F)^T & 0 & 0 \\ * & * & * & * & * & -\gamma_d I & 0 & 0 \\ * & * & * & * & * & * & \frac{1}{\lambda} P_{1i} & \frac{1}{\lambda} \tilde{P}_{3i}^T \\ * & * & * & * & * & * & * & \frac{1}{\lambda} \hat{P}_{2i} \end{bmatrix} < 0 \tag{4}$$

Then the k order RFD filter of the parameter-depended system (Eq. 1) satisfies the inequalities such as InEq. 4 (because of the limit of paper's length, the other two LMI inequalities of input u and bounded interference d are omitted, which are similar to InEq. 4).

The design process of the parameter-depended reduced order FD filter can be transformed to the design of filter coefficient (A_F, B_F, C_F, D_F) , which can make InEq. 3 and the other two LMI inequalities be solvable and satisfy Eq. 5, where $\hat{P} > 0$.

$$\min_{\hat{A}_F, \hat{B}_F, \hat{C}_F, M_1, M_2, K_{11}, \hat{P}, \gamma} \{ \gamma_d + \gamma_f + \gamma_u \} \tag{5}$$

The filter matrices A_F, B_F, C_F, D_F can be derived by means of the Eq. 6.

$$A_F = \tilde{K}_{21}^{-T} \hat{A}_F \tilde{K}_{21}^{-1} K_{22}, B_F = \tilde{K}_{21}^{-T} \hat{B}_F, C_F = \hat{C}_F \tilde{K}_{21}^{-1} K_{22}, D_F = \hat{D}_F \tag{6}$$



Residual Evaluation. After designing of FD filter, the remaining important task for FD is choose a threshold J_{th} , and use the following logical relationship for fault detection:

if $\|r\|_{\tau} > J_{th}$, then alarm the fault, else means no fault. The threshold J_{th} can be found in InEq. 7.

$$J_{th} \equiv \|r\|_{\tau} = \|T_{ru}u + T_{rd}d\|_{\tau} \leq \|T_{ru}\|_{\infty} \|u\|_{\tau} + \|T_{rd}\|_{\infty} \|d\|_{\tau} \leq \gamma_u \|u\|_{\tau} + \gamma_d \|d\|_{\tau} \quad (7)$$

Where γ_u and γ_d come from the optimization problem described as Eq. 5, and $\|d\|_{\tau}$ is the worst disturbance acting on the plant. Thus, in InEq. 7, $\|d\|_{\tau}$ is evaluated off-line while u is assumed to be known and $\|u\|_{\tau}$ is calculated on-line.

3 Parameter-Depended Control System's Fault Diagnosis

An aircraft's feature point model. A FDI filter is designed for the aircraft model in Eq. 1 which is constructed in literature[7]. The coefficients in Eq. 1 is given as below(the feature point is 0.35s[7]) and uncertainties parameter α, β satisfy $|\alpha| \leq 1, |\beta| \leq 1$.

Fault Model. Two components-the servo actuator and pitch angle rate sensor maybe lose effectiveness in the aircraft's control system. The servo actuator fault model can be formed as $f_{out} = L_1(\rho)f_1(t)$, where fault distribution function $L_1(\rho) = B$. When the bias fault of actuator occurs, set $f_1(t) = b$. The pitch angle rate sensor fault model is formed as $f_{sensor} = L_2(\rho)f_2(t)$, when the sensor bias fault occurs, set the fault vector $f_2(t) = c$.

Fault diagnosis system design. The fault diagnosis system for the aircraft's control system is designed according to Fig.2, which can detect and isolate the actuator fault and sensor fault effectively.

The fault vectors (f_1, f_2) in Fig. 2 indicate the actuator fault and the pitch angle rate sensor fault. The vector u describes the control input of system, and d means the disturbance. The fault detection system consists of two parts-the fault detection filter group and the fault diagnosis unit. The fault detection filter group consists of two FD filters, and the mapping of the output residual collection $\{r_1, r_2\}$ and fault collection f_1, f_2 can be configurated according to the requirements of fault diagnosis. The fault diagnosis unit can inquire the given mapping table according to the residual collection produced by FD filter group and fix the faulty mechanism. In order to isolate the actuator fault and the sensor fault, the mapping of residual and fault collection can be set as one of modes in Fig. 3.

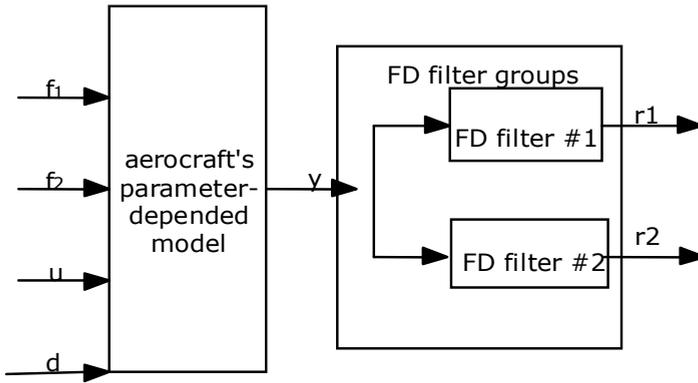


Fig. 2. Structure of the fault diagnosis system for aircraft control system

| | r_1 | r_2 | | r_1 | r_2 | | r_1 | r_2 | |
|-------|-------|-------|-----|-------|-------|-----|-------|-------|---|
| f_1 | ✓ | ✓ | | f_1 | ✓ | × | f_1 | ✓ | × |
| f_2 | ✓ | × | | f_2 | ✓ | ✓ | f_2 | × | ✓ |
| (a) | | | (b) | | | (c) | | | |

Fig. 3. Mapping of residual of FD filters group and fault

(✓ means that residual is sensitive to fault, × means that residual is not sensitive to fault)

The several mappings shown in Fig. 3 can be used for the fault detection of actuator fault and sensor fault. The mapping of the output residual and mechanism fault can be set as Fig. 3(c), that means the residual of fault detection filter #1 is sensitive only to actuator fault and the residual of fault detection filter #2 is sensitive only to sensor fault.

4 Simulation Example

FD filter Design. The detection for actuator fault is designed firstly. The fault signal is simulated as a pulse of unit amplitude that occurs from 30 to 40 seconds and is zero elsewhere. The unknown input d is assumed to be band-limited white noise with power 0.0005 and the upperbound of $\|d\|_r$ is set to 0.15. This value is used to calculate J_{th} using InEq. 7. With these conditions, the 1-order RFD filter data A_F, B_F, C_F, D_F can be obtained from Eq. 6. The faults are usually detected in the frequency range between 0 and 2 rad/s with a -20dB/dec roll-off at the lower frequencies, so the

weighting functions W_d and W_F are set to $W_d = W_d(s) = \frac{0.14s + 0.0024}{s + 0.01}$,

$$W_F = W_F(s) = \frac{0.01s + 4}{s + 20}.$$

Fault Diagnosis System Design. When the actuator and sensor fail at the same time, the fault isolation is required. The two FD filters are designed as the method described in section-*FD filter design* firstly. When the actuator bias fault occurs, the 1-order FD filter 1# can be get from Eq. 6. Meanwhile, when the sensor bias fault occurs, the 1-order FD filter 2# can also be resolved by Eq. 6. The two thresholds corresponding to two FD filters can be resolved by InEq. 7. Then the whole fault diagnosis system can be constructed as Fig. 2.

In simulation, the fault vectors $f_1(3-6s)$ and $f_2(4-8s)$ simulate the actuator and sensor’s bias faults respectively. The value of the control input u , disturbance d , upper bound $\|d\|_r$, weighting function W_d and W_F refer to *FD filter Design*. The two thresholds J_{th1} , J_{th2} can be resolved by InEq. 7, and the two 1-order RFD filters (1#, 2#) data A_F, B_F, C_F, D_F get from Eq. 6 respectively are shown in Table 1 and Table 2.

The output residuals ($\{r_1, r_2\}$) of fault detection filters group are shown in Fig. 4.

Table 1. 1# FD filter coefficients

| t[s] | A_F | B_F | C_F | D_F |
|-------|---------|--------------|--------|--------|
| 0.35 | -0.9459 | -9.0998e-005 | 0.0018 | 1.0000 |
| 19.2 | -2.6573 | -4.8005e-007 | 0.0015 | 1.0000 |
| 33.8 | -2.9999 | -4.6052e-007 | 0.0011 | 1.0000 |
| 131.2 | -3.8660 | -6.0755e-007 | 0.0019 | 1.0000 |

Table 2. 2# FD filter coefficients

| t[s] | A_F | B_F | C_F | D_F |
|-------|---------|--------------|--------|--------|
| 0.35 | -0.3116 | -1.2787e-005 | 0.0071 | 1.0000 |
| 19.2 | -2.6579 | -5.2668e-007 | 0.0160 | 1.0000 |
| 33.8 | -2.9968 | -8.1288e-007 | 0.0238 | 1.0000 |
| 131.2 | -3.8720 | -6.6335e-007 | 0.0273 | 1.0000 |



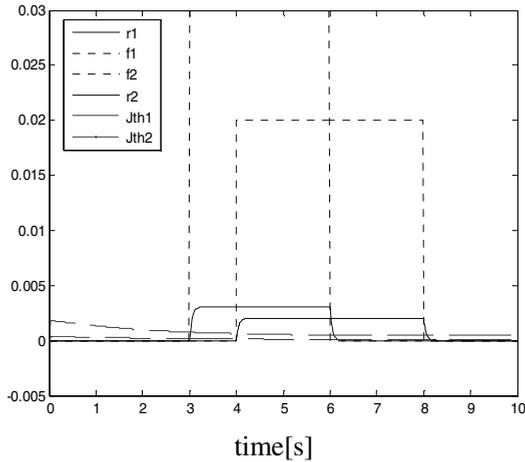


Fig. 4. Output of the 1-order FD filters group

In Fig. 4, when the fault vector $f_1(3-6s)$ occurs at 3s, the residual r_2 keeps zero, but residual r_1 exceeds the threshold J_{th1} quickly. When fault vector $f_2(4-8s)$ occurs at 4s, residual r_2 does not keep zero and exceeds the threshold J_{th2} quickly. When the fault vector $f_1(3-6s)$ disappears at 6s, the residual r_1 returns to J_{th1} quickly, but r_2 is unaffected. When the fault vector $f_2(4-8s)$ disappears at 8s, the residual r_2 will return to the threshold J_{th2} . The effectiveness of the 1-order FD filter can be shown from the simulation result, and the fault detection and diagnosis are realized.

5 Summary

A group of fault detection and isolation (FDI) filters is designed offline for the aircraft's feature-point model which is constructed in literature[7]. The tables of the FDI filters coefficients are given, which are constructed at feature-point 0.35s, and the FDI filters based on other feature-points can also be constructed similarly. Then the tables of the FDI filters are called by the aircraft's control system during the whole flight envelope, and the corresponding fault detection and isolation are realized.

References

1. Melody, J.W., Hillbrand, T., Basar, T., Perkins, W.R.: H_∞ parameter identification for in-flight detection of aircraft icing: The time-varying case. *Control Engineering Practice* 9(12), 1327–1335 (2001)
2. Ibaraki, S., Suryanarayanan, S., Tomizuka, M.: H_∞ Optimization of Luenberger state observers and its application to fault detection filter design. In: *IEEE Conference on Decision and Control, Orlando, FL*, pp. 1011–1016 (2001)

3. Felicio, P., Stoustrup, J., Niemann, H., Lurie, P.: Applying parametric fault detection to a mechanical system. In: American Control Conference, Anchorage, AK, pp. 1024–1029 (2002)
4. Zhou, J., Huang, X.H.: Application of a New Fault Detection Approach to Aircraft's Closed-Loop Control System. In: Xiong, C.-H., Liu, H., Huang, Y., Xiong, Y.L. (eds.) ICIRA 2008. LNCS (LNAI), vol. 5315, pp. 1223–1232. Springer, Heidelberg (2008)
5. Young, M.K., John, M.: A New Approach for Robust and Reduced Order Fault Detection Filter Design. In: Proceeding of the 2007 American Control Conference, New York City, USA, pp. 11–13 (2007)
6. Gahinet, P., Apkarian, P., Chilali, M.: Parameter dependent Lyapunov functions for real parametric uncertainty. IEEE Transactions on Automatic Control 41(2), 436–442 (1996)
7. Zhou, J., Huang, X.H.: Adaptive Fault-tolerant Control in Aircraft's Control System. Journal of Wuhan University of Technology 31(5), 106–109 (2009)

Research of Method about Geomagnetic Sensor Data Processing by Optimal Idea

Chanyuan Liu

School of Mechatronic Engineering, North University of China, Taiyuan,
Shanxi, 030051
lcy@nuc.edu.cn

Abstract. A method of optimal idea, in the paper, is used to process geomagnetic sensor data. The curve fitting by use of the method is more convenient than least square method (LSM). It adapts especially to process nonlinear curve fitting. Circular curve equation is fitted depending on a set of geomagnetic sensor data. It proves that the way is convenient and feasible.

Keywords: Optimal idea, LSM, nonlinear.

1 Introduction

Helicopter model aircraft flying in the air, Real-time calculating the forward direction vector is need. Because of rc helicopter regular level flight, the roll Angle and elevation Angle approximate to zero. So a parallel to the ground of the two-dimensional coordinate system can satisfy the requirement of measurement vector direction. Magnetic sensors installed in the helicopter, through cutting magnetic field lines, When the helicopter level a whirl, output a set of measurement data, In order to calculate the solution direction vector OV. Figure 1. Because of their measurement error, such as sensors. Data center is not the origin of coordinates, From figure 1 It can be seen than error vector is generated because O the center of circle C becomes O1 the center of circle C1, Direction vector OV became OV1. Only finding out circle position O1 of data point curve fitting circle, Just errors are eliminated. The most common method of curve fitting is the best square approximation. To some of discrete measured data, the best square approximation of the curve fitting is also called LSM. The least square method are generally used to handle linear curve fitting, this paper is nonlinear curve fitting of solving the center position according to known data. A similar problem often encountered in engineering calculations. To Deal with the problem, need to use more complex nonlinear least squares, not grasp the engineering staff. But problems such as nonlinear optimization problem is very easy to deal with.

Optimization method is the optimization theory and computing technology in engineering design, and it's essence is multivariate function minimization problem. The idea is very clear to solve this problem with optimization, and need not consider nonlinear problems.

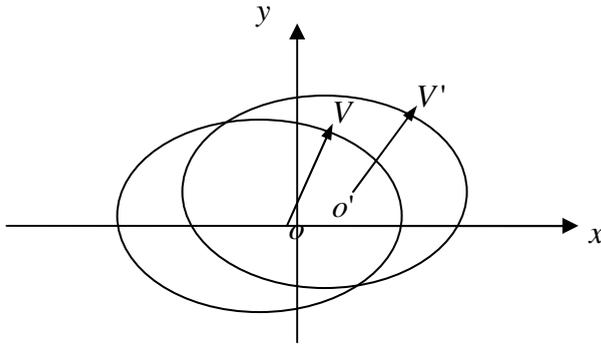


Fig. 1. Schematic diagram of data error

2 Algorithm Description

Least-Square Thoughts

$f(X)$ is obtained by experiment or observation, and it is usually given by the function table $(x_i, f(x_i)), i = 1, 2, \dots, m$. If you ask $\varphi(x)$ approximating function curve $f(x)$, to request $\|\delta\|_2^2 = \sum_{i=0}^m [f(x_i) - \varphi(x_i)]^2 = \min$, if $\varphi(x) \in \Phi = span\{\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x)\}$, make

$$F(a_0, a_1, \dots, a_n) = \sum_{i=1}^m [\sum_{j=0}^n a_j \varphi_j(x_i) - f(x_i)]^2 \rho_i \text{ takes the minimum.} \quad (1)$$

Here, a_0, a_1, \dots, a_n is n undetermined coefficients. ρ_i point x_i of weights, $\varphi_j(x_i), j = 1, 2, \dots, n$ is a function of n linearly independent. If $\varphi_j(x_i)$ meet Haar conditions. The formula A has a unique solution, $a_j = a_j^*, j = 1, 2, \dots, n$,

makes $\varphi(x) = \sum_{j=0}^n a_j^* \varphi_j(x)$. According to the principle of minimum squares approaching $f(x)$ [1].

When $(x_i, f(x_i)), i = 1, 2, \dots, m$ data distribution roughly graphics cannot use $a_j \varphi_j(x_i), j = 1, 2, \dots, n$ A linear combination of the said. It becomes a nonlinear least squares curve fitting approximation problem. More difficult to solve, Be familiar with relevant theoretical [2]. And for this problem can use another thought to look at it, That is, it can be to solve optimization problems as.



Optimization Mathematical Model

The least squares problem, Equation (1) constraints can be expressed as a mathematical model of the optimization problem.

Here, Objective function

$$F(a_o, a_1, \dots, a_n) = \min_{a_i} \sum_{i=1}^m [\sum_{j=0}^n a_j \phi_j(x_i) - f(x_i)]^2 \quad ; \quad \text{Optimization of}$$

parameters is a_o, a_1, \dots, a_n [2].

In order to verify the validity of this approach, take for example equation circle:

$$(x - 3)^2 + (y - 4)^2 = 9,$$

Here, $a_1 = 3, a_2 = 4, a_3 = 3$, Select the fitting data shown in Table 1.

| i | x_i | $f(x_i)$ | i | x_i | $f(x_i)$ | i | x_i | $f(x_i)$ |
|-----|-------|----------|-----|-------|----------|-----|-------|----------|
| 0 | 0 | 4.15 | 5 | 2.5 | 1.042 | 9 | 4.5 | 1.38 |
| 1 | 0.5 | 2.34 | 6 | 3 | 7.16 | 10 | 5 | 6.24 |
| 2 | 1 | 6.23 | 7 | 3.5 | 1.05 | 11 | 5.5 | 2.33 |
| 3 | 1.5 | 1.38 | 8 | 4 | 6.85 | 12 | 6 | 4.11 |
| 4 | 2 | 6.83 | | | | | | |

Optimization mathematical model are as follows, $f(a_1, a_2, a_3) = \min$

$$\sum_{i=0}^{12} [(x_i - a_1)^2 + (y_i - a_2)^2 - a_3^2]^2 .$$

Take numerical iterative formula step length $h = 0.01$. is solved using MATLAB optimization tool, The result is $a_1 = 3.012, a_2 = 4.005$. When the step size A further shortened, Calculation accuracy can be improved. That deal with the optimization problem of nonlinear least squares approximation is very simple and effective.

3 Magnetic Sensor Data Processing

Table 2 is Measured by the geomagnetic sensor model aircraft helicopter level of evaluation as part of orientation data. To find out the data points circle locations O1 can eliminate measurement errors, Figure 1.

Section 2.2, as shown by optimization method, taken the step size $h=0.1$, With round equation $(x_i - c)^2 + (y_i - c_1)^2 = R^2$ fitting beg to circle O1 coordinates, Solution of the problem is $c = 390.00, c_1 = 725.00, R = 265.00$.



Table 2. One revolution of the helicopter horizontal part of the data measured

| x_i | y_i |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 196 | 902 | 389 | 966 | 575 | 874 | 650 | 651 | 443 | 496 | 232 | 577 | 149 | 815 |
| 206 | 911 | 407 | 963 | 576 | 872 | 644 | 628 | 439 | 496 | 205 | 599 | 166 | 860 |
| 218 | 923 | 412 | 964 | 581 | 871 | 642 | 623 | 409 | 495 | 202 | 604 | 170 | 865 |
| 235 | 933 | 449 | 954 | 583 | 866 | 640 | 616 | 409 | 496 | 201 | 608 | 165 | 869 |
| 243 | 938 | 454 | 950 | 602 | 844 | 620 | 582 | 393 | 499 | 187 | 623 | 166 | 874 |
| 249 | 941 | 463 | 947 | 604 | 843 | 615 | 577 | 387 | 499 | 159 | 659 | 195 | 909 |
| 266 | 950 | 497 | 930 | 604 | 837 | 593 | 562 | 356 | 506 | 162 | 665 | 202 | 907 |
| 271 | 953 | 502 | 930 | 613 | 826 | 592 | 558 | 352 | 508 | 156 | 670 | 203 | 910 |
| 311 | 963 | 507 | 925 | 618 | 821 | 565 | 534 | 346 | 510 | 154 | 675 | 214 | 923 |
| 317 | 967 | 511 | 924 | 636 | 790 | 562 | 533 | 340 | 511 | 141 | 723 | 226 | 929 |
| 318 | 964 | 516 | 922 | 642 | 766 | 545 | 524 | 304 | 527 | 137 | 751 | 239 | 935 |
| 338 | 966 | 520 | 918 | 642 | 759 | 508 | 506 | 282 | 535 | 142 | 760 | 239 | 935 |
| 344 | 966 | 545 | 899 | 646 | 754 | 505 | 506 | 280 | 540 | 144 | 797 | 265 | 949 |
| 377 | 967 | 555 | 894 | 657 | 711 | 484 | 500 | 248 | 559 | 145 | 802 | 269 | 950 |
| 382 | 967 | 571 | 878 | 655 | 705 | 466 | 498 | 243 | 561 | 148 | 808 | | |

4 Conclusion

With the optimization method for solving nonlinear least squares curve fitting method is effective and reliable. For engineering workers, Thinking to solve such problems with the concept of optimization is more clear and explicit.

References

1. Li, Q., GuangZhi, BaiFengbin: Principles of Numerical Calculation, pp. 56–57. Tsinghua University Press, BeiJing (2000)
2. Sun, J.: Optimization of Mechanical Design, pp. 59–60. Machinery Industry Press, BeiJing (1999)
3. Shi, Y.: MATLAB language——TOOLBOX Practical Guide, pp. 111–113. Northwestern Polytechnical University Press, XiAn (1999)

Design and Implementation of a Communication Simulation Training System Based on HLA

Chen Bin and Li Juan

College of Electronic Engineering, Naval University of Engineering Wuhan,
China, 430033
sea_chb@sohu.com

Abstract. Simulation training of communication equipment can remove restraints of unfavorable factors well. After researching the relative theory of developing simulation based on HLA, framework of communication simulation training system is established and then Communication equipment federation members is designed. Development of object class and interaction class is to determinate data flow and control flow of federation members substantively. This simulation system uses pRTI 1516 as the bottom support system of federation execution for HLA interface development. The researcher will be economizing on manpower and material resources. Furthermore, it will play a significant role in normalizing operation procedure and upgrading technology level.

Keywords: Simulation, HLA, Communication Train System.

1 Introduction

There are many kinds and huge amount of communication equipments using in various cases. Operators are required to be skilled. But, it is very inconvenient to train in the working equipments. Especially in the process of communication training organized, it needs diversified equipments cooperating very harmoniously. Many constraints should be subject to for system training, such as personnel assigned, condition of equipment used, natural conditions and so on.

So, it is necessary to design a communication simulative training system for training. It can simulate work principle and work process of real equipment via computer. Simulation training of communication equipment can remove restraints of unfavorable factors well. It will be economize on manpower and material resources. Furthermore, it will play a significant role in normalizing operation procedure and upgrading technology level.

2 Technology Basis

High Level Architecture (HLA) [1] is a reusable software frame for building distributed simulation components. The composition of federation is decided based on oriented

analysis and design. HLA is suit to build complicated simulation system for its reusability and interoperability.

Object oriented method is adopted as the basic ideology of HLA. Object model implementing simulation system is developed to acquire high-level interoperability and reuse. In the HLA-based simulation system, federation is defined for the distributed simulation system which achieves a specific simulating goal. It is constituted by some interactional members which constituted by many interactional objects. Object is the basic element of federation.

HLA Rules, interface specification and Object Model Template (OMT) [2] are the main components of HLA. Interface specification is the key. It has defined standard services supporting interoperability of federation members during simulation system running. All these services can be classified six species, i.e. federation management services, declaration management services, object management services, time management services, ownership management services and data distribution management services. They reflect the functions must be realized for interoperability indeed.

Run-Time Infrastructure (RTI) [3] is software implementing of HLA interface specification. It is a software system developed according HLA interface specification. Universal and relatively independent supporting services are provided for simulating applications. The function is analogical to Distributed Operating System.

3 Framework Design

Communication simulation training system needs to achieve information exchange among training plats. Training and examination is included not only for the single equipment, but also union of multiple communicating elements. During training process, all steps and parameters should be recorded to form user operation datum. Then, evaluation and analysis can be done further.

Considering the design requirement of communication simulation training system, the framework is built Combining HLA features. It is shown in figure 1.

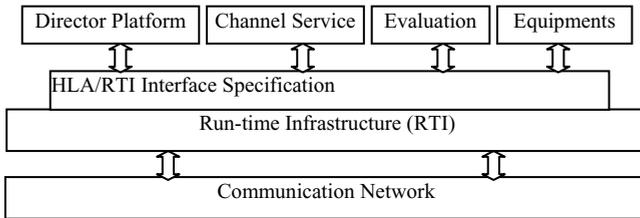


Fig. 1. Framework of the System

1 Director Platform

Director platform provides simulation datum to the simulation system according the scenario. It is the foundation for the system running. Functions including scenario input, edit and file is included.



Federation management is also an important function of the director platform. The whole process of federation implementation is managed and controlled. In the simulation seedtime, management mainly includes correlative condition selected and set for simulation task. During simulation running, it will supervise every step, involving control of advancing states (start, pause and end), surveillance of federation and each member's states, and record of event/ operation/fault.

2 Channel Service

Channel service federation uses logical channel model to simulate the real communication channel. All signals between receiver and transmitter should be processed and transferred via channel service. Thereby, the simulation system will be more consistent with real communication systems.

3 Evaluation and Analysis

Evaluation and analysis component collects and files simulation datum and simulation event. Based upon that, scoring and replay after simulation will be accomplished. Effectiveness analysis and evaluation of the system can be processed further. We can adopt database or file storing record.

4 Communication Equipments

Communication equipment members of federation simulate multifarious functions of Communication equipments. In communication process, user terminal module achieves communication datum connecting. Transmitting module will transform the signals from terminal for channel transferring. Receiving module will transform the received signals form channel inversely. Parameter of transmitting and receiving module can be configured by device control module. Integrated control module achieves device scheduling and control. Working states are monitored in real time.

Every federation members are interconnected as LAN. But, they will not exchange datum directly. Datum exchanging is carried through by interface functions such as updating, reflecting, transferring, receiving and etc. provided by RTI.

4 Design of Communication Equipment

Communication equipments have the similar structure and functions. So, we only need to code one federation for corresponding federation member of communication equipments. During process of simulating, it will incorporate federation execution with different identity representing correlate module of different equipments. It can simplify federation development, and achieve communication among multi-equipments conveniently.

Following HLA, five federation members including terminal, integrated controller, device controller, transmitter and receiver is defined for communication equipment. The interaction of these members is shown in figure 2.

1 Terminal

As a rule, communicating process initiate with terminal federation. After start of simulation, operators should manipulate on the training plat of terminal federation. The message needed to transmit is input to the simulating machine. For the simulation of real device, terminal federation should not only implement the functions simulated, but also provide friendly man-machine interface.

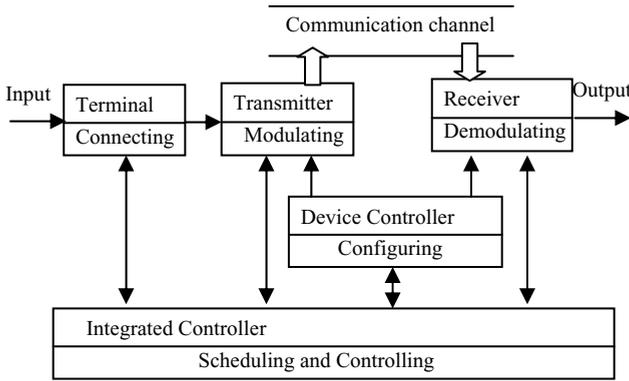


Fig. 2. Interaction of Communication Equipment Members

2 Integrated Controller

Integrated controller federation sends simulation parameters to device controller federation, transmitter federation, receiver federation and terminal federation when simulation initializing. After start of simulation, it will watch running states of other members. Message from others should be parsed and directive will be send to others by integrated controller federation.

3 Device Controller

After simulation beginning, device controller federation should handle command and message from integrated controller federation. After packet analyzer, it will send controller command to transmitter or receiver.

4 Transmitter

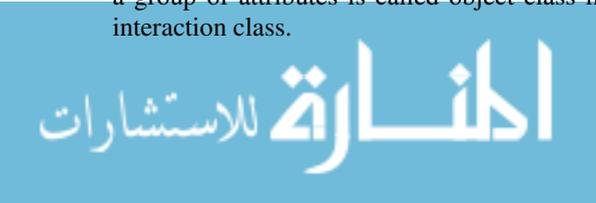
Transmitter federation receives working parameters and communication data from integrated controller federation after simulation beginning. It will parse the packet, and then process the signal. Packet received can be decided its type by the content. If it is working parameter or command send by integrated controller federation, simulating control panel will communication data change according to the direction. Otherwise, that is communication data. In this case, channel signal created module will processes signal by modulating and transforming, and then send to channel service federation.

5 Receiver

Receiver federation receives control packet from device controller federation after simulation beginning. It will make real time shift in working parameters, channel parameters and etc. basing the control packet. When packet from channel service federation is received, receiver federation will process signal by demodulating and transforming, and then send to device controller federation.

5 Simulation Implementation

In the simulation world of HLA, attribute and parameter is the basic cell for communicating within federation and among federation members. Set composing with a group of attributes is called object class in HLA, and Set of parameters is named interaction class.



Design of object class and interaction class is to determinate data flow and control flow of federation members substantively. Each member can publish and order object classes and/or interaction classes to obtain information needed. It can also send message others want at the same time.

1 Object Class

Object class has a certain life cycle. Attributes information of its example can be saved in RTI service. As long as they order the object class, each member joining in federation in midway can always receive s attributes value of all examples of this class.

According to the function analysis of federation members, each simulation device is defined as object class. Corresponding member of federation will publish these classes as shown in figure 3.

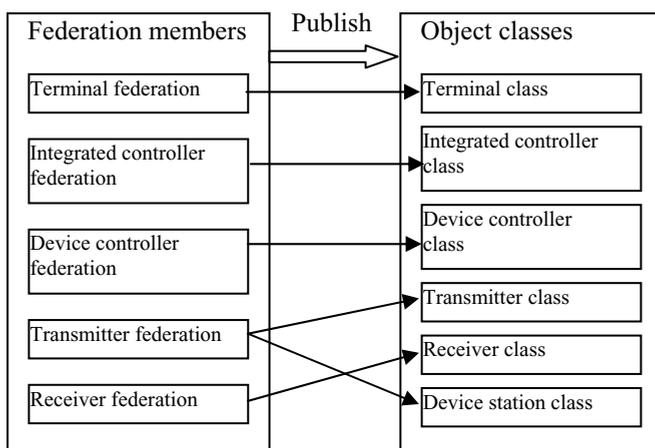


Fig. 3. Object classes and corresponding members of federation

Terminal Class is published by terminal federation. During simulation communication, terminal federation can register many terminal classes to simulate more than one terminal of real devices. Terminal federation takes on responsibilities for updating attribute value of terminal class. So that federation member ordering this class can acquire latest status of simulation example. Only one attribute, using state, is included to express the now using state (busy/ leisure) of terminal federation example.

Integrated Controller Class is published by integrated controller federation. The only one attribute is work state. During simulation, integrated controller federation will update the attribute value. Director federation can order this object class to acquire work state of simulation equipment.

Device Controller Class is published by device controller federation. The attributes include device ID, using state and work mode. Every device controller is corresponding to a receiver and/or a transmitter. Device ID is used to identify this relation. During simulation, device controller registers many examples of this object class to simulate the group of device controller. With same device ID, device controller example is corresponding to a certain receiver example and/or a certain transmitter example.

Transmitter Class is published by transmitter federation. The attributes include device ID, work state, work mode, communication mode, channel No., working frequency and power grade. In simulation communication process, attribute of work mode, communication mode and power grade represent modulation mode, communication mode and power grade of working transmitter separately.

Receiver Class is published by receiver federation. The attributes include device ID, work state, work mode, work bandwidth, channel No. and working frequency. Attribute of work mode expresses modulation mode of receiver in simulation communication process.

Device Station Class is published by transmitter federation and receiver federation. The attributes include longitude and latitude.

2 Interaction Class

According to the signal flowing of communication simulation training system, interaction classes are shown as follows.

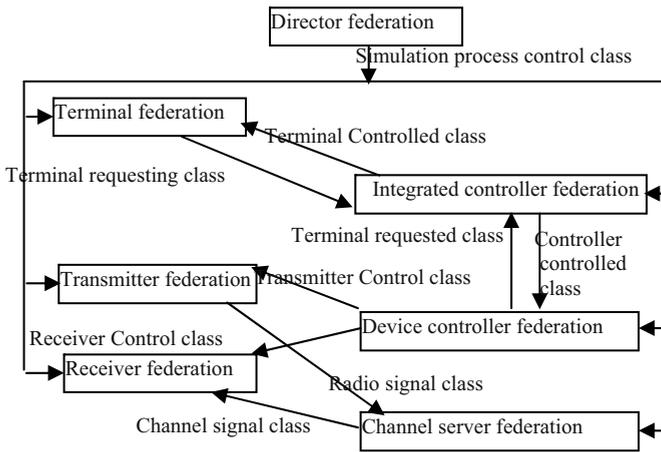


Fig. 4. Interaction classes and corresponding members of federation

In HLA, interaction is a specific action bring by a certain object. It can influence objects of other members. Interaction at real time only occurrences in a certain time, but not last longer than. The member joining in federation in midway can only receive information of this class sent afterward.

Simulation Process Control Class is sent by director federation and received by others. To control simulation advancing is the function of this class. There is one parameter, state, to express the simulation state. The value can be 0(pause), 1(keep on) and 2(end).

Terminal Requesting Class is sent by terminal federation and received by integrated controller federation. This class is used to ask for transmitting to integrated controller by terminal. There are two parameters, telegraph No. and signal type. Telegraph No. identifies terminal uniquely. And signal type identifies operation type of the packet.



Terminal Controlled Class is sent by integrated controller federation and received by terminal federation. It is used to control work state of terminal. There is only one parameter, work state.

Controller Controlled Class is sent by integrated controller federation and received by device controller federation. There are only two parameters, self address and object address. Self address presents address of transmitter and object address presents address of receiver.

Transmitter Control Class is sent by device controller federation and received by transmitter federation. The parameters are parallel with transmitter parameters. When link is creating, this class is used to set transmitter parameters.

Receiver Control Class is sent by device controller federation and received by receiver federation. When receiver is scanning, this class is used to control receiver receiving at a certain channel.

Radio Signal Class is sent by transmitter federation and received by channel server federation. The parameters are the parameters of transmitter in using.

Channel Signal Class is sent by channel server federation and received by receiver federation. Channel server will process the signal from receiver. So it can simulate the real communication channel. Signal processed will send to receiver as the channel signal.

Terminal Requested Class is sent by device controller federation and received by integrated controller federation. When device controller has received packet from receiver, it asks integrated controller for leisure terminal.

3 HLA Interface

This simulation system uses pRTI 1516[4] as the bottom support system of federation execution for HLA interface development. When pRTI 1516 is running, a configuration file, Federation Execution Data (FED), is required.

FED is the result of federate object model development. It is a protocol managed to achieve interoperating of all federation members. It records object classes and interaction classes entering federation interaction during simulation executing. Their attributes, parameters and relative route space information are also filed. In addition, management object model Pre-defined by HLA and other details of federation execution is also recorded.

When simulation is running, RTI will create corresponding federation execution according to the detail datum provided by FED file. In the whole life cycle of federation execution, interaction of federation members is coordinated basing on FED.

In the development of system federation, we can use Digital Weapon Kit (DWK) modeling tool to create, edit and modify FED file according to the design and definitions of those object classes, interaction classes and data types finished at an earlier time. pRTI 1516 provides 143 API functions to realized HLA interface specifications, including six species management services and supporting services. The developer can selective program basing on their needs to realize corresponding service function.

6 Conclusion

HLA as a late development of advanced distributed simulation is used to develop the common simulation technology framework. Interoperability among different simulation application and reuse of simulation components will be supported. Run-Time Infrastructure is used to provided Universal and independent support services. Simulation application level is isolated from low level of supporting environment. Implementation details are concealed from each other. Thereby, each component can be developed independently. So, Up-to-date technology in their respective fields can be utilized to achieve normative functions and service.

After researching the relative theory of developing simulation based on HLA, framework of communication simulation training system is established. According to structure of communication system, the composition of federation is confirmed. In addition, federation object model is developed. The design of system has the characteristics of well expandability and customization. Adopting the idea of modularization design, components in the simulation system is implemented on the base of their functions. Module reuse capability invested by HLA is possess as well as flexible expandability. Coding for special operating act can provide services to simulation tasks of varies types and goals [5]. It is propitious to reduce development difficulty, shorten development time, and satisfy requirement of varies operation training.

References

1. Fu, Z., Wang, Y.: HLA Technique in Computer Simulation. National Defence Expre, Beijing (2003)
2. IEEE Std1516.2-2000, "IEEE Standard for Modeling and Simulation(M&S) High Level Architecture (HLA)-Object Model Template(OMT)". The Institute of Electrical and Electronics Engineers, USA (2000)
3. Moreau, M.: Documentation for RTI, pp. 32–38. George Mason University (1997)
4. Daniel, J., Calvin, O.: Data Distribution Management in RTI1516. In: Spring Simulation Interoperability Workshop, Canada, pp. 12–16 (2002)
5. Fujimoto, R.M.: Performance Measurements of Distributed Simulation Strategies. Transactions of the Society for Computer Simulation 2, 89–132 (2000)

Numerical Simulation on Heat Transfer from Envelope of the Underground Engineering

Wang Li-jun, Miao Xiao-ping, Wang Rui-hai, and Wang Bo

Engineering Institute of Corps of Engineers, PLA Univ. of Sci. & Tech. Nanjing, China
dsfnhvac@gmail.com, mxp@163.com, ruihaiwang@126.com,
Wangbo83331@sohu.com

Abstract. Whether the results of the dynamic heat flux from the underground engineering envelope are accurate, may influence the accuracy of calculating the transient heat load and could affect the initial cost and actual operation of the air conditioning system in the underground engineering. Based on the mathematical modeling of heat transfer in the underground engineering envelope, the influence of the model dimension, boundary condition on the soil surface, initial temperature of the soil, the adiabatic distances far from the envelope and the heat transfer coefficient between the envelope surface and the indoor air, the heat transfer mechanism of the underground engineering envelope was studied in terms of the building structure, style of the envelopes and the difference of the locations. For providing the analysis basis to simplified calculation of heat transfer in the underground engineering envelope.

Keywords: underground engineering, envelope, heat transfer, numerical simulation.

1 Introduction

The heat transfer between the underground engineering envelope and the surrounding soils is clear different from the ground buildings. In summer, the wall surface temperature gradually increased. Heat flow transfer to the surrounding soils, In winter, the envelope surface transfer heat to the cold air when they are in contact with each other. This makes the calculation of heat transfer more complicated and important. It is essential to analyze the underground envelope heat transfer rule in order to optimize the design of energy-saving and the operation management of the air conditioning dehumidification system.

2 Mathematical Model

2.1 Control Equation

In order to explore the process of heat transfer of underground envelope, on the basis of some assumptions, the simplified physical model is shown as Figure 1.

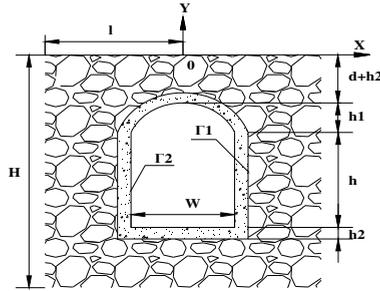


Fig. 1. Simplified physical model of the underground engineering and the surrounding soils

2.2 Engineering Materials and Surrounding Soils

Only the heat conduction equation in these areas:

$$\rho C_p \frac{\partial T}{\partial \tau} = \nabla(\lambda \nabla T) \tag{1}$$

Where ρ is density, kg/m^3 ; C_p is specific heat at constant pressure, $\text{J}/(\text{kg}\cdot\text{K})$; T is temperature, K ; τ is time, s ; λ is thermal conductivity, $\text{W}/(\text{m}\cdot\text{K})$; ∇ is Hamiltonian operator.

2.3 The Air Area of the Off-Wall Lined Interspace

The air has flow and convection heat transfer in the region simultaneously, because its involuntary convection in a confined space, the control equations are Navier-Stokes equation and energy conservation equation[1].

$$\begin{cases} \frac{\partial \rho}{\partial \tau} + \nabla(\rho \mu) = 0 \\ \rho \frac{\partial \mu}{\partial \tau} + \rho \mu \nabla \mu = \nabla \left\{ -\rho I + \eta(\nabla \mu + (\nabla \mu)^T) - \frac{2}{3} \eta(\nabla \mu) I \right\} + F \\ \rho C_p \frac{\partial T}{\partial \tau} = \nabla(-\lambda \nabla T) - \rho C_p(\mu \nabla T) \end{cases} \tag{2}$$

Where μ is velocity vector, m/s ; P is pressure, Pa ; I is unit vector; F is volume force, N/m^3 ; η is dynamic viscosity, $\text{kg}/(\text{m}\cdot\text{s})$.

2.4 Soil Surface Boundary Conditions

If the soil surface temperature changes are known, for the first boundary condition [2]

$$T|_{Y=0} = T_{sur}(\tau) \tag{3}$$

Where $T_{sur}(\tau)$ is soil surface temperature, K .

If we consider the heat balance of the soil surface, the convection heat transfer between soil and the outdoor air, solar radiation to soil and long wave radiation of the soil to the sky should be taken into account.



$$-\lambda \nabla T |_{Y=0} = h_{h,0} \{T_0(\tau) - T|_{Y=0}\} + \alpha I(\tau) - \varepsilon \sigma \{T_{Y=0}^4 - T_{sky}(\tau)^4\} \quad (4)$$

Where $h_{h,0}$ is heat exchange coefficient between the soil surface and surrounding air, $W/(m^2 \cdot K)$; $T_0(\tau)$ is outdoor air temperature, K ; α is the absorptivity of the soil surface to solar radiation; $I(\tau)$ is the horizontal overall radiation intensity, W/m^2 ; ε is the emissivity of the soil to the sky radiation; σ is Stefan-Boltzmann constant, the value is $5.6686 \times 10^{-8} / (m^2 \cdot K^4)$; $T_{sky}(\tau)$ is the sky effective temperature value, K .

3 Project Style

3.1 Partially Underground Buildings

The engineering structures which have other structures in the ground are termed partially underground buildings, the simplified diagram is shown in Figure 2(a).

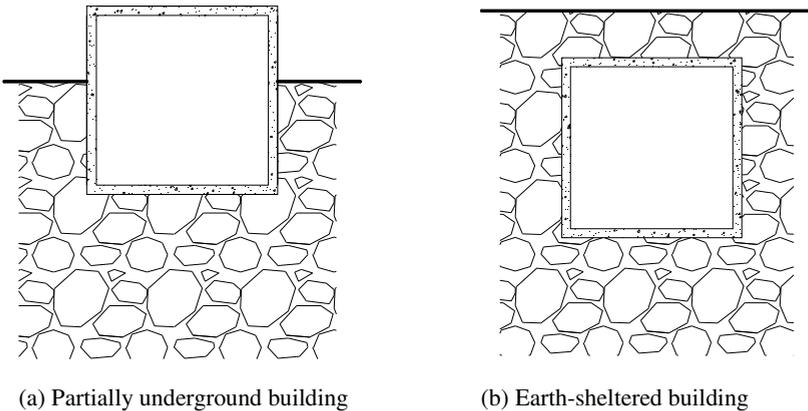


Fig. 2. Schematic of typical underground engineering

Study to the underground parts of the partially underground adherent buildings in Nanjing. The underground engineering envelope wide is 6m, the buried underground depth is 5.3m, the thickness of the envelope is 0.5m, the length is 12m, the engineering around is light clay, the material of the envelope is reinforced concrete. Thermal properties are listed in Table 1.

Table 1. Thermophysical parameters of the building materials and the soils

| material | $\rho / (kg \cdot m^{-3})$ | $c_p / (J \cdot kg^{-1} \cdot K^{-1})$ | $\lambda / (W \cdot m^{-1} \cdot K^{-1})$ | $a / (m^2 \cdot h^{-1})$ |
|---------------------|----------------------------|--|---|--------------------------|
| reinforced concrete | 2500 | 0.92×103 | 1.74 | 0.0027 |
| light clay | 1200 | 1.01×103 | 0.47 | 0.0014 |

Boundary condition on the soil surface is the third kind, the value of the soil surface absorptivity to the sun radiation is 0.5[3]. The emissivity to the sky long-wave radiation is 0.9[4], Outdoor hourly meteorological parameters use the date that are provided by "appropriate meteorological data sets of thermal environment analysis for China building"; The adiabatic distance away from the envelope is 15m; Refer to the information about Nanjing area, the depth and temperature of the soil constant temperature is 20m and 15.4°C respectively. Indoor air temperature is 26°C, the heat transfer coefficient of the envelope surface is 8.14W/(m²·K)[5], and the soil surface heat transfer coefficient is 15W/(m²·K)[6].

COMSOL Multiphysics is a large, advanced numerical simulation software. Widely used in various fields of scientific research and engineering calculations, that is called "the first truly any direct coupling multi-physics analysis software". The three-dimensional model can be reduced to two-dimensional model if the aspect ratio of the underground engineering envelope is greater than 1:2 and the accuracy desire is less than 15% when studying the heat transfer of the envelope[7]. Practical engineering, the aspect ratio is larger than 1:2 due to the construction technology factors, considering the time-consuming of the numerical simulation and heat transfer of the underground engineering envelope is only a part of the load of the engineering, so this paper adopt the two-dimensional model to research the heat transfer. The step time is 3600s[8].

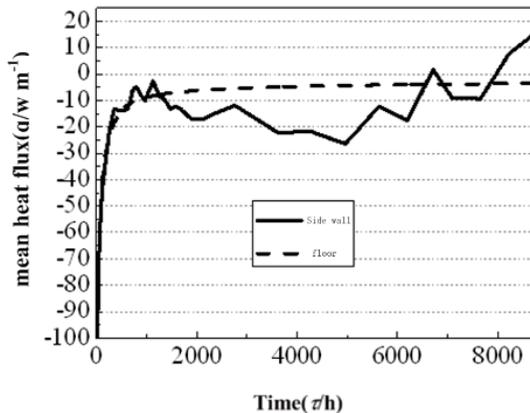


Fig. 3. Heat load per unit of length for the attached underground engineering envelope

Figure 3 and Figure 4 show the heat load per unit of length for the attached underground engineering envelope and temperature profile of the attached underground engineering envelope and the surrounding soils respectively. In Figure 3, the negative value indicates that the envelope heat transfer to the soil, and the positive is reverse. It can be seen from the figure that the wave of heat transfer load of the side walls present roughly triangular function. The heat load per unit length is larger in the beginning more than 1,000 hours, which is to maintain the room temperature conditions

and transfer heat to the cold surrounding soils. In Figure 4, the heat flow path is not all the arc shape between the attached underground engineering and surrounding soils, this rule is only showed in the side walls which are close to the surrounding soils.

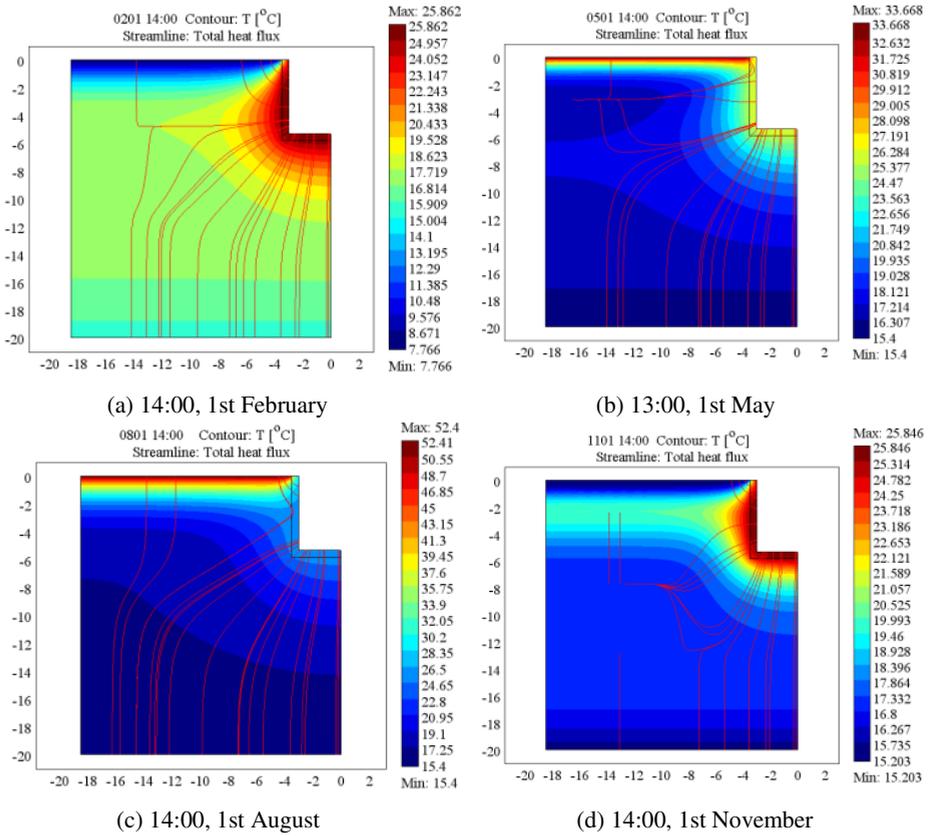


Fig. 4. Temperature profile of the attached underground engineering envelope and the surrounding soils

3.2 Earth-Sheltered Building

Earth-sheltered building is a project which all structures are below ground entirely, such as Figure2 (b). It can be seen from Figure 5 that the annual variation of the mean heat flux through the single underground engineering envelope with different embedded depths. The Figure 5 shows that the envelope engineering are influenced dramatically by the change of outdoor meteorological parameters that the embedding depths are within 2m. To 10m depth, the annual variation of the mean heat flux has been very small, the meteorological parameters may not take into account.



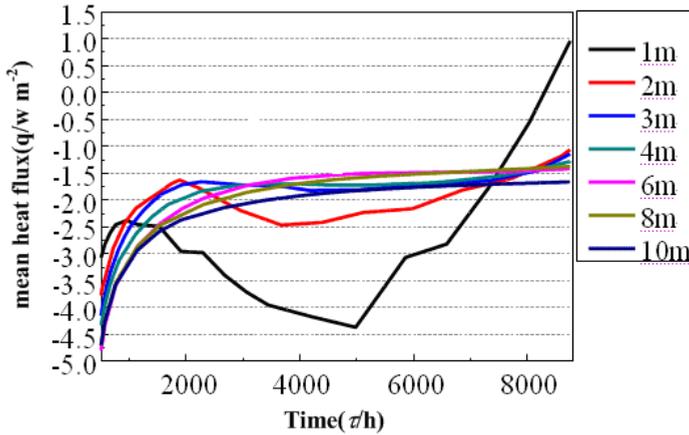


Fig. 5. Annual variation of the mean heat flux through the single underground engineering envelope with different embedded depths

4 Envelope Style

For the adherent and off-wall lined envelopes which filled in stonewalling, due to their similar architectural material thermophysical properties, the lesser resistance, which can be simplified as monolayer adherent. For the off-wall lined envelopes, the interior air sandwich play a remarkable influence in despite of its closure space and almost no flow. But taking into account the insulation performance of the air, the heat flux must show different changes in contrast with the monolayer adherent.

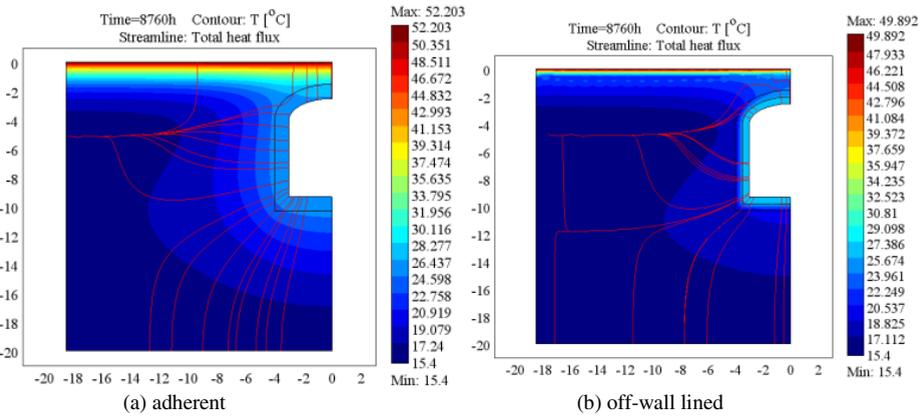


Fig. 6. Temperature profile of the adherent and off-wall lined envelope in the underground engineering (1 year)

Study to a underground engineering in Nanjing, its wide is 6m, 5.3m height, the vault height is 1.5m, the embedded height is 1m. The engineer around is the light clay, the material of the envelope is reinforced concrete. Thermal properties are listed in Table 1.

The conditions of the numerical simulation are consonant with mention above. The thermal parameters of obturation air use the data that are provided by COMSOL software library, its density is the ideal gas density influenced by temperature in one atmospheric pressure, specific heat at constant pressure and thermal conductivity are the experience function of the temperature.

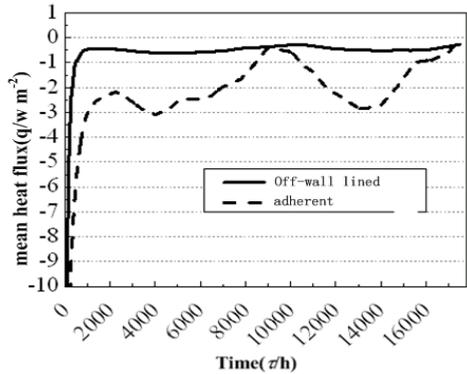


Fig. 7. Mean heat flux variations of the adherent and off-wall lined envelope in the underground engineering (2 years)

Mean heat flux variations of the adherent and off-wall lined envelope in the underground engineering are shown as Figure 6 and Figure 7. It can be seen from the figures, the air sandwich plays an important role in heat preservation, and the heat flux is decreased by approximately one third.

5 Different Region

China is divided into five regions according to "Thermal Design Criterion of Civil Architecture", the chilliness, the cold, the hot summer, the cold winter and the mild and the hot summer and warm winter. In order to investigate the thermal load of the five regions of underground engineering envelope, Harbin, Beijing, Nanjing, Kunming and Guangzhou are selected to analyze. Analyzing the same underground engineering and the surrounding soils to highlight the heat load influence of the underground engineering envelopes in different regions. For constant temperature depth in different areas, even if the surface temperature of the soil is in the greatest volatility in Harbin (21.35°C), the 20m constant temperature depth may meet all the conditions.

The mean heat flux changes of single layer adherent underground engineering envelope in five different Chinese cities in 2 years are shown in Figure 8. It can be seen from the figure, as Harbin in the chilliness regions, larger heat transfer to the surrounding soil and the heat load is larger than other cities. For Guangzhou, the heat flux will transfer to the envelope in summer. A striking dissimilarity in heat flux between the Earth-sheltered underground engineering in different regions.



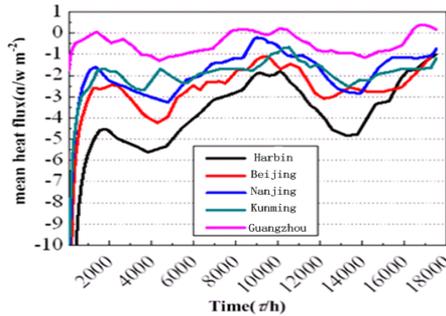


Fig. 8. Mean heat flux variation of single layer adherent underground engineering envelope in five different Chinese cities (2 years)

6 Conclusion

The heat transfer between the underground engineering envelope and the surrounding soils is clear

different from the ground buildings. According to this article, the conclusions are as follows:

(1)The envelope engineering is influenced dramatically by the change of outdoor meteorological parameters that the embedding depths are within 2m. To 10m depth, the annual variation of the mean heat flux has been very small, the meteorological parameters may not take into account.

(2)The air sandwich in the off-wall lined envelope in the underground engineering plays an important role in the heat preservation, and the heat flux is decreased by approximately one third.

(3)A striking dissimilarity in heat flux between the Earth-sheltered underground engineering in different regions.

References

1. Van, B.G., Hoogendoorn, C.J.: Submitted to *Solar Energy* 30(4) (1983)
2. Lloyd, R.M.: TCardiff: submitted to University of Wales. College of Cardiff (1994)
3. Dos Santos, G.H., Mendes, N.: 8th International IBPSA Conference. Submitted to International Building Performance Simulation Association, Eindhoven (2003)
4. Zhong, Z., Braun, J.E.: Submitted to *Building and Environment* 42(3) (2007)
5. Beausoleil-Morrison, I., Mitalas, G., McLarnon, C.: BASECALC(TM): Submitted to International Building Performance Simulation Association (1995)
6. Yuan, Y.P., Cheng, B.Y., Mao, J.F., et al.: Submitted to *Building and Environment* 41(3) (2006)
7. Abdelbaki, A., Amjad, S., Zrikem, Z.: Submitted to Proceedings of the 3rd Renewable Energy Congress, Reading, UK (1994)
8. Yuan, Y.P., Mao, J.F., Du, Y.X., et al.: Submitted to *Building and Environment* 43(6) (2008)

A Reversible Watermarking Scheme for 2D Vector Maps

Dan Wu

Department of Mathematics, China Jiliang University, Hangzhou, 310018, P.R. China
linlin_52@yahoo.com.cn

Abstract. Reversible watermarking is suitable for hiding data in 2D vector maps, because it can remove the embedding-induced distortions after extracting the hidden bits. This paper proposes a novel reversible watermarking scheme based on the difference shifting. Since the vector map includes a sequence of vertices, we embed the watermark into the vertex coordinates by modifying the differences between the adjacent vertex coordinates. The scheme need not record the extra information. The experimental results show that the proposed scheme achieves good performance of the imperceptibility and is robust against some attacks, such as slightly adding noise, translating, slight scaling and rotating.

Keywords: digital watermark, reversible watermarking, difference expansion, difference shifting, histogram shifting, robust.

1 Introduction

In the past decades, digital watermarking has been widely used in areas such as ownership protection, content authentication, distribution tracking, and broadcast monitoring. However, in some applications, especially in the medical, military and legal domain, it is desired to reverse the marked media back to the original media without any distortions after the watermark is retrieved. Under the circumstance, reversible watermarking is desired, which not only extracts the watermark, but also perfectly reconstructs the original host signal from the watermarked work. Recently many reversible watermarking algorithms have been proposed, such as those reported in [1-6]. Tian [7] proposed a high capacity reversible data embedding algorithm, which was based on difference expansion, and the method had been extended by [8-10]. Ni [11] proposed an algorithm based on histogram shifting, which utilized the zero or the minimum points of the histogram of an image and slightly modified the pixel grayscale values to embed the data into the image. Thodi [12] proposed a reversible watermarking technique called prediction error expansion, which better exploited the correlation inherent in the neighborhood of a pixel than the difference expansion scheme. However, most previous works focused on the image reversible watermark. To our knowledge, few works have focused on the 2D vector maps. M. Voigt et al. [13] first proposed the method of reversibly hiding data in vector maps. They modified the integer discrete cosine transform (DCT) coefficients of the map coordinates to hide the data. Most previous works focused on enhancing the robustness [14-15]. Wang [16] proposed two reversible data hiding schemes based on the idea of difference expansion. The first scheme took the coordinates of the vertices as the cover data and hid the data

by modifying the differences between the adjacent coordinates, and the second scheme adopted the Manhattan distance between neighbor vertices as the cover data. However both schemes can only resist against the slight noise attack.

In this paper, we apply difference shifting to embed the watermark in 2D vector maps. According to the characteristics of 2D vector maps data, we use the correlation of map coordinates to embed the watermark. The watermarks are embedded by changing the coordinate's differences between the adjacent vertices of the map. The scheme is reversible and is robust against some attacks, such as adding slight noise, translating, slightly scaling and rotating.

The structure of this paper is given as follows: In Section 2, we analyze the idea of difference shifting. A detailed description of our algorithm is provided in Section 3. Experimental results are presented in Section 4, and Section 5 concludes.

2 The Idea of Difference Shifting

There are two methods used to reversibly hide data in images: difference expansion [7] and histogram shifting [11]. Now we demonstrate the idea of difference shifting as follows.

If x_1 and x_2 are the gray values of a pixel-pair, then the integer-mean m and the difference d are defined as

$$\begin{cases} m = \text{floor}((x_1 + x_2)/2) \\ d = x_1 - x_2 \end{cases} \quad (1)$$

Since this transformation is invertible, the gray levels x_1 and x_2 can be given by

$$\begin{cases} x_1 = m + \text{floor}((d+1)/2) \\ x_2 = m - \text{floor}(d/2) \end{cases} \quad (2)$$

Using the difference d , we can hide the bit b via the following equations

$$\begin{cases} d' = d + K + 1 & |d| < K \quad \text{and} \quad b = 1 \\ d' = d - K - 1 & |d| < K \quad \text{and} \quad b = 0 \\ d' = d + \sigma & d \geq K \\ d' = d - \sigma & d \leq -K \end{cases} \quad (3)$$

where K and σ is the threshold controlled by the user ($\sigma > 2K$).

The watermarked pixel x_1' and x_2' can be calculated by d' and m via (2). When $d > K$ or $d < -K$, we shift the difference d further away from the zero point, which is called the difference shifting, and leave $(2K, K + \sigma)$ and $(-K - \sigma, -2K)$ empty.

3 Watermark Embedding and Extracting

With the case of 2D vector maps, a map is generally composed by large amounts of vertices which are represented by 2D coordinates. The vertices of many maps have high densities, and the positions of two adjacent vertices are usually very close, which can be used to embed the watermark.

Normally, the vertex coordinates are floating point numbers. In order to perform difference expansion and shifting, all coordinates should be first transformed to integers. The integer coordinates (x, y) can be defined as follows:

$$(x, y) = \text{floor}((x_l, y_l) \times 10^p), \quad p \leq p_{\max}. \quad (4)$$

where (x_l, y_l) is the original coordinates with floating point numbers and p_{\max} is the maximum number of digits after the decimal point [16].

Watermark Embedding. The embedding process hides a binary message, which includes the hash of the original map for data authentication. Suppose the vector map is composed by vertices $\{v_1, v_2, \dots\}$, we divide it into vertex pairs $\{(v_1, v_2), (v_3, v_4), \dots, (v_{2i-1}, v_{2i}), \dots\}$, where $v_{2i-1} = \{x_{2i-1}, y_{2i-1}\}$ is the $(2i-1)$ th vertex and $\{x_{2i-1}, y_{2i-1}\}$ is the integer coordinates of v_{2i-1} . For the pair (v_{2i-1}, v_{2i}) , the difference and the integer-mean of two vertices are calculated for x and y respectively

$$\begin{cases} dx = x_{2i-1} - x_{2i} \\ mx = \text{floor}\left(\frac{x_{2i-1} + x_{2i}}{2}\right) \end{cases} \quad (5)$$

$$\begin{cases} dy = y_{2i-1} - y_{2i} \\ my = \text{floor}\left(\frac{y_{2i-1} + y_{2i}}{2}\right) \end{cases} \quad (6)$$

We just take x coordinates as an example since the situation is exactly the same as that in y coordinates. According to the difference, we can decide whether or not to embed information bit b via the equations

$$\begin{cases} dx' = dx + K + 1 & |dx| < K \quad \text{and} \quad b = 1 \\ dx' = dx - K - 1 & |dx| < K \quad \text{and} \quad b = 0 \\ dx' = dx + \sigma & dx \geq K \\ dx' = dx - \sigma & dx \leq -K \end{cases} \quad (7)$$

Then the watermarked coordinates x_{2i-1}' and x_{2i}' can be calculated as follows

$$\begin{cases} x_{2i-1}' = mx + \text{floor}((dx'+1)/2) \\ x_{2i}' = mx - \text{floor}(dx'/2) \end{cases} \quad (8)$$

Watermark Extracting and Original Data Recovering. Given the watermarked vector map, we can get the same sequence of vertices and divide it into vertex pairs as in embedding. Then we extract the watermark and recover the original data in x coordinates as an example. The situation is exactly the same as that in y coordinates.

We denote the sequence of vertices as $\{(v_1', v_2'), (v_3', v_4'), \dots, (v_{2i-1}', v_{2i}'), \dots\}$, and calculate the difference dx' and the integer-mean mx of the pair (v_{2i-1}', v_{2i}') . When $|dx'| \leq (3K + \sigma)/2$, the watermark can be obtained via the following equations

$$\begin{cases} b = 1 & dx \geq 0 \\ b = 0 & dx < 0 \end{cases} \quad (9)$$

In order to recover the original vertex it is necessary to obtain the original difference dx via the following equations

$$\begin{cases} dx = dx' - K - 1 & 0 \leq dx' \leq (3K + \sigma)/2 \\ dx = dx' + K + 1 & -(3K + \sigma)/2 \leq dx' < 0 \\ dx = dx' - \sigma & dx' \geq K + \sigma \\ dx = dx' + \sigma & dx' \leq -K - \sigma \end{cases} \quad (10)$$

Then the original x coordinates can be exactly recovered according to (2), using the restored difference dx and the integer-mean mx . We can extract the watermark and recover the pair (v_{2i-1}, v_{2i}) . We do the same procedure to the other pairs, extract the watermark, and recover the vertex pair.

4 Experimental Results

We have applied our proposed algorithm to a contour map. In the experiments, we can extract the correct watermarks from the watermarked vector map and exactly recover the original map. Comparing the MD5 hash of the original map to the recovered map we find that the two hashes are exactly matched if the contour map has not been attacked. Table 1 is the experimental results of our scheme which include the payload, and the root mean square error (RMSE) of the map. The results show that the payload increases, and RMSE increases with the threshold (K) increasing. Table 2 lists the attacks and the error rates (we chose $p=6$, $K=100$, and $\sigma=110$ in the experiment). According to Table 2, it can be seen that our algorithm is robust against some attacks, such as adding slight noise, translating, slightly scaling and rotating.

Table 1. Payload, and RMSE of different models (Our proposed algorithm)

| | p | K | σ | Payload | RMSE |
|-------------|-----|-----|----------|---------|--------|
| Contour map | 6 | 10 | 30 | 537 | 0.0013 |
| | 6 | 50 | 110 | 1178 | 0.0075 |
| | 6 | 100 | 210 | 2361 | 0.0136 |
| | 6 | 150 | 310 | 3473 | 0.0191 |
| | 6 | 200 | 410 | 4572 | 0.0237 |
| | 6 | 250 | 510 | 5568 | 0.0277 |
| | 6 | 300 | 610 | 6435 | 0.035 |

Table 2. The error rate after different attacks

| Attack | Error rate |
|------------------------------|------------|
| No attacks | 0 |
| expanding 0.25 | 0 |
| Shrinking 0.2 | 0 |
| Tanslating | 0 |
| Adding Gauss noise(slightly) | 1.4% |
| Counterclockwise rotation 3° | 6.99% |
| Counterclockwise rotation 4° | 8.01% |

The original contour map and the watermarked map are shown in Fig. 1 and Fig. 2. Note that no notably visible distortions exist in Fig. 2 compared with map in Fig. 1. In the experiments, we can extract the correct watermarks from the watermarked map and exactly recover the original contour map.

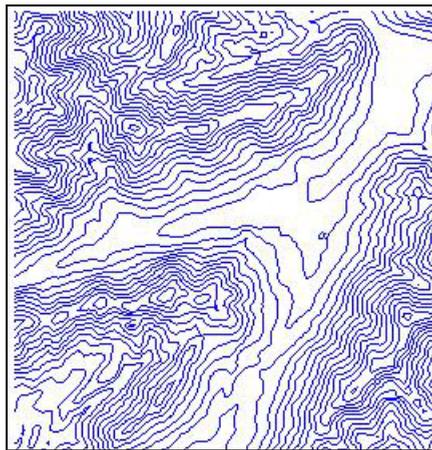
**Fig. 1.** The original contour map



Fig. 2. The watermarked contour map after embedding 2361bits

5 Conclusion

In this paper, we have proposed a reversible watermarking algorithm of 2D vector maps, which is based on the difference shifting. The difference between two vertices coordinates are used to embed the watermark, and the difference of every vertex pair fall into three parts: one is shifted right if the difference is bigger than K , another is shifted left if the difference is smaller than $-K$, and the third part is used to embed the data (K is a positive integer chose by the user). If the watermarked contour map has not been attacked the embedded watermark can be extracted from the watermarked map and the original map can be exactly recovered without any distortion. The distortion caused by the reversible watermark embedding is unnoticeable, and the experimental results show that our algorithm is robust against some attacks, such as adding slight noise, translating, slightly scaling and rotating.

Acknowledgments. The author would like to thank the Foundation of Zhejiang Educational Committee (No. Y200909370) for financial support.

References

1. Honsinger, C.W., Jones, P.W., Rabbani, M., Stoffel, J.C.: Lossless recovery of an original image containing embedded data. US Patent, Docket No: 77102 (1999)
2. Goljan, M., Fridrich, J.J., Du, R.: Distortion-Free Data Embedding for Images. In: Moskowitz, I.S. (ed.) IH 2001. LNCS, vol. 2137, pp. 27–41. Springer, Heidelberg (2001)
3. Fridrich, J.: Lossless data embedding—new paradigm in digital watermarking. EURASIP Journal on Applied Signal Processing, 185–196 (2002)
4. Celik, M.U., Sharma, G., Tekalp, A.M., Saber, E.: Reversible data hiding. In: Proceedings of International Conference on Image Processing, pp. 157–160 (2002)
5. Xuan, G., Zhu, J., Chen, J., Shi, Y.Q., Ni, Z.C., Su, W.: Distortionless data hiding based on integer wavelet transform. Electronics Letters 38(25), 1646–1648 (2003)

6. De Vleeschouwer, C., Delaigle, J.F., Macq, B.: Circular interpretation of bijective transformations in lossless watermarking for media asset management. *IEEE Transactions on Multimedia* 5(1), 97–105 (2003)
7. Tian, J.: Reversible data embedding using a difference expansion. *IEEE Transactions on Circuits and Systems for Video Technology* 13(8), 890–896 (2003)
8. Alattar, A.M.: Reversible watermark using the difference expansion of a generalized integer transform. *IEEE Transactions on Image Processing* 13(8), 1147–1156 (2004)
9. Chang, C.C., Lu, T.C.: A difference expansion oriented data hiding scheme for restoring the original host images. *The Journal of Systems & Software* 79(12), 1754–1766 (2006)
10. Cho, J., Prost, R., Jung, H.: An oblivious watermarking for 3-D polygonal meshes using distribution of vertex norms. *IEEE Transactions on Signal Processing* 55(1), 142–155 (2007)
11. Ni, Z., Shi, Y.Q., Ansari, N., Su, W.: Reversible data hiding. *IEEE Transactions on Circuits and Systems for Video Technology* 16(3), 354–362 (2006)
12. Thodi, D.M., Rodriguez, J.J.: Expansion embedding techniques for reversible watermarking. *IEEE Transactions on Image Processing* 16(3), 721–730 (2007)
13. Voigt, M., Yang, B., Busch, C.: Reversible watermarking of 2d-vector data. In: *Proc. ACM Int. Workshop on Multimedia and Security*, Magdeburg, Germany, pp. 160–165 (2004)
14. Gou, H., Wu, M.: Data hiding in curves with applications to map fingerprinting. *IEEE Trans. Signal Process.* 53(4), 3988–4005 (2005)
15. Gou, H., Wu, M.: Fingerprinting Curves. In: Cox, I., Kalker, T., Lee, H.-K. (eds.) *IWDW 2004. LNCS*, vol. 3304, pp. 13–28. Springer, Heidelberg (2005)
16. Wang, X.T., Shao, C.Y., Xu, X.G., Niu, X.M.: Reversible Data-Hiding Scheme for 2-D Vector Maps Based on Difference Expansion. *IEEE Transactions on Information Forensics and Security* 2(3 Part 1), 311–320 (2007)

Research of Mobile Public Transport Information System Based on J2ME and SVG

Zhang Suqi¹, Gu Junhua², Li Weini², and Gu Qiang²

¹ Tianjin University, Tianjin 300072, China
hebutiti@163.com

² Hebei University of Technology, Tianjin 300401, China
jhgu@hebut.edu.cn

Abstract. With the rapid development of wireless technology and the popularity of mobile devices, mobile GIS become the trend of the development of GIS. Mobile public transport information system is the powerful expansion of intelligent public transport system applied in WebGIS. The paper puts forward mobile public transport information system based on J2ME and SVG which effectively use the bus information of WebGIS. It applies the mobile GIS in public transport service that makes the life of people more convenient, and brings good social and economic benefits, improving the social information construction.

Keywords: J2ME, SVG, mobile GIS, GPS, path-selecting.

1 Introduction

In the sustainable development of the cities economic construction ,traffic problem increasingly, for example the road crowded by automobile caused, vehicle delay, the air pollution and noise pollution ,seriously influences the city environment and development. With the current situation analysis, to attract more people to choose the public transport for the main mode of transit, the government should make well the construction of public traffic information system, releasing timely the bus and road information, to obtain information for people and rational selection of travel paths [1].With the popularity of mobile devices, GPS, wireless communications technology continues to evolve, people access to information no longer visit the website by using the computer in a place, but can be more easily and quickly to get the required information at any time and place. The rapid development of a large number of mobile devices leads development of GIS to mobile computing GIS, and gradually become an important direction of development of GIS.

The paper puts forward the mobile public transport system based on J2ME and SVG by studying the characteristics of mobile GIS. People can locate their own location and query the line, bus site, real-time operating vehicles information anytime anywhere to better plan their travel plans by using the system.

2 System Design

The system choose the J2ME as mobile phone client development platform, and the Mobile SVG as the format of digital maps, using wireless communication technology to achieve mobile devices and server-side for the real-time information interaction.

The system combines mobile J2ME development platform technology, global satellite positioning technology. Wireless communication technology, geographic information system technology and other technologies. The system composes car terminals, wireless communication network, the server-side etc.

The server-side deal with the vehicle geographic data, and thee client's data require, returning the results to the client-side the map-server responses the requests of web server to generate map: database server storage spatial data, bus information data and so on. Mobile terminal achieve to surf the Internet any time any where through the mobile communication technology(GPRS,CDMA,3G etc),and obtain the required maps, bus sites and other information through the HTTP protocol connecting to the server-side. Each bus has on-board GPS terminal to receive the positioning signal, to collect the location data of bus and data continuously were uploaded to the server-side platforms in real time.

3 The Key Technology of System

Locate function. The mobile devices with built-in GPS module enable to locate the position in the map, and to view the surrounding traffic conditions, the building features and other information. When we need to obtain the positive information, the threads of system according to the need set the time interval of reading position information. It avoids too frequent updating data, resulting in the waste of system resources. The system only need to get the latitude and longitude information for determining the location of the holder device. The main thread opens a separate thread to get the data. When it opens the thread, the system carries out operation of request coordinates every three minutes or longer and or shorter the time interval to obtain the current latitude and longitude. After the data obtained, threads enter the waiting state to response to the next request of GPS data. When users have been obtained location information unnecessarily, withdraw the function is back to main thread.

Layers control. The map is layer collection. The first is the plan of each layer for map design. Map layer control is indispensable function, and can choose loading required layers to clearly see the important features of the map.

The domestic and foreign large-scale GIS mapping software are all based on the layer mode to management geographical spatial data. The spatial data is said by dot, line and region. And every geometric element of same characteristics makes up a separate layer. The vector diagram of SVG application the method that manage spatial data with layers [2]. This system identify six layers of the experiment data processed in <g> tags.SVG document is a DOM structure, and we can obtain the layer through traversal the <g> and </g> of SVG document. Control visibility of layer properties is true or false, in order to achieve dynamic layered shows the map layer.

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

Query algorithm for path. This is the basic idea for algorithms based on transfer with least number: transfer twice most; two-way search lines and stations from the starting point A and ending point B separately; compare each site on public transport network site which can be transferred; get all possible paths from A to B; analysis path distance and determine the optimal path. First identify the sets of lines and stations, the set made by inquiring the table in database. The set of lines or stations is like this:

- (1) $L_a(i)$ ($i=1, \dots, n$) represents all lines which through starting sites S;
- (2) $L_b(j)$ ($j=1, \dots, n$) represents all lines which through ending sites E;
- (3) $E_i(x)$ ($x=1, 2, \dots, p$; p is a positive integer) represents stations which through starting sites S in $L_a(i)$;
- (4) $F_j(y)$ ($y=1, 2, \dots, q$; q is a positive integer) represents stations which through ending sites E in $L_b(j)$;
- (5) $H(m)$ (m is a positive integer) represents all lines which through $E_i(x)$;
- (6) $I(n)$ (n is a positive integer) represents all lines which through $F_j(y)$;
- (7) $D(a, b)$ represents the distance between station A and B.

In this Algorithm, assuming that people can accept the adjacent range is 50 meters. $W \leq 50$ represents for nearby stations. It can obtain nearby stations by searching buffer features in the real map within 50 meters, avoiding querying nearby stations from the database.

The flow chart is shown in figure1.

The procedure of algorithm that getting optimal path between two points is as follows:

- (1) Search in the bus lines and sites list, finding out the line set $L_a(i)$ which through station A and $L_b(j)$ which through station B;
- (2) If $L_a(i) \cap L_b(j) \neq \Phi$, then A direct B and jump to (3) to get the optimal routes. Or jump to (4) to transfer;
- (3) If there is only one direct route, then this route is the best one and it can be output. If it needs transfer for one time: first assume the stepping-stone is S1; then calculate the distance from A to S1 and the distance from S1 to B; at last chose the shortest one as optimal routes. If it needs transfer for two times: the way is same as transfer for one time;
- (4) Search in the bus lines and sites list, finding out the station set $E_i(x)$ which are in the line set $L_a(i)$ and $F_j(y)$ which are in the line set $L_b(j)$. if $S1 = E_i(x) \cap F_j(y) \neq \Phi$, then from A to B needs transfer for one time and the stepping-stone set is S1, jumping to (5) to getting the optimal routes. Or jump to (6) to transfer for two times;

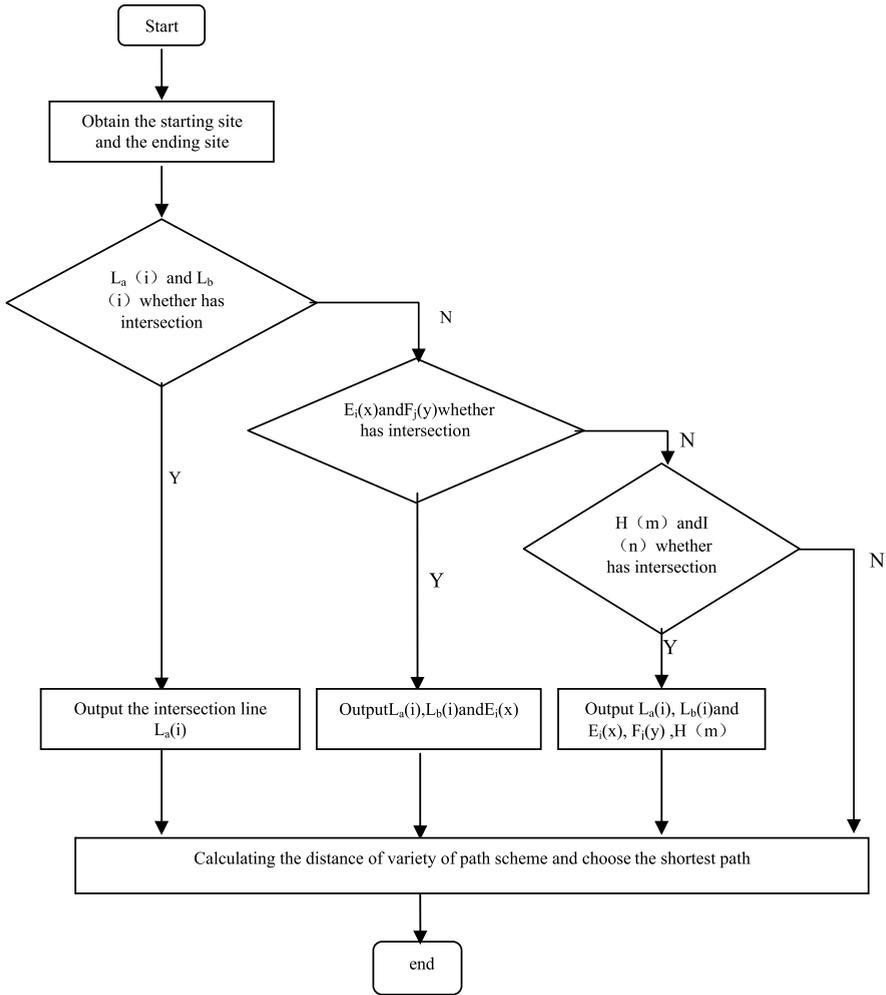


Fig. 1. Flow Chat of Solving the Shortest Path

(5) From starting site A to any site in S1 by bus, then reach the ending site B. Each transfer route has a transfer site and two bus sections. Using the way in (3) to calculate the distance of the two bus sections and chose the shortest one as optimal routes. Output the result and end the algorithm;

(6) Search in the bus lines and sites list, finding out the line set H(m) which through stations in E_i(x) and I(n) which through stations in F_j(y).if R= H(m)∩I(n)≠Φ, then from A to B needs transfer for two times and jump to (7) to get the optimal routes by transferring for two times. Or it needs transfer for three times at least and this can be treated as unreachable, then end the algorithm;



(7) The first transfer station set S_1 is $L_a(i)$ and R intersection sites. The second transfer station set S_2 is $L_b(j)$ and R intersection site. From starting site A to S_1 to transfer by bus, then reach the second transfer station S_2 , at last reach the purpose station B. In this situation each transfer route has two transfer sites and three bus sections. There would have many routes from A to B. Using the way in (3) to calculate the distance of the three bus sections and chose the shortest one as optimal routes. Out put the result and end the algorithm.

Multithreading technology. Normally, in the applications program of mobile equipment, programs trigger a series of corresponding operation through keystrokes of equipment. The programs of public transport information system will has a large number network connection operations, reading flow of data, map or bus information data which waste relatively resources and time. These operations usually take a few steps and waiting time to finish. Therefore, the process that deals with each incident response can build multithreading response to avoid causing dead lock in programs.

So each application program should contain at least one thread: main thread, which is equivalent to entrance of program. Other threads are established through constructors of threads (the Thread) class or instantiation inherited Thread classes. In this system, the J2ME client does not need frequent update data message, so system uses the method realize multithreading interface [3] to control the operation of Internet connection, GPS data read.

Storage of map data and data communications. In J2ME platform, mobile device memory capacity is small and there is no guarantee for a network with data storage quickly. So compared with J2SE platform with huge memory and hard disk capacity, it need consider more factors to solve the problem to reduce the data storage. In J2ME platform, there are two ways to store the data for routine exits with not losing data. They are RMS and file system. By investigating of cell phones, Nokia Symbian40's RMS has 20K memory capacity and MOTOROLA phones generally have 64K or more than 64k space to store data. These are normal cell phones. Normally, in bus information system, the map data is generally not less than dozens of Ks. So, if store the server data in RMS, testing programs will throw exception that the data is over the capacity of storage (Record Store Full Exception). It because there is not enough space for recording data. So if the system supports JSR75, choosing file system to store map data is an effective way. It stores some data into the memory card.

4 System Test

As the map of the north city for example, system makes SVG vector diagram design, and release public transport information in mobile bus information system. Users can obtain their position as well as the transit site information, and can real-time require the bus status information. As shown in figure 2, and figure 3.



Fig. 2. The real-time information of bus status and transit site



Fig. 3. Scheme Diagram of Main Interface

5 Conclusion

Using the J2ME and SVG technology for realizing mobile transportation information system, the experiment shows that the technology is possible. Since the J2ME has cross-platform, network security, code portability and other features, developing mobile GIS bus information system in the J2ME development environment can easily access its location and public transport information. It provides a convenient travel of users. By modifying the system, it can also be used for vehicle navigation. So it has a very broad application prospects.

Acknowledgment. This research is supported by scientists and engineers service businesses (2009GJA10020), application base and cutting-edge technology research plan of Tianjin (10JCYBJC00200).

References

1. Jeffery: Intelligent transport systems for traveler information. *Highways and Transportation* 46(4), 21–23 (1999)
2. Li, Y.: The Research and Application of WebGIS Technology based XML-SVG. North China Electric Power University, Hebei (2006)
3. Lee, E.A.: The problem with threads. *IEEE Computer Society* 39(5), 33–42 (2006)
4. Ouyang, Q., Xu, X.: Architecture of J2ME Platform and Application Development. Wuhan Marine Communication Research Institute. *Computer&Digital Engineering* (02) (2006)
5. Guo, K., Wei, Q.: The Research and Application of embedded bus query system based on J2ME. *Fujian Computer* (2), 22–23 (2009)
6. Kang, M., Peng, Y.: The Key Technology and Application of Mobile GIS. *Bulletin of Surveying and Mapping* (9), 50–53 (2008)
7. Yang, Y.: Positioning Technology for Mobile GIS. *Geospatial Information* (4), 67–70 (2009)

Based on the ExtJS Technology and SSH Framework Authority Management Research

He Yueshun¹, and Luo Min²

¹ East China Institute of Technology, Nanchang, China

² Jiangxi Institute of Computing Technology, Nanchang, China
hys8418@163.com, Huangrp0206@gmail.com

Abstract. Authority Management has become an important part of information system, this article describes the ExtJS technology and SSH framework, and study the tree structure management of the system architecture and operational mechanisms; Carried out analysis and design of information rights management platform, and the application of ExtJS technology and SSH framework to achieve the whole platform for the information required multi-level dynamic management function module, the user operating the distribution rights, distribution of its designated role of rank and authority, the user's dynamic management, and can be specified user for the specified roles assigned to realize the different users on different modules of different operations; Realize the dynamic manage of multi-level departments by using of the ExtJS page's tree structure. Meet the design objective, with good results.

Keywords: ExtJS, SSH, Authority Management, Information Platform.

1 Introduction

Computer network management information system is widely used as a mode of information management in the modern enterprise, in the actual management of managers with different permissions as the user of management information systems, the necessity and importance of authority management must be reflected in the information management system. Rights management is the core of system security, and through permissions settings and maintenance, to prevent the computer system and resources unauthorized access, ensure that only appropriate personnel have access to appropriate services and data. Authority management has been the security of computer systems research and application of hot spots. In particular, the modern application environments, such as the distributed electronic library, heterogeneous information systems, collaboration systems, workflow applications, etc., access to resources not only needs an access control policy to complete; In addition, the resources of the size and complexity in the distributed system, making design a security mechanism to protect system resources become more complex and difficult, it is the concern of every business. So to design a distributed environment to support multiple access control policies and flexible authority management model for research has gradually become a new topic of concern.

Authority management is an important but complex area in the Safety Management. Several typical access control policy: MAC (Mandatory Access Control), DAC (Discretionary Access Control) and RBAC (Role-based Access Control). RBAC expressed a natural organizational structure, its function is divided into the role, but the user is granted one or more roles. Authority is granted to the role rather than individual users, users assigned to a role up access permissions. RBAC's main objective is to simplify the authorization; simply assign a new user role can change the functions of the user. Role for a given function provides a simple security policy that a person can easily determine authorization information. In contrast, a typical model of access list, you must search the entire collection of authority to judge a person's actual authority.

2 Use the Key Technologies

2.1 ExtJS Technology

ExtJS[1], the predecessor of the YUI (Yahoo User Interface), through continuous development and improvement, has become the most complete and mature set of building RIA Web application is a JavaScript library. Using ExtJS build RIA Web application; it has the same standard user interface and mode of operation with desktop applications, and can be across different browser platforms. ExtJS now becomes a perfect choice of developing better user experience of web application program .With the user experience have become increasingly demanding, for Web application developers, ExtJS is undoubtedly the best solution to help developers to quickly achieve a good user interface development.

2.2 SSH Framework

Struts framework is based on MVC (model view controller) model framework, is a part of the Apache Software Foundation's Jakarta project team, a free and open source Web application framework layer. In the normal Struts application, the user request in the processed data will be on display in a new page, namely through transfers Struts the mapping. find Forward method to carry on page skipping, it is very familiar with the use of methods, But when ExtJS page and Struts union another kind of situation appeared. In most cases, ExtJS page do not execute the page jump, but only a request the necessary data to the server business and update the related components of page. Therefore, no need to configure Struts in Action the Forward label, but write business data directly back to the client by HttpServletResponse.

Spring Framework [2] is based on the J2EE implementation of a lightweight J2EE framework. It serves all aspects of the application, provides the basis configuration of Bean, AOP support, JDBC extraction framework, the abstract transaction support, etc. It is also effective organization of the system middle layer objects, eliminating the component object to create and use closely coupled problems.

Hibernate is a powerful object / relational mapping (Object Relational Mapping, ORM) framework [3], the operation of the JDBC access to the database were lightweight package. Using Hibernate for ORM technology, relational database tables

can be mapped into the data object, and to ensure that source of concise and complete object-oriented style.

3 Systems Analysis and Design

Development of information platform with MyEclipse integration Struts, Hibernate and Spring framework, database design using SQL Server 2005. Meanwhile, using the without refresh technology of Ajax to improved the efficiency of the page visit in the ExtJS page.

According to the needs of information platform, powers the system is divided into functional modules, authority, role, grade, Users, department management of six functions. Modules manage which ExtJS tree page to complete the required information platform for the entire multi-level dynamic management function module; Authority management is on the ultimate child nodes of multi-function module of dynamic management, that is, the users need to operate finally the module of dynamic management; grade management completed the distribution of the user operation, different levels of users will have different operating levels; Role function to complete owned grade and authority distribution of the designated role; Users function to complete the dynamic management of users on the entire information platform, and users can specify the allocation of a specified role to achieve the different users in different modules for different operating systems; And can be specified for the specified user roles assigned to realize the different users of different modules of systems for different operating; Department management that is realize the dynamic manage of multi-level departments by using of the ExtJS page's tree structure. The design authority function data structure diagram in SQL Server2005, shown in Fig.1.

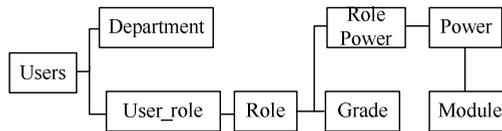


Fig. 1. The Structure of Permissions Function

3.1 System Implementation

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

- Asynchronous implementation to dynamic tree of Department ExtJS.

In the application [4], we are often involved in to display or handle the tree structure of the object information, such as department information, local information, or menu information of the tree structure, the operating system folder information and so on. For the traditional html page to display the tree is more

difficult, to realize that is need to write a lot of JavaScript, especially for asynchronous loading of the tree based on Ajax, the Ajax is not only related to data loading and processing technology, also need to consider inter- browser support and so on, dealing with very troublesome. ExtJS have the tree controls, these controls can be in B / S applications to quickly develop a tree structure that contains information on the application. With these controls in B / S applications, rapid development of the application that contains the tree structure information. Although the default icon of Ext JS [5] to be set display the file and folder for the tree node, but not limited to the concept of file systems, file icons and tree nodes display the contents do not need to defined code, and can be based on dynamically changed . Department management and module management of authority system is by the dynamic tree of ExtJS combined with SSH background framework achieved. Here's an example to departments to illustrate the application of dynamic tree.

ExtJS tree node of the core of dynamically loaded is through Ajax technology to return a group of JSON (JavaScript Object Notation) object to achieve, [(id: 1, text: 'child nodes', leaf: false)] is a group of the most common JSON object form to achieve tree nodes. The id attribute is used to mark the tree nodes, the content of text attributes will be displayed on the node of page, leaf attribute is used to mark the tree nodes have child nodes, if the value is true, then representing the node without child nodes. On the contrary representing the node has child nodes. In view of these basic properties, using SQL Server 2005, designed the field that is corresponding of JSON object form to implement the department dynamic list,as shown in Table 1. The remark field actually represents the node level, we define '1' represented a child node, '0' represented non-child node, child nodes under fatherid (parent node ID) to determine the loading order. In this way, we can accord the level of each tree node to create a dynamic tree of multi-level department, shown in Fig2.

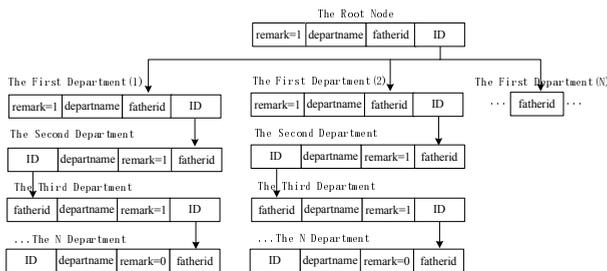


Fig. 2. The Dynamic Tree of Multi-level Departments

According to these characteristics of point, the design ideas of loading tree node in two steps, the first step in accordance with the fathered and the remark to load the entire root node.



Table 1. The design of department table field

| Field Name | Description | attributes |
|----------------|-----------------|------------|
| ID | Automatic ID | id |
| departmentname | Department name | text |
| fatherid | Parent node ID | |
| remark | child node | leaf |

```
// first step to expand the root node (0), loading the root of all fatherid == 1
if (node.equals ("0")) (
// ... Get department name, ID identifies, the parent node, whether there is a
child node
for (int i = 0; i < count; i++) {
// If the Remark == 0 indicates that after this record haven't child nodes
if (1 == fatherid && "0".equals Remark) {
departmentList += (" {id : " + ID + ",text : " + Department + ",leaf : true},");
}
// If Remark == 1 indicates that after this record has child nodes
else if (1 == fatherid && "1".equals(Remark)) {
departmentList += (" {id : " + ID + ",text : " + Department + ",leaf :
false},");} } }
```

The second step, as the dynamic tree node of foreground ExtJS is through asynchronous Ajax achieved, each event of expand the node will be triggers an event and can put the property of the node transmitted to background. When expand the root node, we set the ID of the root node as the search object, corresponding database record fathered, load all the root node corresponding to the next level of child nodes.

```
// ... Get department name, ID identifies, the parent node, whether there is a child
node
for (int i = 0; i < count; i++) {
//load the child nodes that have non-child node
if(node.equals(Integer.toString(fatherid))&& "0".equals Remark) {
departmentList += (" {id : " + ID + ",text : " + Department + ",leaf : true},");}
// load the child nodes that have child node
else if(node.equals(Integer.toString(fatherid))&&"0".equals Remark) {
departmentList += (" {id : " + ID + ",text : " + Department + ",leaf : false},");} }
Finally return to corresponding JSON objects to the foreground and to the tree
structure displayed in the Web page.
departmentList = departmentList.substring(0,departmentList.length() - 1);
departmentList += ("]");
// Finally return a JSON object to the foreground to display tree node
```

- Users - Function module – authority of the load ExtJS dynamic tree
- The background management design of Function module – Authority similar to Department Management using the ExtJS dynamic tree to achieve, but because of the authority system, different users have different authority, that is, different users can specify the distribution visit with the parent node and corresponding child nodes. Therefore, we to improve in the basis of the department tree, the module



database design by adding the module loading order (moduleorder) field, the authority design by adding the authority loading order (powerorder) field, the Users (Users) will be loaded under the authority sequence of load module node in the foreground.

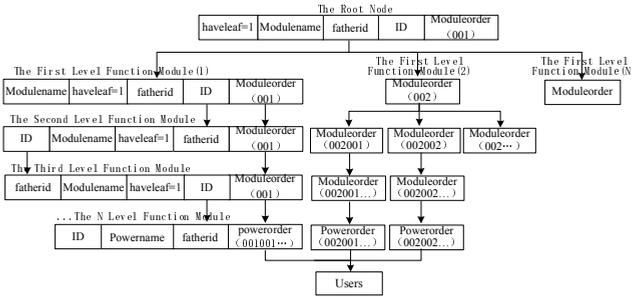


Fig. 3. User-Function Module-Permissions Dynamic Tree

three values that can see from Fig 3. Therefore, we need to intercept the top three in each element of the List

According to these features of the tree node, after users log on, the design concept that load module tree node are divided into three steps, the first step to read the login user has authority (powerorder) from the database, and all the authorities into a List of array powerorderList .

List powerorderList=new ArrayList();//Definition List array powerorderList.add (powerorder);//put into all powerorder of Login user

The second step is that interception and screening of all the elements of the array, and then load the entire root node. For example, we first read an array of {001001, 001002001,001001002,002001,002002001,003,003001 ...}, the Moduleorder of one-level functional module is array to re-arrange, and remove the one recurring element. For the elements of example array that we will filter out repetitive elements, to arrive at a new array of "(001,002,003 ...)" to load the node. So we designed the moduletempList of another array List to store the users need to load the root node. At the same time put the loading order of the node of need the loading (moduleorder) into id attribute of JSON object.

```

for(int k=0;k<powerorderList.size();k++)// Loop through user authorities
    {powerorder=powerorderList.get(k).toString().substring(0,3); // Interception
of the top three
    if(moduletempList.contains(moduleorder)==false&&haveleaf!=0) // Filter
element
        {moduletempList.add(moduleorder); // Store the root node need to be loaded
        moduleList += ("id : " + moduleorder + ",text : "+ modulename+ ",leaf :
false,");}
// Return the JSON object

```

The third step is when expand the root node, we set the ID of the root node as the search object, filter out the elements that the top three is equal the id of root node in the all powerorder array. If we expand the root node id for the 001, you



need to load the array should be {001001, 001002 ...}. And so on, load the node of two-level modules is a 6-bit id, three-level modules is a 9-bit id, until the haveleaf of load nodes is 0, that is end of the load until the final child node.

```

int chilidnodeLength=node.length()+3; // Get the length of child node need to load
for(int k=0;k<powerorderList.size();k++)// Loop through user authoritys
{String havepoweroeder=powerorderList.get(k).toString();
// Find the leaf nodes elements that need to load in the authority array
if(havepoweroeder.length()->=chilidnodeLength)
// Filter out elements that match the length of the root node in all powerorder
if(node.equals(havepoweroeder.substring(0, node.length())))
{if(chilidnodeLength==havepoweroeder.length())// Lower level node is the final child node
(... Add the final child node)
else // lower-level leaf node has child nodes
{powerorder=havepoweroeder.substring(0,chilidnodeLength);
// Find the corresponding modules by Dao layer
List ModuleList= userRoleAgentI.findModule(powerorder);
if(ModuleList.size()->0)
{... Load the lower leaf node } } } }

```

- User - Role - grade dynamic control

Each user will be given a specific role, however role function to complete owned grade and authority distribution of the designated role. User- Role - grade be shown in Fig4. Functional role combined with authority function control user operations dynamic loading of function modules. The user grade function control the content of the user can operation in Web pages.

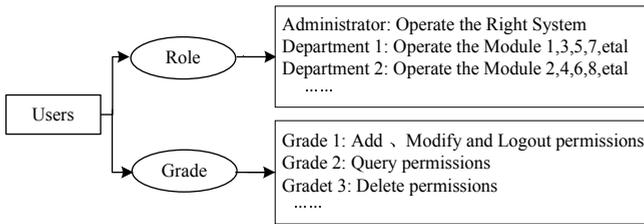


Fig. 4. The Relation of User-Role-Grade

4 Conclusions

Authority management system break the previous function modules designed of the limit fixed, users only need to understand simple load models in the function module, can be according to their needs to add multi-function modules of their require in the authority system . As the dynamic implementation of the function module, for the programmer, only needs to know the specific content of function modules the user required developing the appropriate interface, rather than fixed design requirements

module in the programming. Authority Management System, not only to make the system more flexibility, maintainability and scalability, and allow users to more effective and standard to participation in development process.

This paper mainly on the structure of authority management systems to analysis and design, while achieving the authority management functions, However, the text of the algorithm that retrieve the node has not been optimized, as information platform the amount of information of increase, the system function modules, and the speed of retrieval algorithms will directly affect the B / S structure system shows the efficiency of the network. Therefore, the authority management system will be optimized on the search algorithm to further improve the practicability of the whole system.

References

1. Wei, J., Xia, H., Meng, L.-C., et al.: ExtJS Web application development guide. Mechanical Engineering Press (2009)
2. Chen, X., Sun, W.D., Sun, D.: Based on SSH framework and AJAX technology of the Java Web Application Development. Computer Engineering and Design 30(10), 2590–2591 (2009)
3. Wu, M., Yang, Y.: Based MVC pattern of the Java mainstream framework of the integration technology. Computer and Digital Engineering 37(10), 92 (2009)
4. Gong, P.: ExtJs Practical Concise Guide, <http://wlr.easyjf.com>
5. Frederick, S., Ramsay, C., Blades, S.(Cutter): Learning Ext JS. Packt Publishing Ltd. (November 2008)

Public-Private Partnerships and E-Government

Faizullah Khan, Surat Khan, and Bin Zhang

School of Economics and Management Beijing University of Posts and Telecommunications
Beijing, P.R. China

{khan_kakar, suratkhan}@yahoo.com, binzhang@bupt.edu.cn

Abstract. The rationale to exploit the potential of ICT and e-government are compelling for most of the countries. However, implementation of e-Government projects requires the Capital investment as well as funds required for operation and maintenance of ICT systems on sustained basis. Facing excessive financial constraints governments are not in a position to extend an open-ended financial support to e-government projects. It therefore becomes imperative to find new innovative methods of financing ICT and e-government projects by including the expertise and finance of private sector. A Public-Private Partnership is a legally enforceable contract between a private sector entity and a government body that requires the private partner to deliver a desired electronic public service, and share the associated risks of service delivery. This paper analyzes Public-Private Partnerships in general and e-government PPP projects in particular, and the need for employing PPP for implementation of e-Government projects in Pakistan.

Keywords: Public-Private Partnerships (PPPs or P3), e-Government, Public service, Value for Money, Self-Financing Projects.

1 Introduction

1.1 E-Government

E-Government can be defined as the use of Information and Communication Technologies (ICTs) by governments to enhance the range and quality of information and services provided to citizens, businesses, civil society organizations, and other government agencies in an efficient, cost-effective and convenient manner, making government processes more transparent and accountable and strengthening democracy.[1]. To exploit the potential of e-Government to improve citizen access to service delivery a minimum threshold of technological infrastructure, human capital and e-connectivity are the pre-requisites. Considerable resources are required to establish, expand and constantly update e-networks. E-Government projects are long-term endeavors, requiring large capital infusion in software, hardware, infrastructure and training. A viable financing plan should not only pay for the immediate needs to jumpstart e-government; it must also consider its long-term financing options for the sustainability of the project. [2]

There are various financing models for funding e-government projects including government funding through annual budgetary allocation, issuing bonds by governments, outsourcing and Public-Private Partnerships (PPPs). Public-Private Partnerships are better options because the government is relieved of the financial and administrative burden of providing the service, although it retains the important role of regulating and monitoring the performance of the private partner.

1.2 Public-Private Partnerships

Public-Private Partnerships are partnerships between the public sector and the private sector for the purposes of designing, planning, financing, constructing and/or operating projects which would be regarded traditionally as falling within the remit of the public sector. [3]. The popularity of PPP arose initially out of the government need for financing to meet increasing demand for expansion and rehabilitation of physical infrastructure such as roads, energy facilities, and water and sanitation networks. The reasons of government motivation towards PPPs include potential for Value for Money (VFM), early project delivery, gains from innovation, obviating the need to borrow to finance infrastructure investment, and access to improved services. [3] The PPP model gains acceptance, and is rapidly expanding in to all areas of public life, including ICT resources and e-government. [4]

This paper is organized as follow: section-II is a brief review of PPP and its different schemes and modalities; Section-III is brief review of status of PPP in different countries of the world in the field of ICT and e-government Section-IV presents the status of economy of Pakistan. Section-V is a brief review of e-Government in Pakistan. Section-VI presents suggestions for implementation of e-Government projects in Pakistan through PPP. Section-VII concludes the paper with suggestion for future work.

2 Forms and Models for PPP Projects

P3 arrangements are basically contracts between a private sector entity and the government that call for the private partner to deliver a desired service and assume the associated risks. [5] At the heart of all PPPs is the deployment of private sector capital. Within a PPP framework, this can result in greatly improved value for money (VFM) for the government in terms of the risks transferred to the private sector. The different Schemes and Modalities generally used for PPPs are given in Table 1 [12]. Some of the key PPP models employed for e-government projects include Build-Own-Transfer, Design-Build-Own-Operate, Lease-Develop-Operate and Build-Own-Operate-Transfer etc.

3 Public-Private Partnerships in E-Government: A Global View

Initially started for infrastructure development projects such as roads, rail tracks, airports, hospitals, schools etc, PPPs are now formed for ICT and e-government projects world wide.

Table 1. PPP schemes and modalities

| Schemes | Modalities |
|--|---|
| Build-own-operate(BOO) Build-develop-operate(BDO) Design- construct-manage-finance(DCMF) | The private sector designs, builds, owns, develops, operates and manages an asset with no obligation to transfer ownership to the government. These are variants of design-build-finance-operate (DBFO) schemes |
| Buy-build-operate (BBO) Lease-develop-operate(LDO) Wrap-around addition (WAA) | The private sector buys or leases an existing asset from the government, renovates, modernizes, and/or expands it, and then operates the asset, again with no obligation to transfer ownership back to the government |
| Build-operate-transfer (BOT) Build-own-operate-transfer (BOOT) Build-rent-own-transfer (BROT) Build-lease-operate-transfer (BLOT) Build-Transfer-operate (BTO) | The private sector designs and builds an asset, operates it, and then transfers it to the government when the operating contract ends, or at some other prescribed time. The private partner may subsequently rent or lease the asset from government |

Each of the state governments in the United States has an online website for citizens to access government services. For instance the National Information Consortium (NIC) a private firm provides free portals to 18 states and seven municipalities under PPP model. NIC offers to implement its transactional portal system for no charge, then collect a fee from each user or get a percentage of revenue the system generates. [6]

The Malaysian government e-procurement system or “e-Perolehan” is a secured e-procurement service. E-Perolehan is financed through PPP (BOT scheme) by Commerce Dot Com Sdn.Bhd. Commerce Dot Com Sdn.Bhd receives a transaction fee on each completed sale.[7]

The Indian State of Andhra Pradesh’s pioneering electronic service project (eSeva), first launched in 1999, has created “one-stop-shop” kiosks throughout the state where a multitude of G2C services may be transacted from a single point. eSeva is PPP project between the India’s state of Andhra Pradesh and Tata Consulting Services (TCS) under BOOT model.. eSeva was initially launched in Hyderabad with 43 service centers. Later on expanded to rural areas through three separate kiosk projects, Rural e-Seva, Rural Service Delivery Points (RSDPs), and Rajiv Internet Village Centers.[7][8]

The South African Department of Labor ICT project is another successful PPP Model project. The Labor IT systems include the highly complex payment of salaries and bonuses, and vacation or sick time accrued. By contracting out service provision to the private sector, this project enabled the Government of South Africa to provide prospective and current employees with quick and easy access to information, forms, tests, pay records, time-off records, and credential records necessary to manage their careers. [4][9]

ESDlife was a bilingual portal in Hong Kong, which was developed and maintained through a Design-Build-Own-Operate (DBOO) PPP that implemented the government's Electronic Service Delivery (ESD) Scheme from 2000–2008.

Under this contract, the private operator, ESD Services Limited (a JV between Hutchison Whampoa Limited and Hewlett-Packard HK SAR Limited), was responsible for developing, financing, operating, and maintaining the portal, and the government was responsible for paying transaction fees to the private operator after the transaction level had reached a pre-agreed volume. In turn, the private operator was allowed to expand the portal to include commercial services, such as advertisements and e-commerce, to generate additional income. [10]

4 Economy of Pakistan and PPP

Located in South Asia, Pakistan is the seventh most populous country of the world with a population of more than 165 million. Neighbored by Afghanistan, Iran, India, and China, Pakistan has one of the most strategic geographical locations in the world. Due to US invasion on Afghanistan, Pakistan is the only logistic support route for US and allied troops in Afghanistan. Being first line state against war on terror, Pakistan faced serious domestic political uncertainties and undesirable law and order situation which seriously affected the economy. Table below shows the Annual budget position of country for financial year 2009-10 and 2010-11. [11]

Table 2. Economic status of Pakistan

| | 2009-10 (Million PKR) | 2010-11 (Million PKR) |
|-------------------------------|--------------------------|--------------------------|
| Total Expenditure | 2585557 | 2764437 |
| Total PSDP | 510000 | 663000 |
| PSDP as %age of Total Budget | 20 | 24 |
| Budget Deficit (Loans) | 667094 | 553164 |
| Loans as %age of Total Budget | 26 | 20 |

The Public Support Development Program (PSDP) is mainly supported by external and internal loans and grants. Pakistan's total public debt stood at an estimated PKR. 8160 Billion as of end-March 2010. At this level Public debt is equivalent to 56% of GDP, and 379% of total budgeted revenue of the year.

Keeping in view the economic crisis, Fiscal Constraints, Budget deficit and heavy Public debt, and recognizing the importance of improving and expanding infrastructure assets and services for sustainable economic and social development, the democratic government of Pakistan is now considering supplementary and innovative approaches to infrastructure provisioning by including the expertise and finance of the private sector. Private sector financing can allow government to bring forward projects that might be otherwise being delayed because of budget constraints.

The Government of Pakistan Issued Pakistan Policy on Public-Private Partnerships which was approved by Economic Coordination Committee (ECC) on January-26, 2010. The policy covers PPP projects in areas of Transport and Logistics, Mass Urban Public Transport, Local Government Services, Energy Projects, Tourism Projects,

Industrial projects including industrial parks and special economic zones, Irrigation projects, and social infrastructure which includes education, culture, and health infrastructure.[13] E-Government implementation specifically is not included in PPP policy.

5 E-Government in Pakistan

e-Government is the use of ICT to promote more efficient and effective government, facilitate more accessible government services, allow greater public access to information, and make government more accountable to citizens. The Government of Pakistan introduced e-Government as priority area in its first National IT Policy and Action Plan in 2000. But the policy speaks nothing about the provision of e-Government services to citizens and businesses. [14].The government established e-Government Directorate (EGD) in October-2002, to plan, implement and lead e-Government efforts in Pakistan. EGD under Ministry of IT issued its first E-Government Strategy and 5-year Plan for the Federal Government in May-2005. [15]. EGD and Ministry of IT has completed many projects of automating the record and working of different departments including establishment of websites for almost all public sector departments, legislative and judicial institutions. Although a lot of useful public sector information can be found online, however all the government websites have only English version, which majority of the population could not understand. The goal of e-Government is not merely to computerize government records, or provide some basic information online, but the ultimate goal of e-Government is to transform government and make government citizen-centric. After lapse of 10 years of issuance of IT policy and Action plan, in 2000, the government is providing no basic citizen service online. People still have to travel long distance; spend time, money and effort, to visit government offices for driving license renewal fee, birth/death certificate, land ownership certificate (Fard), voter registration and payment of government taxes etc. They still have to stay and wait in long queues for payment of their utility bills. They still have to face traditional bureaucratic behavior of public sector officials.

The websites are not being updated regularly. For example, the parliament of Pakistan has passed an amendment in which the name of one province, North Western Frontier Province (NWFP), was changed and new name, Khyber Pakhtoonkhwa was given to it. The President of Pakistan signed the amendment on April-19, 2010, after which the amendment is part of the constitution of Islamic Republic of Pakistan. But many of the website of Federal and Provincial governments still publish old name for Khyber Pakhtoonkhwa Province on their websites.

Similarly in 2002, a minimum qualification of bachelor degree for candidature to the membership of Legislature was prescribed, by the then President and Chief Executive of Government of Pakistan. But in 2008 general elections which was second general election after that order, more than 150 candidates for legislature on bogus bachelor degree certificates not only qualified for contesting election, but won the election and became members of different legislative assemblies. This is either due to lack of coordinated system between election commission and degree awarding universities during nomination papers scrutiny, or unavailability of computerized

database of Universities record. This not only causing heavy loss to government exchequer in the form of re-election costs, but also unnecessary waste of time of higher courts to decide the petitions of candidates against each other. The most serious issue is the creation of distrust in citizens about political leadership and government system.

6 PPP for E-Government Implementation

Studies by World Bank and other UN organizations conclude that there is a strong relationship between infrastructure and economic development. Efficient infrastructure creates employment, develops human capital, promotes local and foreign investment and trade, raises standards of living, and improves access to critical services. The mismatch between the investment needs of Pakistan and the resources available from the government has increased reliance on Public-Private Partnerships. But for PPP to be successful an appropriate institutional framework characterized by political commitment, good governance, and clear supporting legislation is needed. E-Government helps reduce corruption, increases openness and trust in government which is fundamental to good governance, and thus contributes to economic policy objectives. The deployment and use of ICT and e-Government require telecommunication infrastructure to provide network access, electrical infrastructure to make ICTs work, skills infrastructure to keep all the technology working, Money to buy or access ICT, usage skills to use ICT, and Literacy skills to read content.

E-Government is one of the areas which are eminently suited for PPP, because the technological know-how, managerial expertise and financial resources required for e-government lies primarily with private sector. International experiences show that if a suitable framework for PPP in e-government is designed it can become a big leap forward in the efficient, convenient and cost effective delivery of public services to citizens. While employing PPP in e-government the principle of “Think Big, Start Small, and Scale Fast” might be followed. Rushing forward with ill-conceived e-Government plans can be a costly mistake, financially and politically. The e-government projects will sustain only when there are revenue models attached with it. Because the private partner has primary interest in revenue from the projects, therefore initially those projects which have proper revenue model must be initiated as PPP. The following e-government projects might be offered initially as PPP for implementation

- 1) Electrical Power Generation sector to make the ICT and e-Government systems work. Pakistan is facing serious electrical energy crisis.
- 2) Design, development, financing and management of Multi-service bilingual (English, Urdu) portals for National and Provincial Governments that facilitate seamless integration of applications implemented by various departments and agencies of the government
- 3) E-Procurement for all levels of Public sector
- 4) Front end services like kiosks and integrated citizen service centers that work as one-stop shop, which will provide ubiquitous and affordable access mechanism for citizens

- 5) Establishment of an integrated ICT network for law enforcing agencies, to improve law and order situation in the country.
- 6) Computerization and integration of employee database for Federal and provincial governments, and online access for employees.
- 7) Computerization and integration of record of all academic institutions of the country so that in case of verification the citizens and government agencies have to access one point.

While implementing the e-government projects, the experience and best practices in other countries that have successfully implemented similar projects, might be considered. While designing revenue model for the PPP scheme, the option of self funding approach, by imposing usage fee on citizens, businesses or government agencies, could be adopted. Such kind of self-financing projects have high success rate world wide.

7 Conclusion

In countries like Pakistan which has a long history of dictatorship, political instability, and large-scale corruption, citizens distrust their governments. The government is facing the challenges of growing demand for services, operational and institutional deficiencies, and limited fiscal resources. Private sector is therefore recognized as a valuable source of new technology, management expertise, and investment capital. If properly planned and designed in collaboration with all stakeholders, the PPP arrangements can bring dramatic improvements in the quality, availability, and cost-effectiveness of public services. E-government is a tool for achieving the goals of a clean, transparent, efficient and accountable public sector. E-government can build trust between citizens and government, which is fundamental to good governance, by enabling citizen engagement in policy process, promoting open and accountable government, and helping to prevent corruption. Public-Private Partnership is therefore best option for Pakistan to harness the benefits of ICT and e-Government. The government of Pakistan must revise the National Public-Private Partnership Policy to include e-Government projects. The government must establish clear set of public policy, legal, regulatory, and institutional frameworks within which e-government PPP transactions can develop and operate. Without safeguarding the interests of private partner and other stake holders the PPP partnerships will not succeed. The benefits of PPP can accrue to all stakeholders. For government PPP minimizes financial capital investment, provides protection against technology obsolescence, speedier implementation and better management of e-Government projects, and better image towards citizens. For citizens PPP provides easy access to services at convenient single window or one-stop shops, any time, anywhere and saves indirect costs and hardship. For private sector PPP provides low risk reliable revenue streams and access to new markets.

References

1. InfoDev e-Government Tool Kit and Portal Chapter-1,
<http://egov.sonasi.com/toolkit/overview1/chapter-1-introduction-to-e-government-and-overview-of-the-toolkit/what-is-e-government>
2. Pascual, P.J.: e-Government. e-ASEAN Task Force UNDP-APDIP, 17, 18 (May 2003)
3. Webb, R., Pule, B.: Public Private Partnerships: An Introduction, p. 2, 5. Department of Parliamentary Library, Commonwealth of Australia (2002),
<http://www.aph.gov.au/library/pubs/rp/2002-03/03RP01.htm>
4. Infodev: Public-Private Partnerships in e-Government: Knowledge Map, p. 3, 5, 18. The Institute of Public Private Partnerships Information for development Program (June 2009)
5. Loew, J., McLindon, M.: "Why are countries interested in P3?" A P3 Primer. The institute for Public-Private Partnerships (2002),
<http://www.ip3.org/pub/publication002.htm>
6. Mimicopoulos, M.G.: E-government funding activities and strategies, United Nations New York, p. 19 (2004)
7. e-government toolkit, "Private Partnership case studies," by infoDev,
<http://www.egov.infodev.org/en/Section.64.html>
8. Kuriyan, R., Ray, I.: Public-Private Partnerships and Information Technologies for Development in India. In: International Conference on ICTD, Bangalore, p. 5 (December 2007)
9. <http://www.labour.gov.za/DOL>
10. Infodev: Public-Private Partnerships in E-government: Hand Book, p. 10. Prepared by Institute of Public-Private Partnership (June 2009)
11. Government of Pakistan, Ministry of Finance, "Annual Budget 2010-11",
http://finance.gov.pk/budget/Budget_in_Brief_2010_11.pdf
12. International Monetary Fund (IMF), Fiscal Affairs Department: Public-Private Partnerships, p. 8 (March 12, 2004)
13. Pakistan Policy on Public-Private Partnerships, "Private participation in public infrastructure projects", p. 4 (January 26, 2010),
<http://ipdf.gov.pk/tmpnew/PDF/PPP%20Policy%20FINAL%2014-May-2010.pdf>
14. Government of Pakistan, "IT Policy and Action Plan", Islamabad, p. 25 (August 18, 2000)
15. E-Government Strategy and 5-Year Plan for Federal Government, Electronic Government Directorate Ministry of IT Government of Pakistan (May 2005)

Reverse e-Logistics for SMEs in Pakistan

Surat Khan, Faizullah Khan, and Bin Zhang

School of Economics & Management Sciences, Beijing University of Posts and
Telecommunications, Beijing, P.R. China
{suratkhan,khan_kakar}@yahoo.com, binzhang@bupt.edu.cn

Abstract. Reverse Logistics is receiving increasing attention in business and research due to the growing environmental concern, economic gain and consumer pressure. Reverse logistics deals with the activities and processes associated with the flows of products, components and materials from users to re-users, or its proper disposal.

The main objective of this paper is to develop and implement ICT enabled Reverse Logistics (Reverse e-Logistics) for the Small and Medium sized Enterprises (SMEs) in Pakistan. A Cloud Computing Model is proposed for implementing the Reverse logistics for the SMEs in Pakistan. The govt. support in establishing the Reverse Logistics model is prerequisite.

Index Terms: Logistics, Reverse Logistics, Small and medium sized enterprises, Cloud computing, Pakistan.

1 Introduction

Due to the growing environmental concerns, consumer pressure and increasing international competition the concept of 'Reverse Logistics' has gained significant attention within the realms of academia, researchers, operations managers and company executives. Reverse logistics (RL) deals with the activities and processes associated with the flows of products, components and materials from users to re-users, to recover its residual value, or to proceed to the proper disposal.

Effective reverse logistics management increases the trust of customers, reduce the environmental pollution of products, reduce the materials costs of enterprises, thereby enhancing its competitive advantages and improve the corporate image.

The paper is structured as follows. In Section-II a brief introduction of Reverse Logistics (RL), its driving forces, processes, and operation models are highlighted. In Section-III a statistics about SMEs in Pakistan is given. Section-IV is a brief introduction of Cloud Computing. Section-V concludes the paper by proposing a strategy to make an ICT enabled Reverse Logistics for SMEs in Pakistan using Cloud Computing Model.

2 Reverse Logistics

2.1 Definition of Reverse Logistics

Logistics is defined by The Council of Logistics Management as: The process of planning, implementing, and controlling the efficient, cost effective flow of raw materials, in-process inventory, finished goods and related information from the point of origin to the point of consumption for the purpose of conforming to customer requirements. [1]

Reverse logistics includes all of the activities that are mentioned in the definition above. The difference is that reverse logistics encompasses all of these activities as they operate in reverse. Therefore, reverse logistics is: The process of planning, implementing, and controlling the efficient, cost effective flow of raw materials, in-process inventory, finished goods and related information from the point of consumption to the point of origin for the purpose of recapturing value or proper disposal. [1]

2.2 Difference between Forward Logistics and Reverse Logistic

The operation of the reverse logistics is completely contrary to forward logistics, Table 1 shows the comparison in brief [1].

Table 1. Comparison of Forward Logistics and Reverse Logistics

| Forward logistics | Reverse Logistics |
|---------------------------------|-------------------------------|
| Forecasting straight forward | Forecasting more difficult |
| One to many distribution points | Many to one points |
| Production quality uniform | Production quality uneven |
| Packaging uniform | Packaging often damaged |
| Destination / routing clear | Destination /Routing unclear |
| Disposition option clear | Disposition not clear |
| Speed recognized | Speed often not considered |
| cost easily visible | cost less directly visible |
| Inventory mgt. consistent | Inventory mgt. not consistent |
| Product life cycle manageable | PLC issue more complex |
| Negotiation straightforward | Negotiation complicated |
| Marketing method well known | Marketing complicated |
| Process more transparent | Process less transparent |

2.3 Driving Forces Behind Reverse Logistics

1) *Economics*. A reverse Logistics can bring direct gains to companies from dwindling on the use of raw materials, from adding value with recovery, or from reducing disposal costs. Returns can be a source for value recovery by re-using the return (possibly in a secondary market), remanufacture its parts to be used as spares or recycling the remaining to recover the materials value. [2] Even with no clear or immediate expected profit, an organization can get (more) involved with Reverse Logistics because of marketing, competition and or strategic issues, from which are expected indirect gains, like green image and improved customer and supplier relations [3].

2) *Legislation*. Legislation refers to any jurisdiction indicating that a company should recover its products or accept them back. Post-retail products are recognized as a serious environmental threat, if disposed improperly. In Europe there has been an increase of environmentally related legislation, like on recycling quotas, packaging regulation and manufacturing take-back responsibility [2][3].

3) *Good Corporate citizenship*. Corporate citizenship concerns a set of values or principles that in this case impel a company or an organization to become responsibly engaged with reverse logistics. Good Corporate Citizenship, distinguish a firm by doing well for other people. Some firms will use their reverse logistics capabilities for altruistic reasons, such as philanthropy [2].

2.4 Return Reasons

Products are returned or discarded because either they do not function (anymore) properly or because they or their function are no longer needed.

1) *Manufacturing returns*, i.e., returns related to the execution of production processes, including: raw material surplus, quality-control returns, production leftovers, and by-products.

2) *Distribution returns*. i.e., returns related to the distribution of production to (potential) customers, including: product recalls, B2B commercial returns, stock adjustments, and- distribution items.

3) *Market returns/Customer Return*, i.e., returns from the users of products, including: B2C commercial (reimbursement guarantees), warranties, service returns (repairs and spare parts), end-of-use returns, and end-of-life returns [4].

2.5 Reverse Logistics Processes

Basic Reverse processes or Reverse Logistics activities

1) *Acquisition (or Collection)* refers to all activities rendering used products available and physically moving them to some point for further treatment. Collection may include purchasing, transportation, and storage activities.

2) *Grading (Inspection and separation)* denotes all operations determining whether a given product is in fact reusable and in which way. Thus inspection and separation results in splitting the flow of used products according to distinct re-use and disposal options. Inspection and separation may encompass disassembly, shredding, testing, sorting, and storage steps.

3) *Re-processing* means the actual transformation of a used product into a usable product/component/material again. This transformation may take different forms including recycling, repair, refurbishing, cannibalization, and, remanufacturing. In addition, activities such as cleaning, replacement, and re-assembly may be involved.

4) *Disposal*. Proper disposal is required for products that cannot be re-used for technical or cost reasons. Disposal may include transportation, land filling, and incineration steps.

5) *Direct recovery (Resale or re-use)*, i.e. recovery without any major processing. (E.g. reusable bottles, containers, most leased or rented equipment). It may also happen for surplus goods, e.g. spare parts that are left over after discarding the original equipment [4].

6) *Re-distribution* refers to directing re-usable products to a potential market and to physically moving them to future users [5][6][7][8][9].

2.6 Actors of Reverse Logistics

- Forward supply network actors (such as Supplier, manufacturers, wholesalers, retailers, service providers)
- Specialized reverse logistics actors (such as recyclers, independent remanufacturers)
- governmental entities (such as European Union and national governments)
- Opportunistic players (such as charity organizations) [4].

2.7 The Operation Models of Reverse Logistics

1) *Self-employed model*. In this model enterprise constructs independent reverse logistics system, manages the recovery and recycles the scrapped products by itself.

2) *Pool model*. In this model, same industry firms cooperate with each other; build common reverse logistics system in the form of joint venture (Including recycling network) and offer kinds of service for cooperative enterprises and even non cooperative enterprises, such as transportation, purchase, sales etc.

3) *Outsourcing model*. In this model the enterprise gives part or all business to the particular firm in the form of paying cost. This model is suitable for the most circumstances, either recovery and repair or recycling the waste products. Small and medium-sized enterprises tend to like the outsourcing more, so that they decrease the operation cost.

3 Small and Medium Sized Enterprises

3.1 SMEs Definition

The only way to reduce poverty in a sustainable way is to promote economic growth, through wealth and employment creation. In developing countries, SMEs are the major source of income, a breeding ground for entrepreneurs and a provider of employment. There is no universally accepted definition of SMEs. Generally, countries define SMEs according to the number of employees or level of assets or turnover or both. [11] According to the Pakistan SMEs Policy, the firm having

employment size up to 250, paid capital up to Pak- Rupees 25 Million and annual Sales up to Pak-Rupees 250 Million are Small and medium-sized enterprises (SMEs) [12][13].

Small and Medium Enterprises (SMEs) constitute 90 percent of businesses in Pakistan., According to more recent estimates there are approximately 3.2 million business enterprises in Pakistan, (44% Rural & 56% Urban), , SMEs account for about 30% of Pakistan's GDP, 15% of investment, and 78-80% of employment of non agriculture labor force. They also play a prominent role as existing or potential producers of export goods [13].

3.2 SMEs' Survey in Gujranwala District in Pakistan

The finding of the survey in brief is as under.

1) *Entrepreneur's Educational Qualification:* 49% SMEs were owned and operated by businessmen having inter or graduate level qualification; only 3% of the SMEs were headed by entrepreneurs having master level qualification, 44% are operate by businessmen having secondary and primary level of education and only 4% were headed by entrepreneurs with no education.

2) *Technical/professional Training of Owners:* 70% SMEs owners and proprietors had not obtained any technical or professional training,. Only 4% had obtained formal training, mostly in manufacturing units, whereas 26% obtained informal training from their mentors.

3) *Computer Usage:* 46% SMEs have computer. 57% SMEs using computer for maintaining books of accounts whereas the remaining 43% for letter writing, email and internet etc. 65% SMEs have purchased and developed their own Accounting software whereas the remaining 35% are using Excel etc for maintaining the accounting records. Further 95% SMEs have no information about the accounting software developed by SMEDA and only 20% of the SMEs expressed their willingness to acquire the SMEDA accounting software.

4) *Power Outages and Production time loses:* 62% of the manufacturing units did not have their own power generation mechanism and thus had to shut their operations during the power outages... [14].

4 Definition of Cloud Computing

4.1 Definition of Cloud Computing

Cloud computing is an on demand service model for IT provision, often based on virtualization and distributed computing technologies. Cloud computing can provide three kinds of service modes, including SaaS, PaaS and IaaS

1) *Software as a service (SaaS):* is software offered by a third party provider, available on demand, usually via the Internet configurable remotely. Examples include online word processing and spreadsheet tools, CRM services and web content delivery services (Sales force CRM, Google Docs, etc) [15].

2) *Platform as a service (PaaS)*: allows customers to develop new applications using APIs (application programming interfaces) deployed and configurable remotely. The platforms offered include development tools, configuration management, and deployment platforms. Examples are Microsoft Azure, Force and Google App engine.[15]

3) *Infrastructure as service (IaaS)*: provides virtual machines and other abstracted hardware and operating systems which may be controlled through a service API. Examples include Amazon Elastic Compute Cloud (EC2), Simple Storage Service (S3), Simple Queuing Service; and, in beta testing, SimpleDB, Terremark Enterprise Cloud, Windows Live Skydrive and Rackspace Cloud. [15][20]

4.2 Deployment of Cloud Models

1) *Private cloud*. The cloud infrastructure is operated solely for an organization. It may be managed by the organization or a third party and may exist on premises or off premises.

2) *Community cloud*. The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be managed by the organizations or a third party and may exist on premise or off premise.

3) *Public cloud*. The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.

4) *Hybrid cloud*. The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds) [16][17][18].

4.3 Cloudcomputing Benefits and Challenges

Table 2, shows the respond of survey, the users perception over the cloud benefits. [19]

Table 2. Cloud Computing Benefits

| | |
|------------------------------------|-------|
| Easy /Fast to deploy | 63.9% |
| Pay only for what you use | 61.5% |
| Less in-house IT staff , costs | 57.0% |
| Low monthly payment | 53.3% |
| Offer the latest functionality | 50.0% |
| Encourage more standard IT | 46.3% |
| Sharing system/information simpler | 43.4% |
| It's the way of the future | 29.1% |

% responding 4 or 5, where 1= not important, 5= very important, n=244.

A useful overview of some of the challenges businesses face in adopting cloud computing is provided in Table. 3 [19]

Table 3. Cloud Computing Challenge

| | |
|---|-------|
| Security | 74% |
| Performance | 63.1% |
| Availability | 63.1% |
| Hard to integrate In-house IT | 61.1% |
| Not enough ability to customize | 55.8% |
| Worried on demand will cost more | 55.8% |
| Bringing back in-house may be difficult | 55.4% |
| Regulatory requirement prohibit cloud | 49.2% |
| Not enough major suppliers | 44.3% |

% responding 4 or 5, where 1= not important, 5= very important, n=244

5 Strategy for Development Reverse e-Logics for SMEs in Pakistan

From the statistics, different surveys and the author's practical experience in the different fields, it seems difficult but not impossible to bring all the SMEs of Pakistan in Global Supply Chain for global competition. SMEs in Pakistan has a number of barriers to the electronic commerce, including:

- Poor infrastructure
- Lack of education,
- Lack of ICT knowledge and skill,
- Lack of awareness of usefulness of ICT technology and e-commerce,
- Perceive high cost and complexity of ICT technology.
- Lack of cooperation in the supply chain,
- On-line transaction security issues, and
- Lack of trust in E-business.

Keeping in view the economic gain, Environmental legislation, and good corporate citizenship image, Every Enterprise should include the Reverse Logistic in the Supply Chain Management Strategy. For SMEs, self-built reverse logistics system will face a higher investment risk. Out sourcing model of Reverse Logistics is best for SMEs. However the SMEs can maintain the self employed RL Model also in special cases. For different product different reverse flow model is suitable, outsourcing Model is more effective in dispose off the waste. The planers and executers can decide the appropriate reverse flow channel for product returns.

Effective information and communication technology (ICT) support is needed to manage return flows. To develop an ICT enabled Reverse Logistics for SMEs, a Cloud Computing solution is suggested. Cloud computing platform, should include supply chain management (SCM), enterprise resource planning (ERP), customer relationship management (CRM), manufacturing resource planning (MRP), enterprise assets management (EAM). Enterprise Supplier Management, enterprise performance management Inventory management, Product Life cycle Management (PLM), bar code identification , radio frequency tag (RFID) , human resource Management etc.

In Pakistan, a Public Private Partnership strategy will be fruitful to develop the cloud computing platform for SMEs, while the Government of Pakistan should make the policy, regulation and take the initiative to establish a sample cloud computing platform, as a facilitator for SMEs by involving more and more Cloud computing software vendors including traditional management software vendors, Internet companies and telecom operators, The Ministry of Science & Technology, The Ministry of industries, The Small and Medium Enterprise Development Authority (SMEDA), The SMEs BANK, The Export Promotion Bureau (EPB), and The small industry development corporations.

6 Conclusion

Due to the growing environmental concerns, economic gain, and increasing pressure of international competition the concept of 'Reverse Logistics' has gained significant attention from within the realms of researchers, operations managers and company executives.

Reverse logistics deals with the activities and processes associated with the flows of products, components and materials from users to re-users, to recover its residual value, or to proceed to the proper disposal.

SMEs sector is the backbone of Pakistan economy. Keeping in view the economic gain (Direct gain and Indirect gain), Environmental legislation, and good corporate citizenship image, every Enterprise should include the Reverse Logistic in the Supply Chain Management Strategy. The SMEs have no potential to establish ICT enabled Reverse Logistics network. Effective ICT support is needed to manage return flows. To develop an ICT enable Reverse Logistics for SMEs, Cloud Computing solution is suggested to ICT enabled the SMEs in Pakistan. The government should launch the Reverse Logistics development policies of environment protection and promote the Reverse Logistic Vision. The Government of Pakistan should take the initiative to establish a Cloud Computing Platform supporting the SME for ICT enabled supply chain Management and ICT enabled Reverse Logistics.

References

1. Rogers, D.S., Tibben-Lembke, R.S.: Going backwards: reverse logistics trends and practices. The University of Nevada, Reno, Center for Logistics Management, Executive Reverse Logistics Council (1998)
2. de Brito, M.P., Dekker, R.: A Framework for Reverse Logistics. Erasmus University Rotterdam, Netherlands (2003); Version April 2003
3. Dekker, R., Zuidwijk, R., Pappis, C.: E-commerce for Reverse Logistics. Erasmus University Rotterdam, Netherlands
4. de Brito, M.P., Dekker, R., Flapper, S.D.P.: Reverse Logistics: A Review of Case Studies, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=411649
5. Payaro, A.: The Role of ICT in Reverse Logistics: A Hypothesis of RFID Implementation to Manage the Recovery Process. University of Padova, Via Venezia 1, Padova, 35131, Italy
6. Reverse Logistics. PricewaterhouseCoopers

7. Krikke, H., Ruwaard, J.B., Van Wassenhove, L.N.: Design of closed loop supply chains: a production and return network for refrigerators
8. Fleischmann, M., Krikke, H.R., Dekker, R., Flapper, S.D.P.: A characterisation of logistics networks for product recovery
9. Fleischmann, M., Krikke, H.R., Dekker, R., Flapper, S.D.P.: A Characterisation of Logistics Networks for Product Recovery. *Omega, The International Journal of Management Science* 28-6, 653–666 (2000)
10. Wu, Q., Liu, S.: Analysis on the Reverse Logistics under E-business Environment
11. <http://www.standardchartered.com.pk>
12. SMEDA (Small and Medium Enterprise Authority), <http://www.smeda.org/>
13. <http://www.unido.org/index.php?id=5812>
14. SMEs' Survey of Gujranwala District by State Bank of Pakistan Banking Services Corporation (Bank)
15. The European Network and Information Security Agency (ENISA). Cloud Computing Security Risk Assessment, <http://www.enisa.europa.eu>
16. http://stephenslighthouse.sirsidynix.com/archives/2009/10/what_is_cloud_c.html
17. <http://csrc.nist.gov/groups/SNS/cloud-computing/index.html>
18. <http://doctrina.wordpress.com/2009/09/06/nists-definition-of-cloud-computing-what-is-cloud-computing/>
19. IDC Enterprise Panel (August 2008)
20. Hoover, J.N., Martin, R.: http://informationweek.com/1182/report_saas.htm

A Teaching Method of “In One or Two Words” in Higher Mathematics

Zhibin Li, Xiaoying Sun, and Jing Chang

Mass Transit Technology College, Dalian Jiaotong University, Dalian, Liaoning Province,
China, 116011

{lizhibin, lsxy1962}@dl.cn, changjingdl@163.com

Abstract. This paper highly generalize and refine some concepts theorems and methods in higher mathematics, after summing up, forming the teaching method of “in one or two words” in higher mathematics, generalizing four principles of this methods , summarizing three techniques of the teaching method which is called “in one or two words”, through many years of teaching experience, I find that the application of this approach can make pupils better grasp the essence of the problem, comb content , discard the dross and select the essential, seize the essence which is in the study of higher mathematics, it is easier to build the relationship among relevant contents and make a good effect in teaching.

Index Terms: Higher mathematics, Generalize, In one or two words.

1 Introduction

Higher mathematics is an important basic course of science and engineering university. Facing to the digital lists, various theorems of the boring concept in our textbooks, how to arouse students' interests, from the easy to the difficult from the outside to the inside discarding the dross and selecting the essential grasping the essence studying solidly. According to many years of teaching experience, the author find that teachers can't rely on “repeating what the two thick textbooks say” to teach their students and students can't depend on “memorizing mechanically”. In this paper, the teaching method of “in one or two words” is designed to prepare with few and precise words, extract the essence of each lesson, express the whole content by the most clear venation, which is not only convenient for teachers to teach, but also for student to understand remember and apply, this method obtain the outline portability and doubly effective of teaching effect. In according with the principle and technique of the teaching method of “in one or two words”, this paper makes the following outline of the contents of higher mathematics textbooks, we'd like to communicate with teachers.

2 Principle

The teaching method of “in one or two words” is initially generalized four principles of these methods, the examples as follows:

2.1 Seize the Essence and Excavate the Quintessence

[Example 1] Vector plane straight line: seize two elements

The two factors of vector: (1) size (module), (2) direction (direction angle).

The two factors of plane: (1) pass one point, (2) normal vector (the vector which perpendicularity to the plane)

The two factors of straight line: (1) pass one point, (2) direction vector (the vector which parallel to the straight line).

[Example 2] Green formula Gauss formula Stokes formula: closeness, partial derivative, direction.

Green formula

(1) Closeness: closed curve integral of the second kind;

(2) Partial derivative: Integrand which is $P(x, y), Q(x, y)$ has first order continuous partial derivative in the closed surface area D which is surrounded by the closed curve L (conclusion: partial derivate deviation)

(3) Direction: L is positive boundary curve.

$$\int_L Pdx + Qdy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) d\sigma \tag{1}$$

Gauss formula

(1) Closeness: closed curved surface integral of the second kind;

(2) Partial derivative: Integrand which is $P(x, y, z), Q(x, y, z), R(x, y, z)$ has first order continuous partial derivative in the closed surface area Ω which is surrounded by the closed curved surface Σ (conclusion: partial derivate sum)

(3) Direction: taking exterior of the boundary surface Σ .

$$\oiint_{\Sigma} Pdydz + Qdzdx + Rdx dy = \iiint_{\Omega} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dx dy dz \tag{2}$$

Stokes formula

(1) Closeness: closed spatial curve integral of the second kind;

(2) Partial derivative: Integrand which is $P(x, y, z), Q(x, y, z), R(x, y, z)$ has first order continuous partial derivative in the curved surface Σ which make the closed curve Γ be boundary curve (conclusion: partial derivate determinant);

(3) Direction: the exterior of curved surface Σ and the direction of curve Γ satisfy the right-handed rule.

$$\oiint_{\Gamma} P(x, y, z) dx + Q(x, y, z) dy + R(x, y, z) dz = \iint_{\Sigma} \begin{vmatrix} dydz & dx dz & dx dy \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} \tag{3}$$



2.2 Excavate the Core Content with Relevance

[Example 3] continuation: definition, existence, equality

Definition: the function $f(x)$ has definition in the point x_0 ;

Existence: the limit of the function $f(x)$ exist in the point x_0 ;

Equality: the limiting value of the function $f(x)$ in the point x_0 equal to the functional value of the function $f(x)$ in the point x_0 .

The function $f(x)$ is continuous in the point x_0 : $\lim_{x \rightarrow x_0} f(x) = f(x_0)$.

Note: if you break one of the three aspects, the function will be discontinuous.

[Example 4] monotony extremum concave-convex: boundary point, tabulation, summary.

Monotony extremum

(1) Boundary point: solving the point (stagnation point) which satisfies the derivative value the function is zero and the point (cuspidal point) which satisfies the derivative isn't existent.

(2) Tabulation: taking the boundary point and cuspidal point as the separation, the domain of definition is divided into several intervals, the first line of the table lists these intervals, finding out the central points of each interval; the second line of the table is the symbol of $y'(x)$ in these central points; the third line of the table is the monotonicity of the function y .

(3) Summary: Writing the monotonicity of the function in every intervals and the extremum of the boundary point.

Concave-convex

(1) Boundary point: the point which satisfied the second derivative is zero and the second derivative isn't existent.

(2) Tabulation: taking the boundary point as the separation, the domain of definition is divided into several intervals, the first line of the table lists these intervals, finding out the points of each interval; the second line of the table is the symbol of $y''(x)$ in these central points; the third line of the table is the convexity and concave of the function y .

(3) Summary: Writing the convexity and concave of the function in every intervals.

2.3 Initiate the Students' Interest

[Example 5] Surface integral: the first step is casting, the second step is substitution, the third step is replacement.

Surface integral of the first kind: $\iint_{\Sigma} f(x, y, z) ds$

(1) The first step is casting: Projecting the curved surface Σ to surface xOy , we can get planar domain D ;

(2) The second step is substitution: From the curved surface Σ , we can solve $z = z(x, y)$ and substitute the integrand, then we get $f(x, y, z(x, y))$;

(3) The third step is replacement: the equality $z = z(x, y)$ which has solved from the curved surface Σ is respectively solved the partial derivative about x, y , we can get $ds = \sqrt{1 + z_x^2 + z_y^2} dx dy$; then we substitute $z = z(x, y)$, $ds = \sqrt{1 + z_x^2 + z_y^2} dx dy$ into the integral.

$$\iint_{\Sigma} f(x, y, z) ds = \iint_D f(x, y, z(x, y)) \sqrt{1 + z_x^2 + z_y^2} dx dy \tag{4}$$

Surface integral of the second kind: $\iint_{\Sigma} f(x, y, z) dx dy$

(1) The first step is casting: Projecting the curved surface Σ to surface xOy , we can get planar domain D ;

(2) The second step is substitution: From the curved surface Σ , we can solve $z = z(x, y)$ and substitute the integrand, then we get $f(x, y, z(x, y))$;

(3) The third step is replacement: Exchanging $dx dy$ into $\text{sgn}(\cos \gamma) dx dy$. $\cos \gamma$ is the direction cosine of tangent plane's normal direction which is synclastic with one-sided curved surface Σ .

$$\iint_{\Sigma} f(x, y, z) dx dy = \text{sgn}(\cos \gamma) \iint_D f(x, y, z(x, y)) dx dy \tag{5}$$

2.4 Be Convenient for Memory

[Example 6] Derivative: all of three are the same; all of three aren't the same.

Importing: "the one doesn't show up, the two don't show up, the three don't show up, not show up, but show up [riddle, the answer to the riddle is forest]".

All of three are the same: " x, x_0 "

All of three aren't the same: x which is below the limit symbol approach x_0 , the numerator of quotient limit is $f(x) - f(x_0)$, the denominator is $x - x_0$.

The function $f(x)$ is derivable in the point x_0 : $f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$.

3 Techniques

Techniques of the teaching method of "in one or two words", the most commonly techniques are subsumption algorithm refining method and visualization mean, the examples as follows:

3.1 Subsumption Algorithm

[Example 7] Multiple integral: projection is the first, Getting across is the second, threshold is the third.

We can use the method of ray penetration to solve the calculation of multiple integral, that is projection is the first, Getting across is the second, threshold is the third.

To make use of rectangular coordinates to calculate double integral

(1) Projection is the first: Projecting the domain of integration to axis x (axis y), getting the interval $[a, b]$, the limit of integral of axis x (axis y) that we can see is called a (lower limit) and b (upper limit);

(2) Getting across is the second: we can get a arbitrary point which belong to the interval $[a, b]$, then we describe a ray which is perpendicular to axis x (axis y) and from the negative of axis y (axis x) to the forward. This ray penetrates the domain D , the boundary curve of the domain D that the ray firstly meets is $y = y_1(x)$ ($x = x_1(y)$), the boundary curve of the domain D that the ray finally meets is $y = y_2(x)$ ($x = x_2(y)$).

(3) Threshold is the third: the boundary curve which satisfies $y = y_1(x)$ ($x = x_1(y)$) in the below (left) of the domain is the lower limit of y (x), the boundary curve which satisfies $y = y_2(x)$ ($x = x_2(y)$) in the above (right) of the domain is the upper limit of x (y).

$$\iint_D f(x, y) d\sigma = \int_a^b dx \int_{y_1(x)}^{y_2(x)} f(x, y) dy \quad \left(\iint_D f(x, y) d\sigma = \int_a^b dy \int_{x_1(y)}^{x_2(y)} f(x, y) dx \right) \quad (6)$$

To make use of polar coordinates to calculate double integral

(1) Getting across is the first: this ray which comes from the origin penetrates the domain D , the boundary curve of the domain that the ray firstly meets is $\rho = \rho_1(\theta)$, the boundary curve of the domain that the ray finally meets is $\rho = \rho_2(\theta)$. From the ray to axis x , the minimum of rotational angle is α , the maximum of rotational angle is β . The lower limit of angle θ is α , the upper limit is β ; the lower limit of ρ is $\rho = \rho_1(\theta)$, the upper limit of ρ is $\rho = \rho_2(\theta)$;

(2) Substitution is the second: getting $x = \rho \cos \theta$, $y = \rho \sin \theta$ into the integrand;

(3) Replacement is the third: substituting $\rho d\theta d\rho$ for $d\sigma$.

$$\iint_D f(x, y) d\sigma = \int_{\alpha}^{\beta} d\theta \int_{\rho_1(\theta)}^{\rho_2(\theta)} f(\rho \cos \theta, \rho \sin \theta) \rho d\rho \quad (7)$$

To make use of rectangular coordinates to calculate triple integral

(1) The first step is casting: Projecting the domain of integration Ω to surface xOy , we can get planar domain D ;

(2) The second step is getting across: we can get a arbitrary point which belong to the planar domain D , then we describe a ray which is perpendicular to surface xOy and from the negative of axis z to the forward. This ray penetrates the spatial domain Ω , the boundary curve of the spatial domain Ω that the ray firstly meets is $z = z_1(x, y)$, the boundary curve of the spatial domain Ω that the ray finally meets is $z = z_2(x, y)$;

(3) The third step is threshold: the planar domain D regards as domain of integration of double integral, we can use the ray penetration to decide the limit of x and y , the boundary surface which satisfies $z = z_1(x, y)$ in the below of the domain Ω is the lower limit of z , the boundary surface which satisfies $z = z_2(x, y)$ in the above of the domain is the upper limit of z .

$$\iiint_{\Omega} f(x, y, z)dv = \int_{\alpha}^{\beta} d\theta \int_{\rho_1(\theta)}^{\rho_2(\theta)} \rho d\rho \int_{z_1(\rho, \theta)}^{z_2(\rho, \theta)} f(\rho \cos \theta, \rho \sin \theta, z)dz \tag{8}$$

To make use of cylindrical polar coordinates to calculate triple integral

(1) The first step is casting: projecting the domain of integration Ω to surface xOy , we can get planar domain D ;

(2) The second step is Getting across: we can get a arbitrary point which belong to the planar domain D , then we describe a ray which is perpendicular to surface xOy and from the negative of axis z to the forward. This ray penetrates the spatial domain Ω , the boundary curve of the spatial domain Ω which the ray firstly meets is $z = z_1(x, y)$ (instead of $z = z_1(\rho \cos \theta, \rho \sin \theta)$), the boundary curve of the spatial domain Ω which the ray finally meets is $z = z_2(x, y)$ (instead of $z = z_2(\rho \cos \theta, \rho \sin \theta)$);

(3) The third step is threshold: the planar domain D regards as domain of integration of double integral, we can use the ray penetration to decide the limit of θ and ρ , the boundary surface which satisfies $z = z_1(\rho \cos \theta, \rho \sin \theta)$ in the below of the domain Ω is the lower limit of z , the boundary surface which satisfies $z = z_2(\rho \cos \theta, \rho \sin \theta)$ in the above of the domain is the upper limit of z .

$$\iiint_{\Omega} f(x, y, z)dv = \int_{\alpha}^{\beta} d\theta \int_{\rho_1(\theta)}^{\rho_2(\theta)} \rho d\rho \int_{z_1(\rho, \theta)}^{z_2(\rho, \theta)} f(\rho \cos \theta, \rho \sin \theta, z)dz \tag{9}$$



To make use of spherical coordinates to calculate triple integral

(1) Getting across is the first: this ray which comes from the origin penetrates the domain Ω , the boundary curve of the spatial domain Ω which the ray firstly meets is $r = r_1(\theta, \varphi)$, the boundary curve of the spatial domain Ω which the ray finally meets is $r = r_2(\theta, \varphi)$;

(2) Substitution is the second: getting $x = r \sin \varphi \cos \theta$, $y = r \sin \varphi \sin \theta$, $z = r \cos \varphi$ into the integrand;

(3) Replacement is the third: substituting $r^2 \sin \varphi d\theta d\varphi dr$ for $dx dy dz$.

$$\iiint_{\Omega} f(x, y, z) dv = \int_{\alpha}^{\beta} d\varphi \int_{\varphi_1(\theta)}^{\varphi_2(\theta)} \sin \varphi d\theta \int_{r_1(\theta, \varphi)}^{r_2(\theta, \varphi)} f(r \sin \varphi \cos \theta, r \sin \varphi \sin \theta, r \cos \varphi) r^2 dr \quad (10)$$

[Example 8] Curve integral: parameter is the first, processing is the second, substitution is the third.

Whether calculation of curve integral of the first kind (the curve integral of arc length) or calculation of curve integral of the second kind (the curve integral of coordinate), variant should be parameterized, when the curve integral of the first kind translates into definite integral, the lower limit is small, the lower limit of the curve integral of the second kind is initial point. That is “variant is parameterized, small and the lower limit is initial point”. Concrete procedures are “curve integral: parameter is the first, processing is the second, substitution is the third”.

The curve integral of the first kind: $\int_L f(x, y) ds$

(1) Parameter is the first: $L: x = x(t), y = y(t)$;

(2) Processing is the second: $ds = \sqrt{[x'(t)]^2 + [y'(t)]^2} dt$;

(3) Substitution is the third: getting $x = x(t), y = y(t)$ and ds into the integral.

$$\int_L f(x, y) ds = \int_{\alpha}^{\beta} f(x(t), y(t)) \sqrt{[x'(t)]^2 + [y'(t)]^2} dt \quad (11)$$

Note: $\alpha < \beta$.

The curve integral of the second kind: $\int_L P(x, y) dx + Q(x, y) dy$

(1) Parameter is the first: $L: x = x(t), y = y(t)$;

(2) Processing is the second: $dx = x'(t) dt, dy = y'(t) dt$;

(3) Substitution is the third: getting $x = x(t), y = y(t)$ and $dx = x'(t) dt, dy = y'(t) dt$ into the integral.

$$\int_L P(x, y) dx + Q(x, y) dy = \int_{\alpha}^{\beta} [P(x(t), y(t)) x'(t) + Q(x(t), y(t)) y'(t)] dt \quad (12)$$

note: α is an initial point.

3.2 The Visual Method

[Example 9] Taylor formula: derivation, algebras, formula.

- (1) Derivation: $f', f'', \dots, f^{(n+1)} f(x_0)$;
- (2) Algebras: $f(x_0), f'(x_0), f''(x_0), \dots, f^{(n)}(x_0)$;
- (3) Formula:

$$f(x) = \frac{f^{(0)}(x_0)(x-x_0)^0}{0!} + \frac{f'(x_0)(x-x_0)^1}{1!} + \frac{f''(x_0)(x-x_0)^2}{2!} + \dots + \frac{f^{(n)}(x_0)(x-x_0)^n}{n!} + \dots + \frac{f^{(n+1)}(\xi)(x-x_0)^{n+1}}{(n+1)!}$$

$$= f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)(x-x_0)^2}{2!} + \dots + \frac{f^{(n)}(x_0)(x-x_0)^n}{n!} + \dots + \frac{f^{(n+1)}(\xi)(x-x_0)^{n+1}}{(n+1)!}$$

We can find every item of the formula id the same, it means that order of derivative degree of power and number of factorial are equal.

3.3 The Essential Method

[Example 10] L' Hospital rule: type, derivation, existence.

- (1) Type: solving the undefined limit which satisfies the fraction $\frac{0}{0}$ or $\frac{\infty}{\infty}$;
- (2) Derivation: the derivative of the numerator function and denominator function are existent in the delete neighbourhood of x_0 ;
- (3) Existence: we can solve the derivative of the numerator function and denominator function, then we can solve the quotient of this formula, finally, we get the limit of the last formula is existent; this limit equals the primary limit.

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} \tag{13}$$

Reference

1. College of Mathematics in Tongji University, Higher Mathematics, 6th edn. Higher Education Publishing House, Beijing (2007)



Ponder on a New Internet Community Space Pattern with Ubiquitous-Network Service

Zhao Hongyu and Chang Sheng

¹ Department of Urban and Civil Engineering, Harbin Institute of Technology Shenzhen
Graduate School, Shenzhen, P.R. China

² Urban Planning & Design Department, Overseas Chinese Town Enterprises Co. Shenzhen,
P.R. China

marthew77@gmail.com, changsheng@chinaoct.com

Abstract. In recent years, the Internet has become the inevitable characteristic of modern urban life. Network carries including communication, learning, business, shopping, entertainment, leisure, media publicity and other city functions. Because of the abundant information, updates quickly, timeliness, and cost-effective, the internet exactly meet the ever-accelerating pace of modern life. But on the other hand, the internet as a life helper has drawbacks in the value recognition of life, trust, and security, and it has caused lots of social problems such as social isolation, urban sub-health, loss of sense of belonging and so on. The urban design in the Internet age calls for a new community space pattern, which is able to satisfy the demands of efficient life, and at the same time access to experience the real human nature. From the perspective of urban designers, we study the life requirements and behavior characteristics of different Internet groups, bring forward some space patterns which are suitable for the internet age, such as showcase-type space, watching-type space, participatory-type space, passing-type space and destination-type space, try to realize the positive combination of real city life and virtual online world, and provide new ideas for urban construction and management in the Internet era.

Keywords: Internet community, lifecycle, space pattern.

1 Introduction

In recent years, the rapid development of Internet technology brings out the magnificent growth of Internet users in China, and the Internet becomes an important part of life in modern city. The urban functions of city life based on modern features such as communication, learning, business, shopping, entertainment, leisure, media publicity and so on has been transferred into Internet inevitably, while Internet has abundant information, updates quickly, timeliness, and cost-effective. its convenient and fast degree exactly meet the ever-accelerating pace of modern life, but the Internet as a life helper has drawbacks in the value recognition of life, trust, and security, and it has caused lots of social problems

such as social isolation, urban sub-health, loss of sense of belonging and so on. The urban design of Internet era calls for a new lifecycle mode of health and sustainability urgently, which is able to satisfy the demands of efficient life, and at the same time access to experience the real human nature, in order to integrate the fast and low rhythm of life, and finally finds the realization approach of new life, care life, health life, beauty life, and enjoy life.

2 Dependency and Characteristics of Internet in China

2.1 Current Level of Network Development

According to "Article 23 of China Internet Development Statistics Report" which was made by the China Internet Network Information Center (CNNIC), at the end of 2008, the scale of Chinese netizen has reached 298 million people, representing 41.9% growth in 2007, the Internet penetration rate reached 22.6%, slightly higher than global average [1].

In the Internet age structure of population, 10-19-year-old Internet users in China in 2008 has become the largest user group. Meanwhile, the proportion of Internet users in the age group increasing the growth rate has exceeded the overall growth rate.

In the academic structure, and in 2007 compared to users in the tertiary and higher education further population has declined, high school, junior high school education continue to enhance the proportion of the Internet increasingly popular to the less educated population.

In the occupational structure, Internet users constitute the largest group is the students, followed by the workers party and government organs and institutions, enterprises managers, staff, professional and technical personnel and other civilian personnel, and accounting for the largest share of China's population of farmers, industry service workers in the proportion of Internet users is relatively low.

2.2 Mass Dependence on the Network

According to China Internet Network Information Center (CNNIC) statistics, Internet applications can be divided into the following categories, namely: Internet media, Internet information retrieval, network communications, network community, network entertainment, electronic commerce, network financial. On this basis, according to the different behavior of Internet users in network applications, it can be sub-divided into several groups, namely: network dependent group, network business group, social networking group, network tasted group, basic application group, self-display group, non-mainstream gaming group. The scale of different group is as shown in Fig. 1.

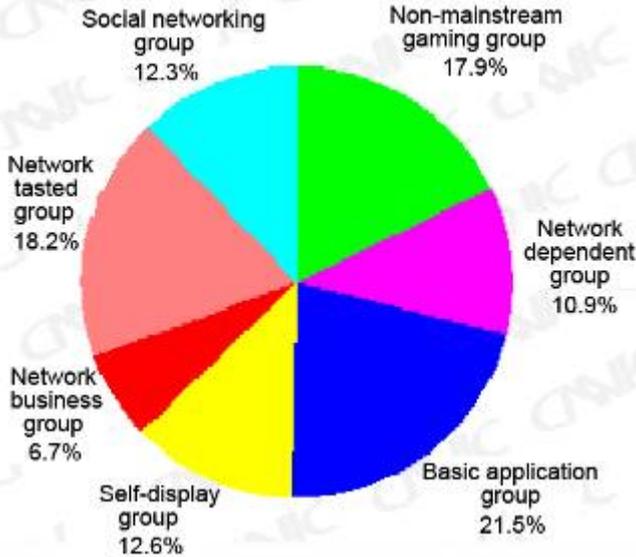


Fig. 1. The scale of different group

3 Social Problems Caused by the Cinternet Dependency

With the rapid development of Internet technology, the network as an interactive media and information channels are incorporated into every aspect of people's lives, however, live in the network to bring convenience and fast, but also to the growing number of social problems, such as: social isolation, lack of face to face, loss of collective sense, the body sub-health and so on.

3.1 Social Isolation

There are new things everywhere on the Internet, and constantly increasing. These new things have tremendous appeal to people's curiosity, which will lead the network to attract the indulgence. The teenager, which occupies the largest number of Internet users, is the major groups of Internet indulgence. Because their mental quality is not strong and self-control is relatively weak, they becomes the multiple group of network mental disorders. The patients often consider the online world as in real life and have no common language with others, therefore appear lonely and depressed, sluggish thinking, reduced self-evaluation and other symptoms, serious even suicidal thoughts and behaviour.

3.2 Lack of Face to Face

Internet can send instant text, voice, image, and can provide multimedia and interactive way for interpersonal communication, through these channels people can confide and help others even when they don't know each other. However, this

socialization is only a virtual socialization, there have machine barriers between this kind of communication. It is just a "human - machine - symbol - symbol - machine - human" form of communication. It has "go socialization" features, and far different from real social situations. Too much online virtual communication not only allow people to vent their emotions, but also make them tend to deep buried their true feelings, do not want to communicate with real people.

3.3 Loss of Collective Sense

In the Internet, everyone can make himself as the centre, use the internet resources, play different roles, design "small productions" totally according to one's personal will. This unfettered and arbitrary performance despite has advantages for showing personality in some extent, but also causes some people completely lost the collective sense because of their extreme independent will. For example: some young people always want to make blockbuster performance in the internet just to show their personality. In order to let the world know they are the main role, they will do various bad behaviours, such as damage other data and sites, invade other systems, and even make terrible network earthquake to block network traffic.

3.4 The Body Sub-health

The internet has brought to people not only spiritual invisible injury, but also physical health damage which can not be ignored. As most of the daily affairs can be resolved through the internet, the time for outdoor activities is far less. While sitting in front of the computer for long can easily lead to various kinds of unpredictable physical harm. Related survey data shows that in recent years, the disease spectrum of Chinese residents has changed significantly, the morbidity and mortality of chronic non-communicable diseases has a rapid increase. In 2004, the number of patients with hypertension is about 160 million, patients with diabetes is about 20 million, and the number of overweight and obesity is about 200 million and 60 million. The lack of physical activity is one of the main reasons. Therefore, increasing communications and outdoor activities is an effective way to solve the many problems above.

4 The Space Strategy of "Contact and Experience"

Regular informal meet is the starting point which can further develop into close relationships. Regular daily meet can promote communications between neighbours, so that people have the opportunity to establish and maintain friendships in a relaxed atmosphere. Therefore, in the internet age, the community, which is most closely with daily life, should combine the life characteristics of different network groups, create different types of space form with network elements to increase the opportunities for regular contact and communication.

4.1 Showcase-Type Space

To the network business group, the ability of quickly access a wealth of information and make comparison between the information is particularly important. In order to

facilitate fast browsing and comparison, besides vast amounts of network information can be retrieved in time, but also need adequate display terminals to release and simultaneous presentation. The interior space due to the size and cost limitations, can not set too many display terminals. In contrast, outdoor space has enough open space as well as building facade which can be used as the display interface, so in terms of number of terminals can be set or view from space area for people to stay, the outdoor space have absolute advantage than the interior space.

In order to increase the visual interface in outdoor space, in the design of urban settlements, which is most closely to daily life, should adopt the idea of “reduce the building height and increase building density.” Because the layout of low-height and high-density can effectively increase the visual interface area in walking scale, therefore provides the possibility for changing more building facades into information display terminals. The more display terminals have, the more outdoor space can play the advantages than interior space, meanwhile more likely to create opportunities to stop people and let them viewing. Therefore create the possibility for short talking and other activities in the shared browsing and comparison of information, through this can realize the promotion of contacts and communications among neighbours (Fig. 2).

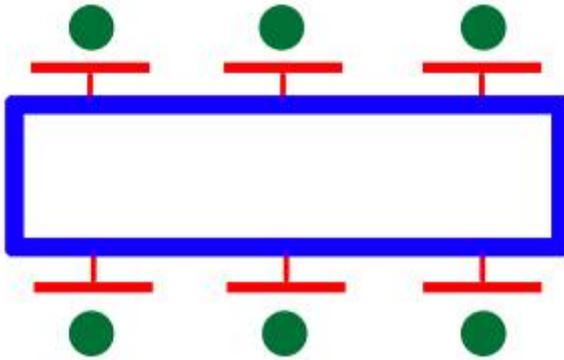


Fig. 2. The space pattern of showcase-type space

4.2 Watching-Type Space

To the basic application group and social network group, outdoor space and indoor space compared to the greatest advantage, is the ability to provide adequate non-private space for people not familiar to watch online media together and talk to each other. People can use the large outdoor display terminal to watch online video, to discuss program, to hold video conference and so on. The open outdoor space and good environmental design can create a relaxed atmosphere for conversation, making communication more likely to arise. Meanwhile, the informal watching and chat groups can attract the attention of passing pedestrians, making them join the group, resulting in a wider range of groups in the exchange.

Affinity watching space should be located in a relatively quiet space in non-traffic zones, with a large network display terminal fastened in a small structure in one side, the other side through environmental design to create a soft enclosure interface. Large

building should not be set around the watching space, because the excessive building scale can make people have a feeling of oppression and tension, thus unwilling to carry out activities in such environments. The spatial characteristics of watching-type space should be relatively enclosed semi-private space, this on one hand is to provide a suitable spatial field sense for stop and long talking, on the other is to make the space has good visibility, thus easy to be found by other pedestrians and let them join the group communion (Fig. 3).

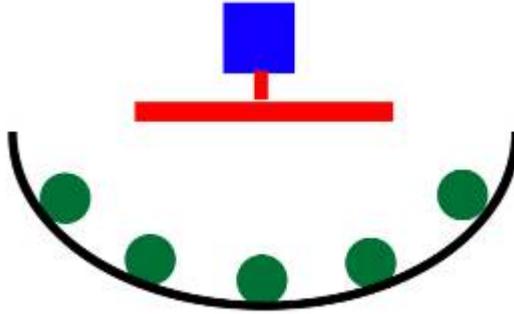


Fig. 3. The space pattern of watching-type space

4.3 Participatory-Type Space

For the elderly users in non-mainstream gaming group and the network tasted group, easy puzzle games is their most widely used network service. Since most of these games require people to participate or have a competitive nature, therefore play with familiar people would greatly increase the fun of the game, this aspect sometimes even become a prerequisite for whether people will participate. Therefore, the outdoor space, compared with the interior space can provide more chance to meet this prerequisite, and thus more likely to promote the game behavior. Once the game begins in the public space, it can attract more pedestrians to the onlookers, therefore achieve a wider range of contacts and exchanges.

Comfortable participatory space should be located at a relatively quiet area of non-traffic zone, the center should set up by a small structure for the game's network display terminal, surrounding the terminal should be provided some seats or such facilities for the participants to stay long, while the external boundary should through environment design to create a soft enclosed interface. The spatial characteristics of participatory space should be a semi-public space with definite boundary, this on one hand is to give participants a relatively static and stable region, on the other hand also give the onlookers a site with definite space field where they can stop and gather. Meanwhile, the participatory space also should have good visibility, easy to find and reach (Fig. 4).

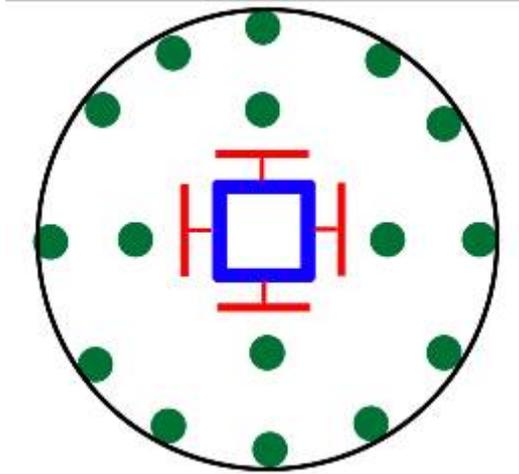


Fig. 4. The space pattern of participatory-type space

4.4 Passing-Type Space

To the network dependent group which is highly rely on network, the most important need is to access to any kind of network service at anytime and anywhere. Besides using internet in spare times, they are even eager to use network on rail or other transportations conveniently. Therefore, if it can be installed enough internet terminals outside, not only the group who are afraid of abrupt network using requirement, but also those who need intermittence internet operations, would be attracted. As the time they have to spend on waiting can be efficiently utilized as transportation time. This also increases the will and possibility to act outside for this group indirectly.

As efficient passing-type space mainly connects with path installment, it should be provided some small rest area for temporary stay to avoid inter-disturb between traffic flow and rest group. This type of small rest stay can be settled on pavement, or also parking place for driveway by widening local road sections. Moreover, it also should be settled some small buildings to connect network vision terminal or to provide network interface in the rest area. The facilities should be as miniaturization as possible, mainly providing efficient internet demands and operating functions. Meanwhile, due to these kinds of small rest-areas shall be adequate in quantities to be functional, thus the layout would shape in a scraggly rhythm, therefore also increase the delightful and attractive of the outdoor space in some ways (Fig. 5).

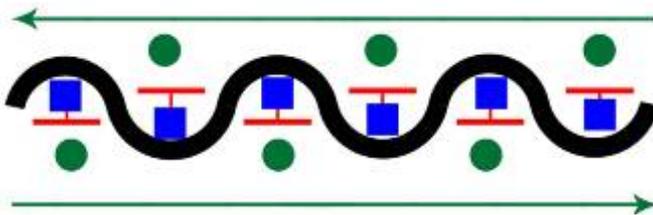


Fig. 5. The space pattern of passing-type space

4.5 Destination-Type Space

The most notable characteristic of destination-type space is that, it can not only provide widest network service for all kinds of groups, but also can provide complicated network technique used in large buildings or in special place, such as virtual cities, network sports, large performance, network parties and so on. As the particularity and non-substitutability of network service provided by destination-type space, in some sense, it has great non-competitive superiority. As the site with destination property, the destination-type space should also be connected with traditional urban functions to increase attraction as possible as it can. Then all kinds of people would like to come and use the unique network service and actualize contact and communication in common use and experience.

Destination-type space should be located in large complex in community center or outdoors sites, and can be easily arrived by varies of transportations. The core region should be settled with rail station or other transportation stations, and inside these stations should be settled with enough network information inspection and service termination. Therefore people not only can achieve the detailed information of surrounding buildings and public space, but also can actualize communication and network mutual service with correlative people as soon as possible. Meanwhile, it should be located underground or over ground all-weather channel, in order to make it convenient to move into surrounding buildings or public space (Fig. 6).

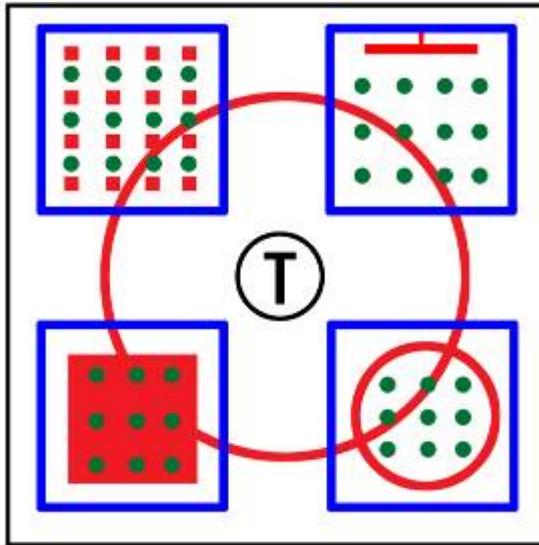


Fig. 6. The space pattern of destination-type space

5 Conclusion

On the basis of above, through adding the internet elements into the traditional urban public space, the new space pattern suitable for the internet age can make the internet

service into all aspects of the city lives and outdoor activities. This will bring the convenience and efficiency characteristics which the network groups require into the physical environment in the real world. Since this kind of outdoor-space-supplied internet service has its intrinsic characteristics, such as naturality, openness, and publicity, it has the unrivalled superiority over the self-internet. And thus, the internet-integrated outdoor environment will be made incomparable attractive. Thereby, people will prefer to go outdoors, live a healthy life and return to the nature of human in the process of contact and experience.

References

1. China Internet Network Information Center, Statistical reports for the Internet development in China, CNNIC 11, 17–18 (January 2009)
2. Maeshall, S.: *Streets & Patterns*. Spon Press, New York (2005)
3. Pu, M.: *Cancer of Urban Life: Problems of Gated Communities in China and Their Solutions*. Time+ Architecture (May 2004)
4. Van Der Burg, A., Dieleman, F.: ‘Dutch Urbanisation Policies: from Compact City’ to ‘Urban Network’. *Tijdschrift voor Economische en Sociale Geografie* 95(1), 108–116 (2004)
5. Taylor, P.J.: Specification of the world city network. *Geographical Analysis* 33, 181–194 (2001a)
6. Sassen, S.: *The Global City*. Princeton University Press, Princeton (2001)

HAZOP Analysis-Based Method on the Risk Assessment of the Main Engine of the 10000TEU Container Ship

Yulong Zhan, Yitong Hao, Yao Yu, and Di Wu

Department of marine engineering
Shanghai Maritime University
Shanghai, China
ylzhan@shmtu.edu.cn

Abstract. HAZOP analysis, developed by the British Imperial Chemical Engineering (ICI) in the 1960s, is a approach for risk analysis and risk assessment [1]. Through continuous improvement and development, it has achieved great success in the risk analysis of the petrochemical industry. However, it is yet to be applied to the ship's risk assessment. This paper will explore the use of HZAOP analysis for the risk assessment of the main engine of the 10000TEU Container Ship, and choose the fuel system of the main engine as the object for the research. Through the HAZOP analysis we can establish the risk of pre-control guidance system for the relevant managers to provide the necessary security and technical support. And to enhance the safe operation level of the 10000TEU container ships.

Keywords: 10000TEU container ships, the Main Engine Fuel System, Risk assessment, HAZOP analysis.

1 Introduction

Since 2007, several 10000TEU container Ships of COSCO have been put into operation. The 10000TEU container ships are one of the large stand the most advanced container ships in the world, which integrate a number of high technology. Its high degree of automation and equipment of ME-type diesel engine, which can make them meet the request of one crew driving and engine room zero people, especially, they are environment-friendly vessels. At the same time, the risks that existing in this kind of vessels are more than that in normal containers vessels due to its poor flexibility and large inertia. It could result in huge economic damages and losses in the event of any security faults occurring in such a huge ship. The main engine seemed as the “heart” of the ship, is the most important part of the 10000TEU container Ship. Therefore, the pre-risk assessment and risk control of it in advance could have a great practical significance to keep the safety of the ship and to reduce economic damages and losses.

We conduct risk assessment on the main engine of the 10000TEU container Ships with the application of HAZOP analysis technique, generating the risk analysis

sub-system of the main engine. In the framework, identify the inherent risks in each unit, and then conduct risk assessment to get risk classification (high-risk, normal-risk and low-risk). Listing the main high-risk items, then use Hazards and Operability analysis (HAZOP) to develop programs to conduct risk pre-control.

2 Hazop Analysis Technigue

HAZOP analysis is a kind of assessment method that can be used to identify design defects, process hazards and operability problems, which are based on an engineering system that can be used for qualitative analysis or quantitative evaluation. The analysis process is to divide the group process into a reasonable analysis node (or process unit) according to certain principles by professionals, and then to identify those potential dangerous deviations, which can also be used for determining the risks in production equipment and processes to seek the necessary measures. By analyzing the diversification of parameters in the process state and the possible biases in the operation control. Next, we must identify the reasons about those changes and deviation to find out the major risks in devices or systems, and the production process. Finally, measures should be taken to correct the changes and deviation. This is the core of HAZOP analysis technique [2].

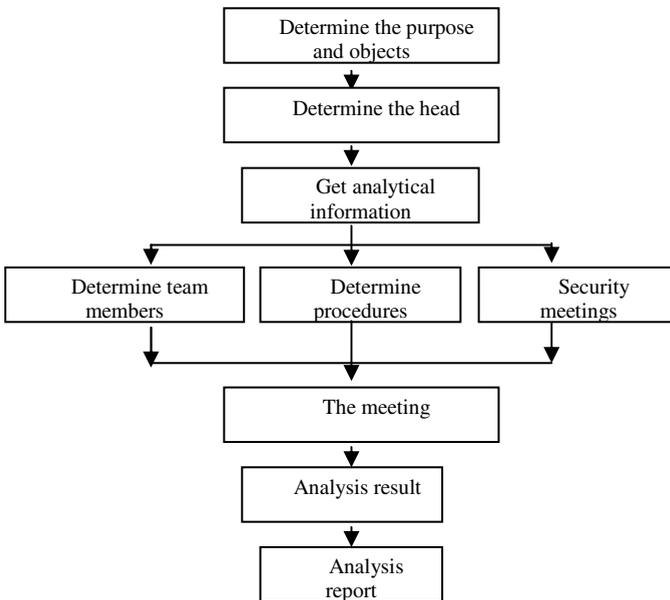


Fig. 1. Analysis of the entire process

HAZOP analysis is in the form of a series of meetings by a professional group, which analyze the problems in the device process and operational risk. The flowchart of the preparatory work and entire process are demonstrated in Fig. 1.

A. *The nodes and bias of the HAZOP analysis*

HAZOP analysis is usually carried out at the design stage, which is to determine the hazards and operational problems that exist in engineering and to take comprehensive measures to reduce the accidents caused by ill-considered. The analysis of the special points called "the node" in the process or operation. By the way we can identify those potential dangerous deviations [3], which are led by guiding words (also known as keywords). One purpose of the use of guiding words is to ensure that analyze all the deviation parameters of the process and then analyze their possible causes, consequences, protective measures and so on. If the protection measures do not meet to the existing security requirement, that is, to put forward feasible measures proposed, and to assess risk rating by the use of risk assessment matrix.

1) *The division of analysis nodes*

Generally, the division of analysis nodes is according to the P&ID diagrams and process, starting from the pipeline from the P&ID diagrams. Considering the following factors as the basis for classification of different nodes:

- The change of the design intent;
- The significant changes in process conditions;
- The change of equipment types, Etc.

In addition, the division of the node should also consider the purpose of HAZOP analysis, a reasonable boundary / cut-off point, the consistency of classification methods and other factors. In short, the principle of division is not rigid, division of the different objects should be flexible in using the principles of dividing node.

2) *Determine deviation*

Deviation is the combination of words and process parameters, generally stated as follows:

Guiding words + Process parameters = Deviations

The guiding word is a kind of simple word which is used for the qualitative or quantitative analysis of the process index and to guide and identify the risk of the process. Such as excess, no, reverse, Etc. Process parameters is the physical and chemical properties related to the normal operation of the process. Such as: temperature, pressure, viscosity, flow, etc.

The advantage of the method could ensure the uniformity of HAZOP analysis, while the problem can be analyzed systematically. With a complete setting of guiding words, you can export all the possible bias, without being missed. The commonly used bias can be seen in Table 1.

B. Risk rating matrix of HAZOP analysis

Hazard Analysis is a method considering the seriousness of each type of failure mode and the influence produced by the probability of each type of failure mode comprehensively. According to the results of hazard analysis, you can find the defects and weaknesses of the device, and develop and implement various improvements and controlling measures to improve its security level.

Table 1. The Commonly Used BIAS

| Guiding words | Process parameters | | | |
|---------------|-------------------------------|-------------------------|--|-----------------------|
| | Flow rate | Temperature | Pressure | Liquid level |
| NO | No flow | * | * | No level |
| Excessive | Excessive flow | Temperature is too high | The pressure is too high | The level is too high |
| Reduction | Insufficient flow | Temperature is too low | The pressure is too low | The level is too high |
| Accompany | Contaminated fluid | * | * | * |
| Part | Reduce the fluid mass content | * | * | * |
| Reverse | Adverse current | * | * | * |
| Abnormal | Sampling | * | Pressure relief or pressure fluctuations | * |

Notes: * means that there is no meaningful deviation

1) The classification of HAZOP fatal degree

The purpose of the classification of fatal degree [4] of the HAZOP is to classify the product and the fault mode according to the comprehensive influence produced by the seriousness and the probability of the fault mode. In this way, we use it to evaluate the every possible failure of the system comprehensively. The seriousness of the final influence of the system product is called as the fatal degree. When we analyze the influence of the fault mode, we need to classify the product's fatal degree to be the HAZOP fatal grade. It divided the degree based on the product's worst potential effect on the system. Generally, it can be divided into four grades which are shown in Table 2.

Table 2. HAZOP Fatal grade

| Fatal Grade | Description |
|----------------|---|
| A . Disastrous | It can cause death and the destruction of the system. |
| B . Fatal | It would cause significant harm to persons, major economic loss or lead to serious damage to the failure of system function. |
| C . Critical | It would cause minor injury to personnel, certain tasks of economic loss or lead to a slight delay or degradation of system function. |
| D . Mild | It cannot result in personal injury, economic loss or some damage in the fault system, but it will lead to unplanned maintenance. |

2) *The classification of HAZOP risk level and evaluation index*

The classification of HAZOP risk level and evaluation index is shown in Table 3.

Table 3. Risk level classification

| The possibility of failure mode occurs rating | Severity level | | | |
|---|-----------------|--------------|-------------|------------|
| | A Disastrous | B Serious | C Medium | D Minor |
| A frequent | 1 | 3 | 7 | 13 |
| B likely | 2 | 5 | 9 | 16 |
| C sometimes | 4 | 6 | 11 | 18 |
| D seldom | 8 | 10 | 14 | 19 |
| E Impossible | 12 | 15 | 17 | 20 |

Table 4. Assessment indices

| Risk Assessment index | Risk grade | Evaluation criteria |
|-----------------------|------------|--------------------------------------|
| 1~5 | High | Not accept |
| 6~9 | Serious | Do not want (generally not accepted) |
| 10~17 | medium | Acceptable (but subject to review) |
| 18~20 | low | Acceptable |

The definition of risk assessment indices is shown in Table 4.

Table 5. The HAZOP analysis log of the fuel system

| Bias | Reasons | Consequences | Counter-measures | Possibility/Severity | Risk level |
|--------------------------------|---|--|--|----------------------|------------|
| Insufficient fuel oil supply | 1.Pipe fouling | 1.Pump wear intensify; 2.Power shortage; 3.Lost power when serious | 1.Clean or replace the filters; 2.Dredge pipes; 3. Clean oil tanks. | A/C | 7 |
| | 2.Spilt pump fault | No influence temporarily | 1.use the hand fuel pump; 2.repair the spilt pump | B/D | 16 |
| Poor quality of purification | 1.Filter fault | 1.increased the abrasion of the fuel injection equipment; 2.Poor combustion; Damage to the combustion chamber components; 3. Power shortage, economy down. | 1.Exchange the double-filter for use; 2.Clean the double- filter; | B/C | 9 |
| | 2.Insufficient time to Release residual | 1.increased the abrasion of the fuel injection equipment; 2. Lose power | 1.Enhance release residual; 2.Lengthen setting time. | C/B | 6 |
| Poor quality of atomization | 1.High-pressure oil pump fault; | 1.Incomplete combustion 2.Damage to the combustion chamber components; 3. Economic decline | 1.Suspend fuel oil offered to the fault cylinder; 2.Replace the plunger and the sleeve; 3. Check out the outlet valve parts. | C/B | 6 |
| | 2.High-pressure pipeline failure; | 1.Cause fire in the Cabin for the Rupture of high-pressure pipe; 2. Damage to the combustion chamber components. | 1.Suspend fuel oil offered to the fault cylinder; 2. Replace tubing and gasket when need. | D/A | 8 |
| | 3. Injector fault. | Insufficient power or lose power | 1.Suspend fuel oil offered to the fault cylinder, 2. Grind or replace the fault injector timely | A/C | 9 |
| Fast Cut-off valve Abnormality | 1.Cannot close | Cannot be closed in emergency situations | Repair in time | D/B | 10 |
| | 2.Cannot open | No power output | Rush repair | D/A | 8 |

3 The Example of the Application of Hazop Analysis to Assess Risk on the 10000 Teu Container Ship Main Engine

A. The risk analysis sub-system framework of the main engine

It has been mentioned in the above that with the application of HAZOP analysis technique, in the subject of the main engine, generating the risk analysis sub-system framework of the main engine. And identify the inherent risks in each unit, and conduct risk assessment.

The risk analysis framework of the main engine can be divided into six sub-systems [5], (see in Fig. 2).



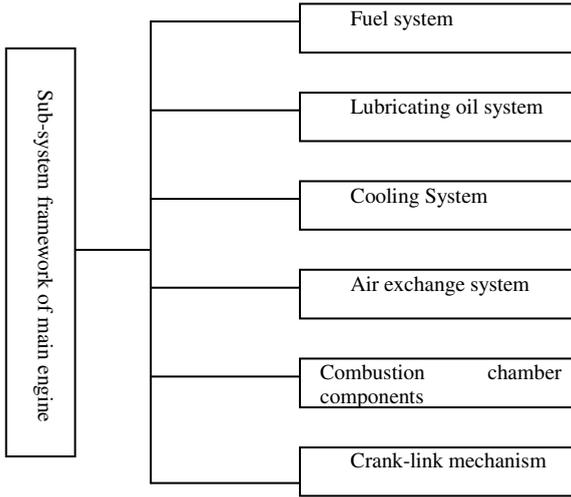


Fig. 2. The sub-system framework of the main engine

B. The example of the application of HAZOP analysis to assess the fuel system

The function of the fuel system is to supply the main engine with the fuel uninterruptedly which can meet the requirements. The fuel system is among the most important systems of the main engine. Major equipment: fuel storage tank, filter, fuel lightening pumps, daily fuel tank, quick closing valve, machine with fuel pumps, fuel pump, high pressure tubing, injectors and so on. The following, we will use the fuel system HAZOP analysis to analyze the fuel system, to identify deviations, and adopt the better response measures in order to maintain the normal operation of the main engine and increase the operating efficiency of the main engine[6].

According to the above-mentioned method of determining the HAZOP deviations, we consider the fuel system in the following several deviations: insufficient fuel oil supply, poor quality of purification, poor quality of atomization and fast cut-off valve cannot activate properly. Through the measurement and evaluation of fuel system, we developed the HAZOP analysis log of the fuel system, shown in Table 5.

In Table 5 we can see the potential risk of the fuel system, also its severity and risk level. To our delight, the measures that used to solve the corresponding risk are given in the table too. From the Table 5, we can get that with the application of HAZOP analysis method we can establish the risk of pre-control guidance system of the 10000 TEU container ship main engine system, even the risk of pre-control guidance system of the 10000 TEU container ship.

4 Conclusions

In this paper, we have used the HAZOP analysis method to analyze the fuel system of the 10000TEU containership's main engine effectively. We did analyze on the reason, the consequence and the measure of node bias. The method can be used to do analysis

on ship's Risk source effectively. Today, Ships' quantity and the tonnage increase gradually. Risk assessment is urgent to large ships, especially to the 10000TEU container ships and it's systems. Using HAZOP analysis method on the risk assessment of the 10000TEU container ships and it's sub-system create a new way for further risk assessment of ships. Although there are some problems that need solved for the use of it on the risk assessment of ships and need further study. I believe it will be extensively used for Risk assessment of ships.

Acknowledgment. The authors wish to thank scientific research fund 09-22 of Shanghai Maritime University.

References

1. Zhang, B., Zhao, D.: Reserch on Improvement of HAZOP Analysis. China Safty Science 17(10), 89–93 (2007)
2. Zhao, W.: The Core of HAZOP Analysis. Science Press, Beijing (2008)
3. Wells, G.: Hazard Identification and Risk Assessment.Ieheme.UK (1996)
4. Guimarães, A.C.F.: Hazard and operability study using approximate reasoning in light-water reactors passive systems. Nuclear Engineering and Design 11(236), 1256–1263 (2006)
5. Jian, L.: The Introduction of the 10000TEU Container Ships. Shanghai Maritime University, Shanghai (2008)
6. Huang, S.: Marine Diesel Engine. Dalian Maritime University Press, Dalian (2006)

'Peer to Peer' Network Graphics Art Design

Lv Jiefeng and Ge Xiazhi

Department of Art and Design, Wuhan University of Technology, Wuhan, Hubei Province,
China

freeant@tom.com, gexiazhi@126.com

Abstract. This article introduced the current phenomenon of cultural transmission in the Internet, then analyzed the reasons for this phenomenon and its transmission characteristics. Peer to peer technology enriched the network graphic art design, but also brought out some drawbacks, the paper insisted to improve the quality of network graphic art design by means of culture, technical approach, law and moral regulation, and at last a new direction for the development of network graphic design was brought out, that is, the combination of new design forms.

Keywords: Art design, 'peer to peer' network graphic, creativity, communication.

1 Introduction

Peer-to-peer (P2P) originally referred to a new technology of network, known as Peer internet technology. P2P technology allows users to directly connect to others' computers for file sharing and exchange, in addition, it takes effect in deep search, distributed computing and collaborative work. It is characterized by weakening or removing the links between servers and intermediaries in traditional network communication, this new technology relies on participants' computing power and bandwidth in the network to achieve the interaction between terminals, thus makes it more easy, direct and interactive to communicate.

P2P net communication mode is based on IM (Instant Messenger) software, if the 'peer to face' network communication mode based on website. When the server becomes 'transparent' access in the net communication mode which based on the IM software, users can be the publisher and recipient at the same time, because they connect with each other 'one on one'.

The development of network art design opens a new chapter with this platform. A richer and more interesting creative approach will emerge. It not only reflects more novel and multiple features, but also brings us more and more in-depth thinking.

2 Creative Techniques in ‘Peer-to-Peer’ Web Graphic Art Designed

2.1 Change ‘Characters’ into ‘Graphics’

The information dissemination on internet of ‘peer to peer’ was initially carried out in text form, it focus on content instead of form. Some ‘Nosy Parkers’ found that using ASCII characters as Modeling elements can be combined into a variety of vivid, colorful, interesting and emotional graphics, such as : ^ _ ^ (happy), -_-!(sweat), O (^ o ^) O (cheers), (T_T) (tears). oΓ_ looks like a frustrated man kneels on the ground, to show a frustrated or depressed mood, this kiind of expressions and actions began to become popular symbols of the network communication.

Thus, serving the basic character as the elements of graphics is becoming one main creative technique in P2P network graphic design, and it spreaded rapidly through P2P instant messenger. In the process, this technique is constantly enriched and re-created, the simple expressions and actions have changed into more complex and concrete images, and now they are still used and disseminated. (Figure1).

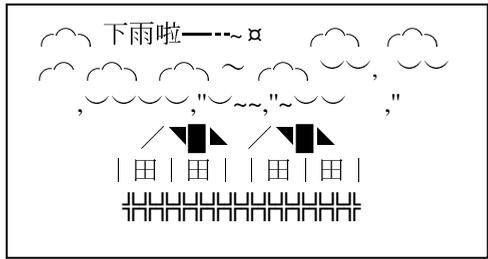


Fig. 1.

As a creative technique of network graphic art and design, changing the characters into graphics is to ignore the text meaning of character, and changes the text symbols into graphic ones, according to the characters’ features of visual form, it is more vivid and easier to resonate to form the basic elements of characters and change them into abstract graphics than to use the characters only. In addition, the character graphics are used widely, and they always have special meaning and are interesting to use.

2.2 Change ‘Graphics ’ into ‘Characters’

Because of the user's emotional appeal, the expression of the network communication becomes more colorful, on the assumption of technological development. In China, most IM users communicate by sending and receiving a simple icon instead of a large passage in Tencent QQ, the icons are simple, but full of a lot of informations and can be interpreted into a variety of meanings according to different contexts, such as 🍀 (I'm so unlucky!), 🏃 (Make great efforts), 💋 (Give a kiss). In addition to these official expressions, users can also design their own graphic symbols, or share more in the internet.



Regarding these readable graphics as the characters and coherencing them, the graphics can be formed into a quite interesting 'text' which has visual illustrations and rich colors.

This kind of 'story narration' is particularly favored by communicators in the internet, though it seems nascent, first, it is simple to use, for you just need a click to select the expression instead of inputing lots of characters; second, the 'story narration' is more intuitive, vivid and interesting, just like the character graphics; what's more, communicating by graphics is a process of 'empathy'. And in this process, the communicators resonate.

2.3 With Sound and Video

Multimedia in digital environment, can be applied in instant messenger in the network, which not only includes texts and images, but also animations, audio, video and some other media channels.

Animation is a common way of P2P network communication, even if there is a few frames of GIF images, static expressions can change into coherent actions.



Another kind of animation occurred, such as the "winks" on MSN. No matter what technology is based on, or how creative and unique the visual style is, their principles are the same, that is, using multimedia to convey and deepen the effect.

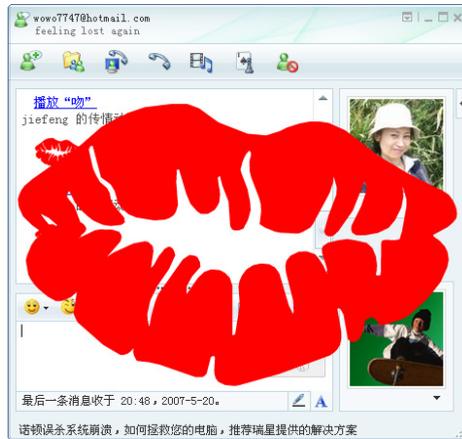


Fig. 2. Animation with sound and video in MSN

3 Features of 'Peer-to-Peer' Wed Art Design

3.1 Cultural Identity: National Movement and the Grassroots

It has become a vigorous national movement that integrate art design elements into information interaction consciously or not. Each user of "peer to peer", may be the

audience of the network art design, and also may be the main part of creation and dissemination in the networks art design. Because of this, there is no utilitarian, representation of elitism, or competition for the right of saying in 'peer to peer' network art design right from the start. In this network, each user is equal, what they need is enthusiasm, inspiration, humor and wit, and a little time for such pursuits, which is the wealth the class of 'grassroots' has, so 'peer to peer' network art design is popular in public.

Grassroots in culture is the vitality of "peer to peer" networks art design, only rooted in the lives of ordinary people, it would be full of inexhaustible creative subjects, be widely approved by society, survive and develop.

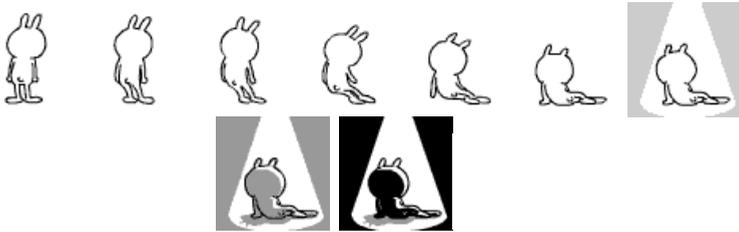


Fig. 3. 'lost' and 'crazy': Animation of Tusiji which is designed by a 22-year-old college student and popular n China

3.2 Ceation Features: Creative Spirit and 'Spoof Doctrine'

A major feature of "per to peer" network art design is the emphasis on originality, which means that pay attention to novelty and innovations. The solution is to open a new path to pursue creativity. Innovation is needed unprompted, social recognition measured by prevalence and the achievement of 'self-fulfilment' become the driving force of creation.

On the other hand, in the process of pursuing for new ideas, their creative technique is " loony tone" style, any images, traditions, concepts, and models, which are deep-rooted, can be used to transfer, collage, ridicule and even subversive, Some of them, though considered as the subject that is "difficult to accepte", are still used openly and legally. It is undeniable that this 'non-mainstream' approach has become 'mainstream', and an important embodiment of 'creative spirit' in network art design.

 (Indecent graphics of official expression in QQ)



Fig. 4. Spoof expression animation of panda



Fig. 5. modified graphic photos of children

3.3 Dissemination Features: The ‘Chain’ Style to the ‘Net’ Style

If the communication structure of network art design based on website is ‘root’ type - spread from a source point to the endpoints, the basic structure of communication in ‘peer to peer’ network art design can be called ‘chain’ structure which is much simpler. However, there is no difference between ‘source point’ and ‘endpoint’ in this structure, each source point is also the endpoint; The endpoint in this chain structure is considered as the source point in the next one. Then the recirculating network formed, the source points and endpoints disappeared eventually.

In the ‘peer to peer’ special mesh mood, in network art design, the boundaries of creators, publishers and receiver have been blurred, each one may have many different identities, and influence one another. Meanwhile, the breadth and speed of communication in the network art design are improved because of the ‘chain reaction’.

4 New Forms, New Thoughts

4.1 The Uneven Quality of Art

The creation and dissemination of ‘peer to peer’ network graphics art design is a spontaneous national movement, the same as the other folk art forms (for example, the controversial Northeast Errenzhuang Opera at present), its contents and forms are in uneven quality, some works of low level, vulgar style, flubdub, and even sexual violence, come up along with innovative ideas.

This may be related to low technical threshold, or related to the subjective quality of grassroots, what’s more, it reflects that people’s aesthetic sense and creativity need to be training and guided while popularizing the network art design. Compared to the modern art design, P2P network graphic art design should be considered as fertile soil where art design grows, instead of being considered as ‘low class’ and kept a respectful distance, then promote it to a healthy, noble, artistic direction.

4.2 The Lack of Ethical Regulation

Properties of private rights is particularly prominent for the dissemination of ‘peer to peer’ graphic art design is between private persons, in addition, the public power can not be involved, the phenomena against morality would appear inevitably. Although it is not illegal, these phenomena may infringe some other people’s rights, result in the pollution of the network environment, and some social problems, so it should be paid attention to as early as possible.

We should solve this issue by both compulsory means such as legal, security, and ethical means of creating social attitudes, public opinions, moral cultivation and harmonious environment, in this way, graphic art design can be 'vulnerable to the attack'.

4.3 Protection of Intellectual Property

Intellectual property has become the focus with the spread of 'peer to peer' network communication., freedom and sharing is always the spirit of the network, but the protection of intellectual property rights have emerged,for example, a Mr.Chen obtained criminal penalties on Hongkong for the first time because of disseminating pirate movies through BitTorrent in the internet. 'Peer to peer' network art design is also facing the same torture: whether the intellectual property right exist in the spontaneous, 'anonymous' and 'peer to peer' network art design; how creators to take advantage of and protect their intellectual property; the protection in what degree can keep the 'peer to peer' network art design from influence.

These problems must be thinking, the corresponding system of laws and regulations should be formed after the achievement of practice and consensus, it is believed that the problem can be solved if we can protect the creators' legitimate rights and interests, and also promote the characteristics and advantages of network to motivate the development of network art design.

5 Summary

The "peer to peer" network art design shows distinctive features and significant advantages as a new form of design and communication, it has a unlimited prospect. We should face this problem positively, and make full use of this broad platform to promote the diversified development of contemporary art design and its dissemination.

References

1. Lv, J., Chen, H.: Research of Network Graphic Art Design. Theory Monthly, 127–128 (September 2000)
2. Lv, J., Chen, H.: Restudy of Network graphic art design. Journal of Wuxi University of Light Industry (1), 150–153 (2001)

Design of Musical Performance Mode Based on Cooperative Learning

Yumeng He

The Music Institute of Liaoning Normal University
Dalian, Liaoning, 116029 China
yumeng.dalian@yahoo.com.cn

Abstract. Although most studies of cooperative learning and training for music performance focus on dimensions of musical systems, recent studies suggest intelligent factors of performance are also represented. The purpose of this paper is to discuss a mode of “intelligence musical performance” approach to establishing the new music performance learning and training. This approach will include human music mapping, thereby overcoming limitations of conventional knowledge and experience approaches like expert in applying to human ware. Thus, in order to establish intelligent systems new approach is required to effectively deal with non-logical aspects or empirical knowledge of human beings and organizations.

Keywords: Music learning, intelligent performance, non-optimum to optimum, cooperative learning.

1 Introduction

Musical Performance (MP) are an intelligent process of musical works realize that harmony, tonality, phrasing, and meter combined with traditional and modern musical performance models (Palmer, 1997; Gabrielsson, 1999; Palmer & Pfordresher, 2003). In some cases it is very difficult to draw a line between intelligent and non-intelligent natural and artificial systems. For example, biological adaptation or any kind of evolution can be presented as learning intelligent ability or non-intelligent process. It is difficult to determine when learning system became an intelligent system. All intellectual activities are triggered by the goal. All kinds of intellectual activities in the musical performance area are based on learning and training of music-knowledge and performance experience, but performance intelligence is not knowledge as well it is not experience. Performance knowledge is a “tool” of music learning and intelligent performance is a “behavior”. If you don’t understand the goal of the performance training, you are not capable to reach it. Ability to performance is an important music intellectual ability that can improve music works.

Researchers and practitioners have long been concerned with three fundamental issues involving musical performance. The first issue involves what people performance--the identifiable knowledge and skills outcomes of from accumulated performance experience. The second issue involves the process of learning (i.e., just how do we learn?), what are the sequences of performances and activities that cause

or facilitate learning? The third raining issue is a more practical one and involves intelligence for musical performance (i.e., designing and building learning model or learning machines to facilitate the training process). The fundamental idea behind the concept of a level for musical performance is a simulated situation designed to create personal experiences for players that serve to initiate their own process of inquiry and understanding. The objective of this paper is to analyze virtual training model and to discuss a method for designing learning and training systems based on intelligent musical performance.

The rest of this paper is organized as follows: In Section 2 we give an outline of the design of intelligent musical performance methodology. In Section 3 we describe the overall structure of the learning and training. The final section 4 the paper concludes and points out further work.

2 Overview of Music Learning

In this section, we present a brief summary of the music learning, and then present an overview of the implementation of cooperative learning mode (CLM).

A. *Music concept Based on Brain*

Music and the human mind have a unique relationship that is not yet fully understood. As Hodges forwards, By studying the effects of music (Hodges, 2000, p. 21), likewise, through music we are able to discover, share, express, and know about aspects of the human experience that we cannot know through other means. Musical insights into the human condition are uniquely powerful experiences that cannot be replaced by any other form of experience. While the effect of music on the critical aspects of learning, attention and intelligence may be a relatively new area of focused research, the human brain may very well be hardwired for music. Exactly where this hardwiring might be located would be difficult to say. For example, even though there is an area in adults identified as the auditory cortex, visual information goes into the auditory cortex, just as auditory information goes into the visual cortex. That is why certain types of music can stimulate intelligence recall and visual imagery (Nakamura *et al.*, 1999).

This process of selection continues as the rest of life is played out. This is the process of learning, selecting, connecting and changing our neuronal patterns (Edelman, 1992; Zull, 2002). Music plays a core role in this process. Jensen contends that, “music can actually prime the brain’s neural pathways” (Jensen, 2000b, p. 246). The brain has the capacity to structurally change throughout life. As Begley describes, “The actions we take can literally expand or contract different regions of the brain, pour more juice into quiet circuits and damp down activity in buzzing ones” (Begley, 2007, p. 8). During this process of plasticity, the brain is expanding areas for functions used more frequently and shrinking areas devoted to activities that are rarely performed.

Further, in the late 1990’s neuroscientific research discovered that the structure of the brain can change as a result of the thoughts we have. Because people have stored representations of songs and sounds in their long-term intelligence, music can be imagined. When a tune is moving through your mind it is activating the same cells as

if you were hearing it from the outside world. Further, as we have noted, when you are internally imagining a tune, the visual cortex is also stimulated such that visual patterns are occurring as well (Sousa, 2006).

More specifically, knowledge association is focused on as a sample of human cooperative ability. This system is designed to allow human cooperative abilities to cooperate with musical logical functions, thereby facilitating human creative activities to music performance in real world environment. In order to do this, a guide function of human intelligence or association is devised and incorporated into the musical performance system. This is called association guidance. The knowledge association process (from initial knowledge or situation to association knowledge or situation is divided into two processes. When initial knowledge of performance is given, the human role is to related words or phrases, called keywords, as indicators/hints or cues to already be stored in the cooperative system to help the performer. The cooperative system first role, which is called association guidance, is to provide suggestions to stimulate or guide this human music concept.

B. Mode from non-optimum to optimum

Thus, an analysis process of the experience of the music performance with optimum and non-optimum is calling interactive sub-optimum. The concept of the intuition non-optimum is quite comprehensive. From the viewpoint of problems' entity, non-optimum means unfeasible and unreasonable; from the viewpoint of problems' behavior, it means non-ideal and non-good; from the viewpoint of problems' capacity, it means ineffective and abnormal; from the viewpoint of problems' change, it means obstacles, disturbance and influence. There exists a serious of non-optimum experience from the entity of the problem to the change of the problem, which causes non-optimum category. As to every kind of human intelligent analysis, there is the individual non-optimum category as well as the common non-optimum category. The so-called individual non-optimum category is decided by the characters of the problem, while the common non-optimum category is an objective entity. Figure 4 is a method of from non-optimum correction to optimum. There are three attributes of the recognition to the sub-optimum problem, experience, intuition and knowledge. The attribute of experience reflects the recognition to the characteristics of the object's behavior. Here the selection of the factors of the decision-making is discussed from the experience attribute's viewpoint. Performance experience provides sub-optimum for the problem. When the recognitions are different, the sub-optimum is different as well. The tracing to the system's conditions of the past can propose a sub-optimum [7]. See Fig 1.

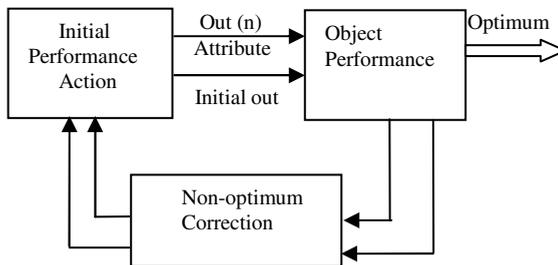


Fig. 1. Learning pattern in non-optimum to optimum

3 Cooperation of Performance Skills

Another source of evidence for cooperative dimensions of intelligence for performance is the conditions under which performers generalize what they know from one performance situation to another. In typical transfer of learning tasks, participants learn one task and then perform a second task. The ease with which they perform the second task is thought to reflect what was learned in the first task. The mirror symmetry of hand and finger movements in piano performance provides a convenient venue in which to test transfer of cooperative learning. Consider the sequence of finger movements 5-4-3-2-1 in the right hand (where thumb = 1), used to press adjacent keys on a piano; this sequence becomes 1-2-3-4-5 when the same keys are pressed by the left hand. Thus, the same melody can be played with different hand and finger movements in piano performance (See Fig.2).



Fig. 2. Finger movements in piano performance

In addition, different melodies (different pitch sequences) can be performed with the same hand and finger movements. Although the transfer of musical skills to math, spatial reasoning, and other tests of intelligence has been studied (Rauscher, Shaw & Ky, 1995; Schellenberg, 2004), fewer studies have examined the transfer of learning from one musical task to another musical task. Are well-learned cooperative movements transferred from one melody to another? Palmer and Meyer (2000) measured transfer of pianists' hand and finger movements from one practiced melody to a novel melody. When the second (novel) melody required the same finger movements as the initial melody performed, pianists were able to play it more quickly. Comparisons across age and skill levels indicated that the more advanced performers showed greatest transfer across cooperative movements, suggesting that they were able to generalize from one hand and finger set to another. The least skilled pianists showed no ability to generalize beyond the particular finger sequence learned.

Does knowledge of a set of cooperative movements transfer across temporally distinct patterns in music performance? For example, can a pianist performing one rhythmic pattern with a set of finger movements such as 1-2-3-4-5 generalize those same finger movements to produce a different rhythmic pattern as fluently as the first pattern? Using a similar transfer of learning task, Meyer and Palmer (2003) showed

that pianists' finger movements did transfer across different rhythmic patterns: pianists could perform novel rhythms with the same finger movements as quickly as the first-learned rhythms. In addition, there was rhythmic transfer: pianists could perform well-learned rhythms with different hand and finger movements as quickly as with the first-learned movements. There was no interaction between rhythmic transfer and cooperative transfer; the time at which keys were pressed were remembered independently of the cooperative features that produced them. Similar findings were obtained when pianists transferred from one melody to another that differed in meter and cooperative movements. No interactions were observed among the temporal (meter and rhythm) and cooperative structures: retaining temporal structure from one melody to the next facilitated speeded performance more than retaining cooperative movements, and cooperative features played a smaller role in transfer of knowledge across melodies. These findings are consistent with the general view that representations of timing in sequence production are not defined primarily in terms of cooperative features (MacKay, 1982, 1987; Semjen & Ivry, 2001), at least for performers with moderate to high levels of musical experience.

4 Conclusions

Research in music performance is beginning to document the nature of intelligence for the cooperative aspects of performance. Performers' hand and finger movements, as well as conceptual intentions, are encoded in intelligence for performance and tend to have independent effects on pitch accuracy. Performers' intelligence for melodies (specific pitch sequences) and finger/hand movements generalize in transfer tasks; furthermore, the cooperative and melodic information transfer independently. Skilled performers show more transfer of learning across melodies that required different cooperative movements than novices. Mental practice shows evidence of cooperative components that facilitate intelligence for performance. Finally, motion capture techniques of measuring music performance are beginning to document the time course of anticipatory movements.

One ramification of these findings is that intelligence for performance is flexible; performers can apply what they know about cooperative movements to different performance situations, and behavioral and neural changes result, as seen in the plasticity with which no musicians learn aural-cooperative associations. Flexibility of cooperative movements is essential to the interpretive nature of music performance; otherwise, significant additional practice would be necessary before a musician could perform a familiar piece with an alternative interpretation. A second ramification is that performers differ in their individual abilities, as evidenced in interpretive effects on intelligence, in transfer of learning from one melody to another, and in mental practice benefits. Mental practice is appropriate for study of brain states, measured in EEG, fMRI, and MEG studies, because of its avoidance of motion "artifacts". Scientific interest in applying imaging methods and motion capture techniques to music performance suggests that answers may soon be found to the interesting question of how cooperative aspects of music performance are represented in intelligence.

References

1. Bangert, M., Altenmüller, E.O.: Mapping perception to action in piano practice: A longitudinal DC-EEG study. *BMC Neuroscience* 4, 26–40 (2003)
2. Chaffin, R., Imreh, G., Lemieux, A.F., Chen, C.: Seeing the big picture”: Piano practice as expert problem solving. *Music Perception* 20, 465–490 (2003)
3. Coffman, D.D.: Effects of mental practice, physical practice, and knowledge of results in piano performance. *Journal of Research in Music Education* 38, 187–196 (1990)
4. Dalla Bella, S., Palmer, C.: Tempo and dynamics in piano performance: the role of movement amplitude. In: Lipscomb, S.D., Ashley, R., Gjerdingen, R.O., Webster, P. (eds.) *Proceedings of the International Conference on Music Perception and Cognition*, pp. 256–257 (2004)
5. Dell, G.S., Burger, L.K., Svec, W.R.: Language production and serial order: A functional analysis and a model. *Psychological Review* 104, 123–147 (1997)
6. Rake, C., Palmer, C.: Skill acquisition in music performance: Relations between planning and temporal control. *Cognition* 74, 1–32 (2000)
7. Driskell, J.E., Copper, C., Moran, A.: Does mental practice enhance music performance? *Journal of Applied Psychology* 79, 481–492 (1994)
8. Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., Taub, E.: Increased cortical representation of the fingers of the left hand in strong players. *Science* 270, 305 (1995)

Building Distance Education and Training System for Police by Making Use of Information Technology

Shuxia Liu

Department of Scientific Research
Liaoning Police Academy
Dalian, Liaoning, China
maillsx@163.com

Abstract. Based on the theoretical research and actual developing status of police training and education in the China, this paper discusses distance education and training system for police by making use of information technology. The objective of this research is to building of distance education and training system for public security police based on the background of "Great Education" and "Great Training" for police officers. In this paper, a review in requirements from informatization training of police is done, and the technology aspect of distance education is to be analyzed. Also, general design principles are to be reviewed. According to all these issues, recommendations to designing a distance education training system for public security police.

Keywords: Great Education, Great Training, distance education, training system.

1 Introduction

The Party Committee of the Ministry of Public Security has required that under the background of big training and basic training, it is of great strategic significant to actively seek for and build "Great Education" and "Great Training" work system with police characteristics, so as to take on new missions, meet new challenges and satisfy new demand from the people toward public security work. One of the Major issues before us is how to actively advance the establishment of "Great Education" and "Great Training" system, convert the empirical and extensive education and training to standardized, precise, scientific education and training, change the decentralized and single-type education and training resources toward that of intensive, highly efficient type, gradually build Distance education and training system for police and enhance the overall quality and law enforcement capabilities of police officers by creatively making use of network information resources.

2 Problems Proposed and the Background of the Issue

2.1 The Basic Content of "Great Education" and "Great Training"

"Great Education" and "Great Training" are rich in meaning. First, as to the time of training, it means not temporary training to meet requirements from the work, but

life-long training. Therefore, "Great Education" and "Great Training" extend one-time education and training to life-long education[1]. Second, as to the target of education, it was designed for education and training of all the police officer. Therefore, according to "Great Education" and "Great Training", all the police officers are to be educated, not partially. Third, as to educational institutions, according to the idea of "Great Education" and "Great Training", the education and training of police officers shall not merely rely on Special Police Education Institutions, but will also integrate other educational institutions including police colleges, special police training institutions and even social education agencies to provide police officers with education and training at different levels, make the training work more efficient and finally achieve the training goal. Fourth, as to the ways of training, according to the idea of "Great Education" and "Great Training", all the effective ways or methods including independent studying, academic education, non-academic education, vocational training and overall quality training shall be taken. Fifth, as to the goal of the education and training, according to the idea of "Great Education" and "Great Training", such education & training are no longer only a matter of making a living or instruments and means of seeking for material gains, but are to perfect the character of police profession and to realize the purpose of serving the people.

Therefore, compared with closed education and training, "Great Education" and "Great Training" advocate diversified, socialized and compound training by integrating all available educational resources coming from police training institutions, police colleges and social education agencies, which transcend traditional schooling and training. Such education and training also enhance the quality of the police force, increased their fighting strength and kept the continuity, stability and development of public security work[2].

2.2 Background of Distance Education and Training for Police Officers

It is indicated in the 17th Congress of CPC that we should keep on wholesale training of cadres, greatly advance the quality of our cadres and call on the party and society to establish the idea of great education and great training. Early in the year 2008, Minister Meng Jianguo stress on public security directors' meeting that we should actively build great education and great training work system conforming to the practical situation of public security with our own characteristics and that we should deploy and design training work according to the strategic situation. In June, 2008, Zhou Yongkang, Secretary of national Politics and Law Committee specifically propose that we should reform the enlisting and distribution systems of politics and law colleges, draw lessons from various vocational training institutions and establish long-acting training mechanism with the characteristics of our own police. In August, 2008, the Ministry of Public Security held a meeting in Nanjing on which "three great construction" were confirmed and they are "informatization construction, normalized law administration construction and construction of humorous relationship between police and people, which means that public security work has entered into information age. The training work on police officers must be directed by learning and practicing scientific thought of development, supported by advancing informatization construction, assured by strengthening the normalized construction of law administration, sustained by constructing harmonious relationship between police and people. We should further

push the "three great construction", carry out the spirit of "Great Education" and "Great Training" proposed by Minister Mang Jianzhu, draw lessons from various training and education institutions both abroad and home, made the most of modern technology to facilitate teaching and training, continually make innovation on the ways and means of public security training, carry out in its entirety the requirements of "three must training", satisfy the requirements of various training missions and constantly advance the overall quality and practical ability of the police officers[4].

3 Practical Experience of Distance Education and Training System for Police Officers

3.1 Distance Education and Training in Developed Countries and Regions in the World

The training of United States Army was completed by life-span learning. In the study plan of land force, many different means and technologies were used to achieve synchronous or non-synchronous exchange between students and teachers, and standardized single, collective and self-study training were available online for soldiers and troops. It can provide academicians, officers and troops essential information at any time and anywhere. The missions of such plan was used to help culture officers with high potential and strong adaptive capacity through searching current network learning (E-learning) in life-time learning process. Through providing soldiers and officers with training and education at all hours and at anywhere and constantly updating the training content to enhance the combat readiness level and back land force transformation.

Hong Kong Police Force and Japanese Police Department also use such network training mode for training of police officers. In the promotion training for police officers of Hong Kong Police, network was used as the learning platform and you should past the basic training subjects before entering into the centralized training of next stage. Japanese Police Department has its own network institutes for police officers, training plan and arrangement for police officers at different levels were published online, so they can log onto the internet to learn, seek for information and advance their professional skill at any moment[5].

3.2 Lessons from Distance Education of Liaoning Police Officer College

Liaoning Police Officer college has undertaken Distance education of universities and colleges level, 75% of the courses (with priorities given to elementary courses and specialized core courses) are to be completed online by local universities and 25% of the courses (with priorities given to special courses and specialized core courses) are to be completed on Golden Shield Web by public safety departments. All of these courses are to be arranged according to the needs of grass-roots work and complying with the characteristics of public security industry.

The college has developed distance education web that is mainly composed by three platforms: the first is teaching platform on which courses like foundational public security courses and theory curriculum are given, this platform also provide course Introduction resources (teaching resources, course Introduction and teaching

video), online learning resources(teaching courseware, network courseware and video courseware) and online work, Online counseling, online testing and discussion; the second is interactive platform on which there are experts forum and online question-and-answer aiming at training of police leader' management decision-making capacity; the third is educational administration management platform on which application, registration, management of school roll, tutorship, reservation can be completed online. This kind of education has been reported by Guangming Daily titling "propping up the future". It is this kind of institutional innovation that drives education and training of police officer towards modern distance education.

3.3 Lessons from Basic Training Platform of Liaoning Police Officer College

Under the leadership and panning of the Public Security Department of Liaoning Province, in August , 2007, Liaoning Police Officer college successfully developed "Basic Training Platform for Police Officers of Liaoning Province"(hereafter shortened as "Platform") that is divided into four functional modules as studying, training, testing and management. The studying module functions like a large-scale electronic library that provide various kinds of professional knowledge, so police officers can study and practice all by themselves in their office or at home. Through the training module, police officers can review and test the professional knowledge and professional skill they have learned online. The testing platform shall be able to adapt to different categories of police forces, police forces at different posts and different contents and meet the requirements from intelligent testing and automatic testing. The management platform is mainly used to manage files and two databases were establish, namely electronic training file and health records, so that we can learn the police officer' training work and health records at any time. This four platforms reflect the characteristics of information age, make good use of the advantages of fast and wide-coverage network and enrich the training methods.

Required conditions for conducting online education and training of police officers are available. Relying on Police Computer Network, related departments in Beijing, Shanghai and other major city have organized "online training". The training content was put on the Web, so that grassroots police can participate in such training at any time and in anywhere. Related departments in Shanghai have also developed assessment and examination software to better facilitate the development of online education and training. Countrywide, online education and training for police officer has just started.

4 Building and Design of Distance Education and Training System for Public Security Police

4.1 Building of Distance Education and Training System for Public Security Police

1) *Category of Education and Training*

In accordance with article 3 of People's Policemen of Public Security Organ Training Regulation, the category of education and training in public security college and all

level public organs is composed of Junior Education & Training, Promotion Education & Training and Professional Education & Training. According to the current work requirements, it can also be specifically classified into several training topics, such as primary education and training (cub police training), development education and training (police officer training, with rank promotion), intermediate police officer training, cub police officer training, leader training, with rank promotion, director in county district training etc.), practical application training (training topics includes position professional training, Olympic security officer, maintaining social stability and dealing with emergence etc.). Meanwhile, focusing on three foundations and five functions, it also sets up the following training topics, such as public security application, law enforcement standardization, mass work, psychological quality, police tactical skill education, gross root practice, dealing with complex situation, public opinion guidance and control etc.

2) Course Model of Distance Education and Training

The Curriculum system of police officer distance education and training is setting up in accordance with different marine species and position quality standard. The core of position quality standard is the core ability, which includes three aspects, such as career view of value, business knowledge skills and competence. In the building of course system of distance education, it should abide by the following guidelines: focusing on practice, professional, properly supplement theories, strengthening skills and thinking training, and be constructed according to those five knowledge modules, such as fundamental knowledge, basic skill, police skill, mental training and position knowledge. There are two course system models in the construction of course system, including knowledge-oriented and practice-oriented. The target of knowledge-oriented model is that "grasping systemic theoretical knowledge, developing understanding and thinking skills". Those courses can take the model of online learning and training in accordance with features of online education [6].

3) Examination Mode of Distance Education and Training

Examination result is one of standards to evaluate effectiveness of education and training, and also a fast, objective and justice method to evaluate education and training results. It can be divided into different level examinations, such as provincial (ministry) universal examination, municipal (county) self-study examination and marine species quizzes in accordance with organization organs. The difference of those examinations is examination organization authority limitation. According to the exam model, it can be divided into three types, including objective, subjective and operational exam. The evaluation of police officer's results should adopt various parameters, such as amount of surfing in the Internet, handing homework condition, amount of on-line communication and results, and various weight coefficient should be set up to evaluation and investigation. Eventually, specific evaluation results can be obtained. The keystone of examination system is the construction of examination content which should embody the target of education and training.

4.2 Design of Distance Education Training System for Public Security Police

Online independent training for public security police is to choose learning schedule, learning resources and methods according to specific learning targets and evaluate

one's achievements to help them experience the complete process of "determining mission-stipulating plan-implementing plan-control inspection-evaluation feedback". Firstly, there must be a resource construction system which can guarantee implementation of various trainings to build resource integrated system and course setting evaluation model to ensure most standard guarantee for training task. Furthermore, there must be a carrier for implementing training plan-education training system and a education and support service quality insurance system based on professional, standard and 'feedback -improvement mechanism'; finally , to establish a network education assistant management platform which integrates " education training management", "education resources construction", " academicians support service" and "technology support guarantee" these four sub-systems and be of specific management ,resource configuration and quality control alarming functions. "Education training management " system including training planning, learning management, archive management and test affair management etc..

"Education resources construction" system including course construction, information collection, inspection and type-in; "academicians support service" including assistant management, consulting management, blog management, BBS and platform management and other assistant support services; technology support guarantee includes "system monitoring and running maintenance, technology equipment maintenance and project development", which intake every specific working link to a system with flexible organization and coordinated working distribution; it connects with every link in process of education training with closed feedback system and realizes the standard and specific network education training and service management , it ensures the effective implementing of education and training plan with system construction, establishment of coaches team and quality evaluation these substantial organizations.

4.3 Guarantee of Remote Education and Training for Police Officers

With the establishment and perfection of three stage training network for police officers, the education and training are developing toward intensification and scale. There are generally over 50,000 police officers in each province, so the online learning platform shall be designed to meet the needs of over 100,000 police officers and relevant guarantee system shall also be established to meet the requirements of large-scale training [7,8].

1) Guarantee of Training Content and Instructors

Speaking of a certain course, large numbers of auxiliary materials are required to support the implementation of online training. Different segments (see Chart 1) shall be designed for the learning process of individual trainee, and in designing such segments, materials of different kinds, knowledge points, courseware and training goals shall be separated or integrated for the development of resource pool based on knowledge points. The physical objectives of learning and relevant knowledge points shall be defined for trainee to meet the requirement of diversified training. Priorities shall be given to "Instructor System" and the diversified and multi-functional teaching team combining "instructor system and teacher system" shall co-exist to guarantee the quality of the training.

3) *Guarantee of Remote Management of Police Officer' Training Experience*

Training experience refers to training and education police officers have taken since they started to work. With the reform of the police promotion and training system, training and employment have already been closely linked. To advance the quality and ability of people's police has been kept in step with their promotion. Promotion shall also follow the principle of "promotion after testing, promotion together with testing, incumbency with training and complement training ". Training at various levels shall be linked with appointment, examination and evaluation. To strengthen the management of training experience is more pressed and intense than ever before. After "Nanjing Meeting", activity for the purpose of promoting the police officer' ability of information application have been organized all over the country. We should make the most of modern interest technology, especially, the resources on Golden Shield Web to reform traditional training management of the administrative order type and step forward toward new management of serve type. We should also elevate the training enthusiasm of police officers by transferring from "being required to training to wanting to be trained", and provide foundation for scientific management of training experience. Training records management mainly include police officer' essential information, training experiences, health level, annual situation and real-time training notification. Through the establishment of training records management, police officer' training experience can be inquired and counted easily. On one side, police officers can inquire their own training records and make their professional layout; on the other side, public security departments can realize scientific, systemic, standard, convenient and efficient management of police officer' training records [9].

5 Conclusions

Today' world depends on information technology, so training and education of police officers shall be established on an platform that is open to all the police officers for their life-ling training. Different training institutions should exchange their fruits and lessons to promote sharing of training resources of different kinds among police forces and also to realize sharing of high-quality training resources on large scale. Sharing of training resources on network could bring about a lot of benefits, all the cases and problems encountered in real work can be directly converted to instructional resource through network, so all the police officers can share that resources to considerably advance such training. At the same time, we can make the most of information technology to make innovation in Management Model and Supporting Systems, build and perfect long training mechanism of competition, constraint, examination and evaluation, form standardized management finer operation, so as to constantly promote police officer' learning consciousness and advance the level of education and training and the quality of training.

References

1. Phillip, B.: Distance Education in Africa: New Technologies and New Opportunities. JICA Report (February 23, 2000)
2. China Internet Network Information Center: SemiAnnual Survey Report on Internet Development in China (January 17, 2001), <http://www.cnnic.net.cn/develst/e-cnnic200101.shtml>
3. Collis, B.: Tele-learning in a digital age. The future of distance learning. International Thomson Computer Press, London (1996)
4. Jesshope, C., Heinrich, E., Kinshuk: On-line Education using Technology Integrated Learning Environments. Massey University, New Zealand; Kempainen, P.: Network Intelligence: A Paradigm for the Service and Networking Convergence towards Universal Communications. IEC Annual Review of Communications (June 2000)
5. Koschmann, T.: Computer-supported problem based learning: theory and practice of an emerging paradigm. Lawrence Erlbaum Associates, Mahwah (1996)
6. Stahl, G., Sumner, T., Repenning, A.: Internet Repositories for Collaborative Learning: Supporting both Students and Teachers. In: Proceedings of the Computer Supported Collaborative Learning Conference (1995)
7. Liu, S.: Thinking and Exploring on Distance Education Training Mode for Public Security Police. Public Security Education (2008) (Chinese)
8. Yang, Q.: Innovation and Practice of Network Education Management Mode Aiming at Improving Cultivating Quality of Talents. Summary of Liaoning Province Education Achievements (December 2008) (Chinese)
9. Shanghai Public Security Senior Colleague Theme Team. Research on Occupational Public Security Education Training System. Journal of Shanghai Public Security College (December 2007)

Research on Mathematica Experimental Teaching Reform of High Vocational Education

Jing Ruan

Wenzhou Vocational & Technical College, Wenzhou, Zhejiang, China
ruanjing1979@126.com

Abstract. This paper describes the construction content of the experimental test platform for Mathematica. We also introduce detailly the functional requirements of platform and the technology and developing tools to be used with the platform construction.

Keywords: Online testing system, B/S, developing tools, ASP.NET.

1 Introduction

With the development of information technology, more and more exams exploit the paperless examination- examination on computer. Paperless examination can achieve the separation of teaching and examination, and manage the affairs automated fully, which can effectively make use of the hardware and software resources of campus network, so as to exert their fullest effect, for the better school's teaching, research, management services. The objective and transparency of the exam is achieved by this test method. The automation generating test paper reduces the intensity work of the teachers, trains the ability of students using the computer, enhances the awareness of students' mathematics practicing, and save a lot of material, human and financial resources.

The Mathematica math test marking exam papers in college is quite difficult and hard to measure, and the topic is not a scientific response to the various sections of knowledge. Each test, the scoring results are known after the teachers scan and score, the evaluation has a significant lag, which is detrimental to the improvement of student's self-evaluation capabilities. With the rapid progress of computer technology, the intelligence degree of educational software also has made substantial progress. The education conception of modern society and the academic teaching experience make the computer technology teaching play a special function in education teaching. Naturally, the combination of Mathematica mathematics experiment test and computer technology, the construction of Mathematica mathematics experiment test are becoming feasible and necessary.

2 The Overview of Mathematica Mathematics Experiment Test Platform

The mathematica mathematics experimental test platform is an online examination system provided for Mathematica mathematics experiment test. That is, for the students do the the Mathematica mathematics examination on computer terminals in connection to the network through the network environment. The papers of Mathematica

mathematics test platform are extracted from the database server dynamically after the generation of questions; the exam's answers are transmitted through the network and recorded in the server; after the papers submitting, the marking work is done automatically by a computer, the final results of candidates are given directly. To achieve the above ultimate goal, the function of the Mathematica mathematics experiment test platform consists of papers section and exams management section. The papers section based on papers database to generate examination papers automatically, record the answer of the papers, and complete the marking of papers or other functions. The exams management section is mainly for centralized management of students and examination information.

3 The Requirements of System

3.1 The Requirements of Function

(1) the generation of paper After students login, the test platform selects randomly a number of topics composed of a set of complete papers according to the knowledge points built by the paper base. The benefits of random test paper: it can guarantee the coverage of papers and avoid cheating among students.

(2) the accuracy of judging the paper

After the students clicking the keep the answer record in the examination, the teachers can monitor the examination process of the students, and observe the understanding situation of students' knowledge from the students' answering speed .

(3) the display of papers and answer record

The teachers can always check the papers of students and the answer records saved currently, so as to enable the teachers to understand the answer situation of students, which can be used as a final basis for paper analysis.

(4) Deriving the final results

When all the students exit the examination system after handing in the examination paper, the teachers can click on the button export results to export the class students' results in Excel file format. This function can meet the statistical performance of teachers and allows teachers to export the format of the results .

(5) the management of class and students

The system should allow more than one class simultaneously, and has the capability of at least 300 people online at the same time. Where the class is allowed to add and remove, and students are not allowed to delete in a non-empty class, the classes and students are on the examination can not be deleted.

(6) record and query the examination information

The system should have the ability to save the historical data of 3-5 years, which facilitate the teachers to query and analysis the historical data.

3.2 The Description of System Function Structure

(1) The functional block diagram of system

The function block diagram of Mathematica mathematics examinations platform system is shown as Figure 1. The system is divided into the three target user interfaces of administrators, teachers and students. The students interface has the papers browsing function, answer function and change passwords function, etc; the teacher interface can

manage courses, to analysis the results of the examination flexibly. The administrator interface allows administrators to manage the examination, as well as the backup management of data .

(2) The login process of system

The examination system requires rigorous security authentication to determine the categories of users and permissions on user login.

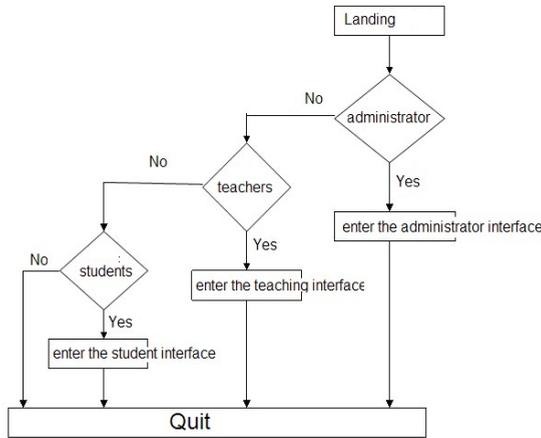


Fig. 1. The functional block diagram of system

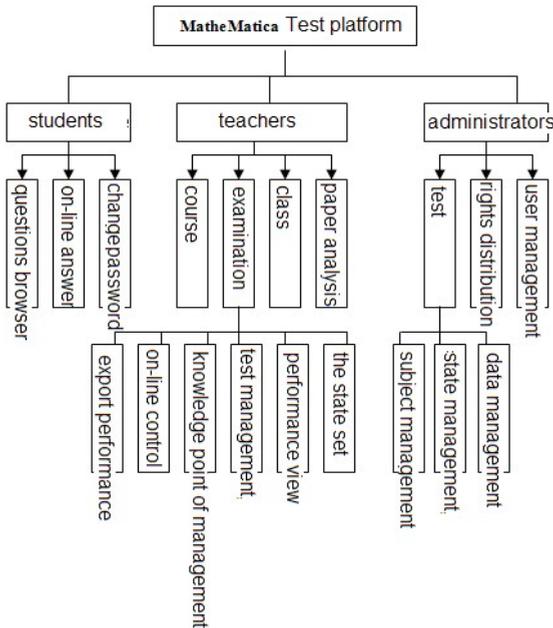


Fig. 2. The process of login and validate

4 Core Technology and Development Tools

The development of Mathematica mathematics examinations platform is the online examination system of B / S structure, whose development involves the ASP.NET technology and the stored process of database. The development tools employ SQL Server 2000 and Visual C #. NET.

4.1 Core Technology

(1)ASP.NET technology

ASP.NET is part of Microsoft.NET. As a strategic product, it not only the next version of Active Server Page (ASP), but also it provides a unified Web development model. ASP.NET is a compiled environment based on '.NET', which is compatible with any. NET- languages (including Visual Basic. NET, C # and JScript. NET) .

The Managed Code such a new concept is introduced in a new ASP.NET, across the entire Windows development platform. The Managed code runs in NGWS Runtime. And NGWS Runtime is a run time environment that execute the manages code, such that the programming is easier.

This system uses a total of 51 stored procedures, all the stored procedures of this system are shown as follows:

| 存储过程 51 个项目 | | | | |
|----------------------------|-----|----|------------|----------|
| 名称 | 所有者 | 类型 | 创建日期 | |
| sp_ClassUpdate | dbo | 用户 | 2008-12-22 | 12:02:19 |
| sp_ClearTestGrade | dbo | 用户 | 2008-12-22 | 12:01:58 |
| sp_CourseAdd | dbo | 用户 | 2008-10-28 | 11:10:17 |
| sp_CourseUpdate | dbo | 用户 | 2008-10-28 | 11:10:17 |
| sp_deleteClass | dbo | 用户 | 2008-10-28 | 11:10:18 |
| sp_deleteCourse | dbo | 用户 | 2008-10-28 | 11:10:19 |
| sp_deleteKnowledgeByID | dbo | 用户 | 2008-10-28 | 11:10:19 |
| sp_deleteStudent | dbo | 用户 | 2008-12-22 | 12:03:01 |
| sp_deleteTest | dbo | 用户 | 2008-10-28 | 11:10:20 |
| sp_deleteTest1 | dbo | 用户 | 2008-12-22 | 12:01:51 |
| sp_deleteTestState | dbo | 用户 | 2008-10-28 | 11:10:21 |
| sp_GradeTestWithNewStud... | dbo | 用户 | 2008-12-22 | 12:02:26 |
| sp_IsHaveTestStateStart | dbo | 用户 | 2008-10-28 | 11:10:21 |
| sp_KnowledgeAdd | dbo | 用户 | 2008-10-28 | 11:10:21 |
| sp_KnowledgeByCourseID | dbo | 用户 | 2008-10-28 | 11:10:22 |
| sp_KnowledgeUpdate | dbo | 用户 | 2008-10-28 | 11:10:22 |
| sp_selectClassRoom | dbo | 用户 | 2008-10-28 | 11:10:15 |
| sp_selectClassRoomByUser | dbo | 用户 | 2008-12-22 | 11:51:04 |

Fig. 3. The list of stored procedures

4.2 The Development Tools

The system front uses Microsoft's Visual C #. NET as the main development tools, and the system back-end database employs Microsoft SQL Server 2000. The development tool Visual C #. NET can achieve a seamless link with Microsoft SQL Server 2000 database.

(1) Visual C#.NET

Visual C #. NET is a comprehensive set of tools, which can be created XML Web Services and Microsoft.NET applications for Microsoft Windows and Web. The Visual C #. NET is created based on a strong tradition of language C + +. This system uses the Visual C #. NET toolset of Microsoft Visual Studio 2005 integrated development tools, using C # as developing tool language.

(2) SQL Server 2000 database

In the development of this system, the SQL Server 2000 database is a comprehensive, extremely powerful database. It has the advantages of high performance, highly reliable and scalable, which can be used for large-scale online transaction processing, data warehousing and e-commerce fields.

5 The Description of System Interface

The Mathematica math test platform has three clients (students, teachers and administrators), there are a total of 79 interfaces, which are described briefly as follows:

5.1 The Administrators' Port

The administrators mainly manage the test courses and data backup and can browse most information of the system.



Fig. 4. The administrator interface

5.2 The Teachers' Port

The teachers can manage all the knowledge points of the examination curriculum.

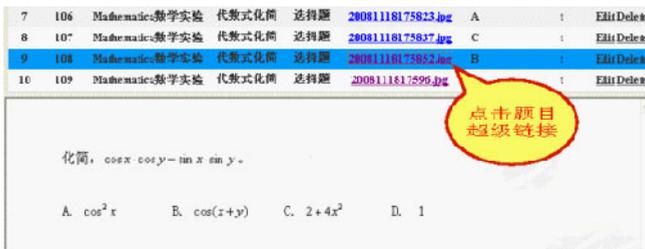


Fig. 5. The teacher interface

5.3 The Students' Port

The students' port is largely the on-line testing interface, students can click the subject lists to browse and answer papers, click on the confirm answer button to save the answer, click the handed out button after answering all the questions.



Fig. 6. The student interface

6 Conclusion

All in all, the construction of MatherMatica mathematics on-line test platform, whose efficiency is very significant. And it promotes educational reform, improve the level of curriculum construction, speed up the teaching of modern process, which will have a positive impact. The MatherMatica mathematics on-line test platform provides teachers with a more flexible form of the examination, to solve the heavy examination affairs. And it improves the scientific proposition, not to do the scoring, statistics, analysis after the test. Computer can automatically graders, sending the results to the database directly to statistics, sorting, summary. The Paperless examination greatly improves the efficiency. At the same time, the paper is composed randomly, there exists no leakage of examination papers and exam cheating, such that the fairness of the examination and the security of answers are effectively guaranteed. System will evaluate the student's answer. The students can get the results at short notice, combining with their own situation to make a comprehensive evaluation of the subjects. Unlike the traditional marking of examination, the other teachers know the results after scoring. And the evaluation has a significant lag, which is not conducive to student self-evaluation capabilities.

References

1. Li, M.-M., Yi, D.-C.: Research of algorithm of general examination system on opening test library and generating paper and judging paper. *Computer Engineering and Design* 26(10), 2743–2745 (2005)
2. Zhang, P., Wang, J.-Z., Wu, Q., Zou, F., Gao, Y.: Secure examination system based on rough set theory. *Application Research of Computers* 27(03), 1012–1014 (2010)
3. You, X.-E.: Design of Bidirectional Identity Authentication of Examination System Based on Campus Network. *Computer System & Applications* 19(04), 19–22 (2010)



ERP Electronic Sand Table Based on Web Service

Zhang Changsheng¹ and Zou Huasheng^{2,*}

¹ College of Physics & Electronic Information Engineering of Wenzhou University, Wenzhou, Zhejiang, China

² Wenzhou University Oujian College, Wenzhou, Zhejiang, China
zoufan99@163.com

Abstract. In this paper, the designing thought of a parameter-driven teaching platform of ERP electronic sand table is introduced that established by Web Service technology under the network environment. At the same time, the architecture and main flow of this platform are described in detail.

Keywords: ERP, electronic sand table, web service, parameter-driven.

1 Introduction

The ERP sand table course is the business management skills training courses that collects the set of informative, interesting, confrontational together. The ERP sand table courses requires a combination of manual sand table or electronic sand table to simulate business operations, where the ERP manual sand table is the teaching aid with sand table plane figure and other chips (money, raw materials .etc objects); while ERP electronic sand table is a set of software platform that teachers and students can simulate the business operations on a computer through the network. In ERP electronic sand table, it is usually display various of the routine departments in the form of a window, and often there are some animations to enhance the fun. Relative to hand sand table, there are more people anticipating in electronic sand table. What's more, teachers are to control students more easily, which can prevent student cheating. At present, the common ERP electronic sand tables are almost based on JAVA or. NET technology to achieve the B / S structural software systems. There exists some problems in these electronic sand tables. For example the rate is slow, the configuration is difficult and market environment variables are not flexible to configure in the ending of teachers and students. Therefore, in this paper, the designing project of teaching platform based on ERP electronic sand table by making use of Web Service technology and parameter-driven technology under the network environment is proposed. The ERP electronic sand table is easier to install arrangement and the realistic market environment can be simulated by the parameter-driven mechanism.

* Corresponding author.

2 The Designing Thought of ERP Electronic Sand Table

In this paper, the most important feature of the proposed ERP electronic sand table is that the electronic sand table platform of a high simulation market based on full parameter-driven is created in the network environment by using Web Service technology.

The ERP electronic sand table teaching platform can be used with hand sand table, it can also be used separately as an electronic sand table. Each group of students prepare a student client computer, which can operate with manual sand table in synchronization, so as to prevent student cheating or an operation error caused by the non-familiar with rules. Teachers can monitor operations of each group completely at the teachers' ending, and count various statistical data and reports. This platform is under the premise of not reducing the effect of teaching groups, using the Web Service technology to achieve accurate and efficient data collection and evaluation information, such that the ERP sand table model focuses on more strategy studies and the competitive will become fiercer.

Overall, the basic feature of ERP electronic sand table based on the parameter-driven is that it operates the database server directly on the network through the Web Service technology. And sending information to the student side in the form of template file. In the students' ending it select the Windows Forms interface, so it has a good interaction. For the frequent configuration information that need to be obtained from the server, the competition rules system is in the form of template file to send information from the server to the student side, which can reduce the network data transmission and increase the response speed of the student side. Teachers can modify the basic parameters simply at the ending, so as to control the competition rules and competition intensity. Compared to other methods of developing ERP electronic sand table teaching aid, this system has such features: simple, convenient, fast response, high simulation authenticity market, easy to develop and use.

3 The Summarize of ERP Electronic Sand Table System

The ERP electronic sand table is to achieve an electronic sand table teaching platform based on the idea of parameter-driven in a network environment through Web Service technique.

3.1 The Architecture Structure of System

In the development of this system, a three-layers architecture is employed. The three layers are mainly the logical three-layers. Namely, data access layer, business logic layer and presentation layer (user interface layer).

Data access layer mainly establishes linking with Microsoft SQL Server 2000 database management system, to implement operations functions on the data in the database. Such as query, delete, modify, etc. The business logic layer is the middle layer between data access layer and presentation layer, whose internal contains a large number of business rules. Meanwhile, the system in the business logic layer applies Web Service technique, dealing with business rules, verifying the legitimacy by Web

services, so as to make the data skip firewall easily by direct transmission. Thus the system can expose some data processing functions to the client application conveniently. While the presentation layer as a input window for teachers and students, that responses to the operations of user-level quickly.

3.2 The Flow of System

The main **users** of the system are teachers and students, teachers are responsible for setting class information and competition parameters. The system automatically generates competition configuration template according to the parameters set by teachers, and on the basis of this standard to generate competition rule files and each initial state data. When after initialization of competition rule files and each group data, students can begin the exercise simulation of sand table, students are to operate in accordance with the order of teachers' end issued the task order. The teachers' end can monitor the operations of each group and analysis the racing data after the competition ends.

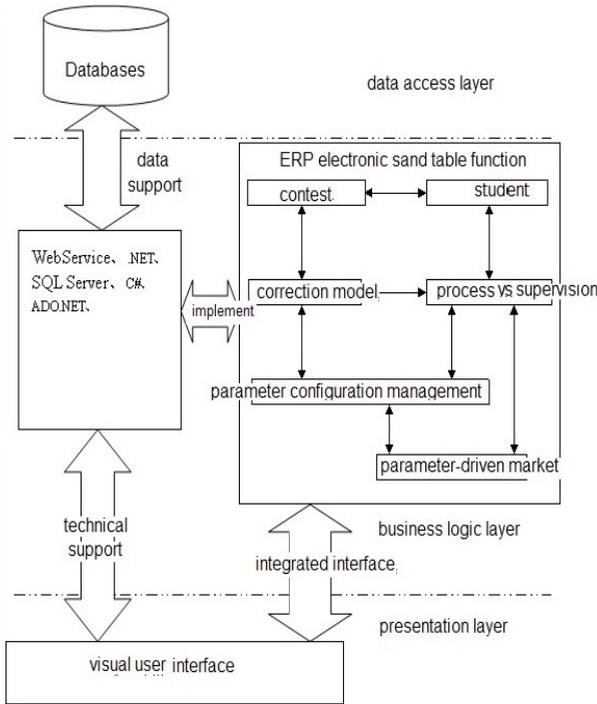


Fig. 1. The distribution figure of system function and relative techniques



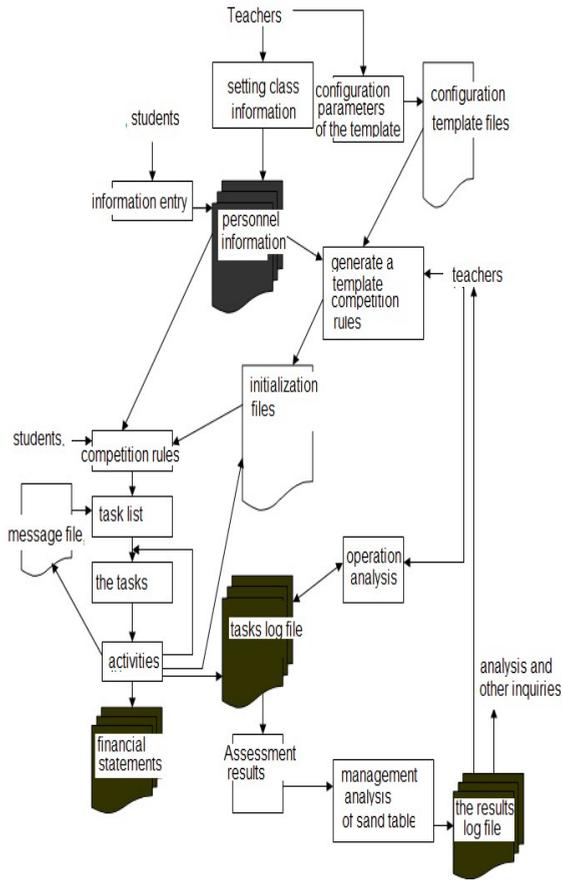


Fig. 2. The flow of system

The description of flow is shown as follows:

- ① After the teachers login, the classes basic information is added and class configuration parameters are set. At the same time, the parametric contest template is configured by teachers, where the parametric contest template reflects the designing idea based on the parameter-driven in teaching platform electronic of this sand table.
- ② After the teachers configure the parametic competitive template, the system generates the configuration template file, and then the system generates the competitive rules of the game and the initial state file.
- ③ At first students add their own login information in the class that teachers have set up before, and set their own role.
- ④ Then students log the system on a specific role again to start model the enterprise's running activities. Students need to follow the steps that required in the task list to operate activities, while teachers and students can exchange the information of user-level.



⑤ When each group completes the simulation of all the enterprise's operational year, the system will calculate automatically the result of each group after the teachers input the command of performance assessment and the performance in the market, financial, growth of each group, then save the results as a data file .

4 The Introduction of Core Technique

4.1 The Parameter-Driven Technique

The parametric designing idea is used in the development of ERP electronic sand table's teaching platform, which requires a parameter-driven mechanism. This parameter-driven mechanism operates on parameter data based on the ERP sand table rules. Through this parameter-driven mechanism, a number of core data of the sand table can be parametric modified. However, in the process of modifying, we must also satisfy the constraints of the sand table. It needs to have a driven- means that has correlation between the bindings to restrain the linkage. Teachers adjust and control taking consideration on the basis of the comprehensive ability of the trained students. In the sand table rules, a lot of public parameters that can vary independently are only a relatively small number, called as the main parameter or the main constraint; other constraints can be determined by the pre-designed allocation ratio and competitive difficulty degree or can establish relationship with the main constraints, called as secondary constraints.

The main focus on how to extract the necessary parameters of the platform's development, how these parameters can drive the sand table model. It needs to build the mathematical model between the number of simulation enterprises and market competition, and then take comprehensive consideration. The final effect is that a specific market environment is obtained by taking considering of the ability of students and based on the teacher end's parameters setting, so as to conduct a proper competition for students in the appropriate market environment. This system can solve the problems that many students face the same kind of market in exercising; and also resolved the problem that students of different abilities face the same market in exercising.

4.2 The Web Service Technique

(1) The concept of Web Service

Web Service is also called XML Web Service. Web Service can receive the request from which other systems of Intranet or Internet passes over. It is the lightweight and independent communications technology. Its available through SOAP provides software services on the Web, using WSDL file to describe, and register by UDDI.

① XML : (Extensible Markup Language). The basis of SOAP is for short-term temporary data processing and World Wide Web.

② SOAP : (Simple Object Access Protocol). It is the communication protocol of XML Web Service. When the user find your WSDL description document through UDDI, he can call one or more operations in your built Web services via the SOAP. SOAP is a specification of the XML document form calling the methods, which can support different underlying interfaces, like HTTP or SMTP.

③WSDL : (Web Services Description Language) WSDL file is a XML document that describes a set of SOAP messages and how to exchange these messages, which is generated by the software automatically in most cases .

④UDDI (Universal Description, Discovery, and Integration): UDDI is a new project mainly for Web services provider and users. UDDI is a mechanism according to the description document guiding system to find the appropriate service. UDDI using SOAP messaging mechanism(standard XML / HTTP) to publish, edit, browse and find some registration information. It employs the XML format to encapsulate different types of data and sends to the registry or return the required data by registry.

(2)The advantages of Web Service

The main goal of Web Service is the interoperability of cross-platform. To achieve this goal, Web Service is the standard based on XML (Extensible Markup Language), XSD (XML Schema) entirely and other independent platforms. It is a new platform that creates interoperable and distributed applications. Thus there are many advantages of using Web Service:

①The communication of acrossing firewall

The middle-layer components of system use Web Service technology, you can call directly the middle-layer components from user interface. Calling Web Service, we can directly use the Microsoft SOAP Toolkit or the SOAP client with '.NET', we can also use our own developed SOAP client ending, and then connect it with applications. It not only shortens the development cycle, but also reduces the complexity of code, and can enhance the maintainability of the application.

②The integration of application

The application often needs to obtain data from the running program on a host; or sends the data to the host or other platform applications. Even in the same platform, a variety of softwares produced by different manufacturers often need to integrate. The XML Web services provide the ability of exchanging messages using standard protocols (HTTP, XML, SOAP and WSDL) in a loosely coupled environment. Messages can be structured, with the type and it also can be loosely defined.

③ The integration of B2B

The greatest advantage to achieving the integration of B2B by Web Service is that it can easily achieve interoperability. As long as exposing the business logic, let it be a Web Service, you can make any designated partner call such business logic, regardless of what platform they run on the system, what development language they use, which greatly reduces the time-consuming and cost on the integration of B2B.

④ The reuse of software and data

Web Service allows the reuse of code, the data behind the code can also be reused. Using Web Service, it only need to call a remote Web Service. The function of several applications can also be integrated through the exposure of Web Service. You can easily put all of these functions integrated into your portal site to provide users with a single and friendly interface. You can use the function provided by the third-party Web Service in applications, at the same time, you can provide your own application to others through the Web Service.

⑤ Interoperability

Any Web Service can interact with other Web Services. As the major suppliers support SOAP(Simple Object Access Protocol) protocol, the conversion of trouble is avoided between CORBA, DCOM and other agreements. You can use any language to write Web Service, developers don't need to change development environment, you can use Web Service.

5 Conclusion

From a technical point of view and the development of the system itself, the Web Service with the advantage of the loosely coupled nature of Web services, good encapsulation, and the application integration of cross-platform has attracted the majority of professionals of all ages. It can simplify the integration of system, achieve the connectivity of system and transaction processing quickly and easily. And this implementation does not need to consider the application environment platform, so as to improve the utilization of resource and reduce the cost of development. So far, only the Web Service can integrate the data that dispersed in a variety of systems and information silos unify, which reduce the cycle of information flow greatly, improve the efficiency of business processes and bring great benefits. In this paper, the middle layer of ERP sand table teaching platform is employed the Web Service technology, a Web Service project is established. The Web Service function is provided outside the Web Services project independent of other layers. In the middle layer the Web Service technology is used to build the system's business logic layer. Making use of the advantages of the Web Service technology to expose a variety of business logic, so as to provide data interface for the layer and other applications of the ERP electronic sand table, so that the existing ERP electronic sand table is easier to install and configure. At the same time, the designing idea based on parameter driven to allows teachers to set the main parameters to adjust the simulation state of the market before the simulation operation, so as to create more flexible market environment for students, thereby to enhance the teaching effect of ERP sand table.

References

1. Xu, D.-M., He, H.: Introducing ERP Sand Table Simulation and Improving Practice Teaching. *Experiment Science & Technology* 5(5), 97–98 (2007)
2. Wang, Q.-Z., Ying, L.-J.: Teaching Research on ERP Sand Table Training Course. *Journal of Zhejiang Business Technology Institute* 7(2), 51–53 (2008)
3. Huang, M.-J.: Exploring and Practising About ERP Sand Table Simulation Course. *Journal of Wuxi Institute of Technology* 4(4), 46–47 (2005)
4. Ruth, A.: Healing and Transfromationing Sand play: Creative Processes Become Visible. *Open Couert*, 73–76 (June 1993)

A Mathematical Modeling of the Best Design of Pipeline Construction*

Ruan Jing**, Lin Bin, and Xiang Haifei

Wenzhou Vocational & Technical College, Wenzhou, Zhejiang, China
ruanjing1979@126.com

Abstract. According to the distance between the two refineries on the same side of the railway and their distance from the railway, this paper gives out the design of the lest cost on laying pipelines, considering the cost of shared pipeline and non-shared ones is same or not. Considering the different cost of laying pipelines between urban and suburban, we establish a more realistic model and a multi-objective optimization model.

Keywords: Pipe laying costs, Fermat point, symmetric method.

1 Introduction

At present, the transportation mode of distance oil mainly depend on pipeline in China, there are some advantages in pipeline transmission, such as big traffic volume, meanwhile, there are some disadvantages, such as high investment, the oil is carried to the oil refinery by pipelines. If we try our best to build common pipelines which efficiently reduce cost, so we must consider the case of the common pipelines.

In the paper, for current situation and development prospect of oil refining plant, we take the two oil refinery be build at the same side of the railway as example. At the some time, we need to build the station at the railway lines. To take the transportation of finished oil for an example, the pipeline links the refinery with the port, according to the different situations of the distance between the two refineries and the railway line. We should consider the two cases for the cost of common pipeline and that of non-common pipeline. Given taking into account the two cases of different cost, the cheapest design scheme of building pipeline is proposed, what's more, taking into account the impact cost of pipeline construction, such as change fee and compensation fee, we have also established a construction mode of the difference between city and suburban. In addition, taking into account environmental pollution, personal safety and other factors, based on this situation, we also present multi-objective optimization model in pipeline construction.

* Foundation items: The Project of Wenzhou Vocational and Technical College (WZYJG1018).

** Corresponding Author: Ruan Jing (1979—), female, Master, Lecturer.

Research Interests: Mathematical modeling and its applications.

These types of models have a certain universality and applicability; we hope to put forward the cheapest general mathematical models and methods in pipeline construction.

2 Setting-Up and Solution of Model

2.1 The Same Model between Common Pipeline and Non-common Pipeline

Let A,B point be the two refineries, the H point on behalf of the railway station on the L line, the P point represents the convergence point of two oil refineries, coordinates were about $(x, y) (x \geq 0, y \geq 0)$, finished oil be transported by the convergence point in common pipeline to the H station, such as Figure 1.

Because the fees of common pipeline is equivalent to that of non-common pipeline, then how to build cheapest cost of pipeline construction is equivalent to finding the P point and H point at the railway, so such making the minimum point $PA + PB + PH$ to the sum distance from P point to A, B, H points.

For this problem, it can be traced to 1640; Pierre de Fermat put forward the following problem: "Given three points A, B, C in the plane, and find that the minimum distance $PA + PB + PC$ ". This is the famous "Fermat's problem" In the mathematic history. In particular, when the three points are located at non-collinear, then the smallest point $PA + PB + PC$ is called the Fermat point ΔABC .

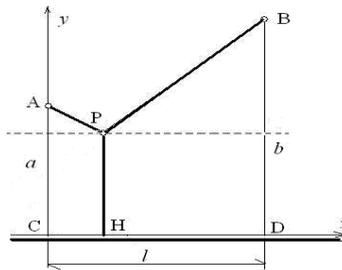


Fig. 1. The oil refinery and distribution map of railway lines

Theorem 1. For the three vertices in a triangle, if the sum of distance of each point minimized, then the point is called the Fermat point of the triangle.

(1) If the three interior angles of a triangle are less than 120° , then the connection point from the three distance line can divide equally the Zhoujiao of the Fermat point. Therefore, Fermat point of the triangle, also known as the conformal center of the triangle.

(2) If an interior angle of the triangle is not less than 120° , then this is the vertex of the obtuse angle is the smallest distance point.



Let F be the total cost of pipeline construction, the fee of construction for the $k \cdot 10000$ yuan/km, so the Fermat's model is as follows:

The objective function: $\min F(x, y) = k \cdot (PA + PB + PH)$
 $= k \cdot [\sqrt{x^2 + (y-a)^2} + \sqrt{(x-l)^2 + (y-b)^2} + y]$ s.t. $\begin{cases} 0 \leq x \leq l, \\ 0 \leq y \leq a. \end{cases}$

the solution of the above model solution, in paper[1] has made a complete answer and proof, Table 1 be succinctly expressed the solution of pipeline construction under the same cost of common pipeline and non-common pipeline, the corresponding graph is as showed Figure 2.

Table 1. The solution of pipeline construction under the same cost of common pipeline and non-common pipeline

| l Scales | P coordinates | the cheapest cost |
|-------------------------------------|---|--------------------------------------|
| $0 < l \leq \sqrt{3}(b-a)$ | $P(0, a)$ | $k \cdot [a + \sqrt{l^2 + (a-b)^2}]$ |
| $\sqrt{3}(b-a) < l < \sqrt{3}(a+b)$ | $P(\frac{\sqrt{3}(a-b)+l}{2}, \frac{a+b-\sqrt{3}l}{6})$ | $k \cdot \frac{a+b+\sqrt{3}l}{2}$ |
| $l \geq \sqrt{3}(a+b)$ | $P(\frac{al}{a+b}, 0)$ | $k \cdot \sqrt{l^2 + (a+b)^2}$ |

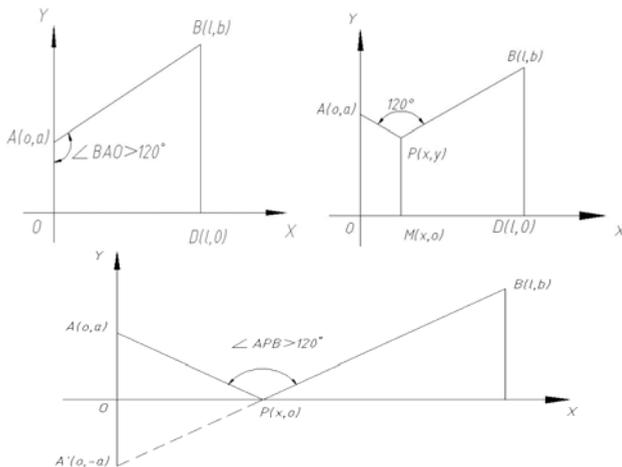


Fig. 2a. $0 < l \leq \sqrt{3}(b-a)$ **Fig. 2b.** $\sqrt{3}(b-a) < l < \sqrt{3}(a+b)$

Fig.2c. $l \geq \sqrt{3}(a+b)$



2.2 The Model of the Different Cost of Common Pipeline and Non-common Pipeline

Because the construction model of the different cost of common pipeline and non-common pipeline is similar to the above-mentioned construction under the same cost of common pipeline and non-common pipeline, so we can build a generalized Fermat models, the fee of non-common pipeline construction for the $k \cdot 10000$ yuan/km, that of non-common pipeline construction for the $\lambda k \cdot 10000$ yuan/km, where $1 \leq \lambda < 2$ should be met in the actual situation, then the problem of the cheapest cost of pipeline construction is equivalent to solving the following model for the promotion of the Fermat .

The objective function: $\min F(x, y) = k \cdot (PA + PB) + \lambda k \cdot PH$

$$= k \cdot [\sqrt{x^2 + (a - y)^2} + \sqrt{(c - x)^2 + (b - y)^2}] + \lambda k \cdot y \quad (1) \quad s.t. \begin{cases} 0 \leq x \leq l, \\ 0 \leq y \leq a. \end{cases}$$

For the solution of the above mode, We can use “geometric symmetry method”, straight line L_1 made by P point parallel the x axis, if A point is A' symmetry points on a straight line, then A' coordinates is $(0, 2y - a)$. If the optimal solution for the above model P point, we can find that at the three A' 、 P 、 B point are located at the same collinear as shown in Figure 3.

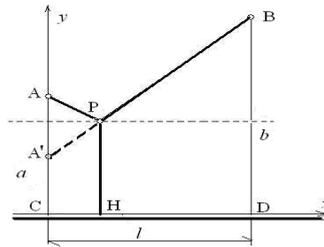


Fig. 3. Geometric symmetry method

The objective function:

$$\min F(y) = k \cdot A'B + \lambda k \cdot PH = k \cdot \sqrt{l^2 + (2y - b - a)^2} + \lambda k \cdot y \quad (2)$$

Find the minimum objective function, based on the extreme value theorem, there can be

$\frac{dF}{dy} = 0$, such that:

$$y = \frac{1}{2} \left(a + b - \frac{\lambda l}{\sqrt{4 - \lambda^2}} \right) \quad y = \frac{1}{2} \left(a + b + \frac{\lambda l}{\sqrt{4 - \lambda^2}} \right)$$



Where $y = \frac{1}{2}(a+b + \frac{\lambda l}{\sqrt{4-\lambda^2}}) > \frac{1}{2}(a+b) > a$, According to the Theorem 1, there are P point as the same to A point, for P point in $\triangle ABM$, there are $0 < y < a$, that is $0 < \frac{1}{2}(a+b - \frac{\lambda l}{\sqrt{4-\lambda^2}}) < a$, such that;

$$\frac{(b-a)\sqrt{4-\lambda^2}}{\lambda} < l < \frac{(b+a)\sqrt{4-\lambda^2}}{\lambda}$$

The binary objective function (1), the partial derivatives about x and let the result equal to zero, we can get

$$x = \frac{l(a-y)}{a+b-2y} \text{ or } x = \frac{l(a-y)}{a-b} \text{ (} x < 0 \text{ delete)}$$

let $y = \frac{1}{2}(a+b - \frac{\lambda l}{\sqrt{4-\lambda^2}})$ input $x = \frac{l(a-y)}{a+b-2y}$, we can get

$$x = \frac{l}{2} - \frac{(b-a)\sqrt{4-\lambda^2}}{2\lambda}, \text{ then } P \text{ coordinates is } (\frac{l}{2} - \frac{(b-a)\sqrt{4-\lambda^2}}{2\lambda}, a+b - \frac{\lambda l}{\sqrt{4-\lambda^2}})$$

The optimal solution of the above model of pipeline construction as showed Table 2.

Table 2. The optimal solution of the above model of pipeline construction

| l Scales | P coordinates | the cheapest cost |
|---|---|---|
| $0 < l \leq \frac{(b-a)\sqrt{4-\lambda^2}}{\lambda}$ | $P(0, a)$ | $k \cdot [\lambda a + \sqrt{l^2 + (a-b)^2}]$ |
| $\frac{(b-a)\sqrt{4-\lambda^2}}{\lambda} < l < \frac{(b+a)\sqrt{4-\lambda^2}}{\lambda}$ | $P(\frac{l}{2} - \frac{(b-a)\sqrt{4-\lambda^2}}{2\lambda}, a+b - \frac{\lambda l}{\sqrt{4-\lambda^2}})$ | $k \cdot [\frac{\lambda(a+b)}{2} + \frac{4l-\lambda^2 l}{2\sqrt{4-\lambda^2}}]$ |
| $l \geq \frac{(b+a)\sqrt{4-\lambda^2}}{\lambda}$ | $P(\frac{al}{a+b}, 0)$ | $k \cdot \sqrt{l^2 + (a+b)^2}$ |

3 The Extension of Construction Model

For more complex cases, the specific location of two oil refinery is shown in Figure 4, where A factory in the rural areas (Figure I in the region), B Factory is located in urban areas (Figure II in the area), two regions line with the dashed line in Fig. each letter of

the distance (unit: km). at the same time, we consider pipelines construction in urban areas need to increase the compensation for demolition and construction and other additional costs, then how can the cheapest cost of pipelines construction?

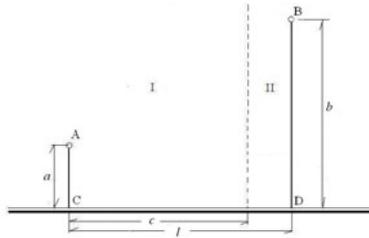


Fig. 4. Refineries and distribution of suburban rail lines

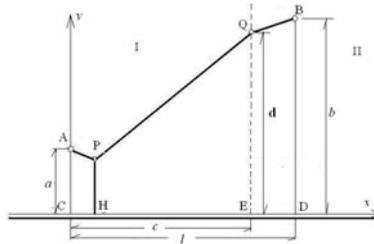


Fig. 5. The figure of pipeline construction

Given the fee of pipeline relocation and that of project compensation and other expenses for $f \cdot 10000$ yuan/km, pipeline from the starting point B of intersection with the suburban line as $Q(c, z)$, as showed in Figure 5. Also, consider the part of the pipeline rates, ranging from the situation. Based on the problem, Let k_1, k_2, k_3, k_4 be the rate for the section of AP, PQ, PH, BQ pipe, respectively, in order to make the cheapest cost of pipeline, we have established the following model.

The objective function: $\min F(x, y, z) = k_1 \cdot PA + k_2 \cdot PQ + k_3 \cdot PH + (k_4 + f) \cdot QB$

$$= k_1 \sqrt{x^2 + (a - y)^2} + k_2 \sqrt{(c - x)^2 + (z - y)^2} + k_3 y + (k_4 + f) \sqrt{(l - c)^2 + (b - z)^2}$$

For example, suppose $a = 5, b = 8, c = 15, l = 20$, the For example, the construction cost of oil transmission pipeline is 56,000 yuan per kilometer in A Factory, the cost of finished oil transportation for the B Factory is 60,000 yuan per km, the cost of common pipeline is 72,000 yuan per kilometer, the cost of demolition and additional construction compensation is 215,000 yuan per kilometer, then the cheapest model of pipeline construction is given.



The objective function:

$$\begin{aligned} \min F &= 5.6 \times AP + 7.2 \times PH + 6.0 \times PQ + (6.0 + 21.5) \times BQ \\ &= 5.6 \times \sqrt{x^2 + (y-5)^2} + 7.2 \times y + 6.0 \times \sqrt{(x-15)^2 + (y-d)^2} + 27.5 \times \sqrt{5^2 + (8-d)^2} \end{aligned}$$

For the ternary function model in The above model, computing the optimal d value and x value have made the minimum objective function value, if calculation can be to meet in the function. We can get the value d 、 x 、 y ;

Using LINGO, Mathematica or Matlab software programming, we get $Q(15, 7.2795)$ and $P(6.733378, 0.138901)$, such that the minimum cost F is 2,519,690 yuan.

4 Conclusion

In this paper, the mathematical model the optimal solution of pipeline construction depends only upon economic benefit in the process of pipeline construction. It is design aspect in the management of pipelines, In order to make the maximum economic benefits of the pipeline, meanwhile, we need to consider the quality of the equipment, technology operation processes operating processes, quality of construction and execution management, and many other aspects.

References

1. Chu, B.N.: A generalization of Fermat point in triangle. Secondary Mathematics Teaching 5, 32–33 (2006)
2. Lin, J.L.: Some Notes on Fermat's Point. Journal of South China University of Technology 32(1), 93–95 (2004)
3. Xie, J.X., Xue, Y.: Optimization Modeling with LINDO/LINGO software. Tsinghua University Press, Beijing (2005)
4. Hu, Y.Q.: Operational Research Tutorial. Tsinghua University Press, Beijing (2003)

Trust-Based Privacy Authorization Model for Web Service Composition

Jun Zheng¹, Zhiqiu Huang², Jun Hu², Ou Wei², and Linyuan Liu²

¹Quality Engineering Technology Center, China Aeronautics Polytechnology Establishment, Beijing, China

²College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China
liulinyuang@163.com

Abstract. The rapid growth of web applications has prompted increasing interest in the area of composite web services that involve several service providers. This increased use of composite services has meant that more and more personal information of consumers is being shared with web service providers, leading to the need to guarantee the privacy of consumers. This paper proposes a trust-based privacy authorization model for service composition, it uses privacy authorization policies to specify the privacy privileges of services. Then it utilizes the trust relationships among services to make privacy authorization decisions. Comparing to the traditional privacy access control approaches, this model can make the fine-grained authorization decision, thus efficiently protecting consumers' privacy.

Keywords: Privacy, Access control, Web Service composition, Trust.

1 Introduction

The past decade has witnessed a phenomenal growth in the number of users who routinely use the web to obtain information, conduct research, or carry out financial transactions [9]. While the ability of the web to provide customized information and financial services has boosted personal and business productivity, it has raised significant concerns regarding consumer information privacy. Once personal information is disclosed online, it may be abused or accessed without authorization, thus increasing the risk of illegal collection, usage and disclosure of personal sensitive data [5]. Hence, the issue of protecting the data privacy of consumers has become of crucial.

Privacy is the right of individuals to determine for themselves when, how, and to what extent private information is communicated to others [3,6]. Privacy concerns are fueled by an ever increasing list of privacy violations, ranging from privacy accidents to illegal actions [4,10]. In order to accomplish the business goals, a service consumer has to release some private data to support the collaborative work; on the other hand, in order to meet the needs of security, a strict control mechanism is required to protect these private data [11].

Currently, there are many researches about the privacy access control approaches in service composition. However, most of them use the method of “all-or-nothing” to make privacy authorization decision, namely, it authorizes either all privileges or none. But, in fact, due to the untrustworthy nature of the Web service, a set of services with the same functions may have different privacy disclosure risks because of their different trust degrees. Take the shopping scenario for example, the seller with the higher trust degree is more trustworthy in protecting the privacy of customers, on the contrary, the seller with the lower trust degree may disclose the privacy information to the other unauthorized party for their own interests.

To overcome the above drawbacks, this paper presents a trust-based privacy authorization model for Web service composition, this model utilizes authorization policies to specify the privacy privileges of services, introduces trust relationships among services into composition design. The composition makes the decision on privacy authorization according to the trust degree of the services.

The remainder of this paper is arranged as follows. Section 2 gives the method of computing the trust degree of services. Section 3 proposes a trust-based Web service privacy authorization model. Section 4 analyzes the privacy authorization process. Closely related work and comparison are reported in Section 5. Section 6 concludes the paper with proposals for future work.

2 The Computation of the Services' Trust Degree

The participants of the services composition are a set of strange services from different organizations, so they lack the necessary trust between each other. Therefore, when the consumers disclose the private data to the providers, it requires that the providers have a certain trust degree. The consumers release privacy privileges according to the trust degree of providers, the higher the trust degree is, the more the privacy privileges will be released.

Definition 1: Trust Degree. Trust degree is a measure indicating the trustworthy degree of the services. Let S be a set of finite services of a service composition. For any a service $s \in S$, funtion $ftd(s)$ is used to obtain its trust degree, where $ftd(s) \in [0,1]$, 0 indicates that it is untrustworthy, 1 indicates trustworthy. The larger the value is, the higher the trust degree is.

The trust degree of the services is evaluated by the users. The users give scores to the trust degree of the services after using services. The computation process can be clarified into two steps: firstly computing the score a user gives to a service, then combining the scores that all of users give and thus obtaining the whole trustworthy value of the services. The specific steps are as follows:

(1) Computation of the local trust degree

In order to compute the local trust degree of a service, we firstly let the score the user i gives to service j be $r_{i,j}$, $r_{i,j} \in [0,1]$. When the user i considers that the service j leaks much information, the score $r_{i,j}$ would be relatively low. Otherwise, if the user i considers that there is little information leaked, the score $r_{i,j}$ would be relatively high. Then we introduce a history trust degree $R_{i,j}$, which refers to the score the users gave to

the service j in the past. Based on $r_{i,j}$ and $R_{i,j}$, we use the following formula to compute the local trust degree of the user i to the service j :

$$R_{i,j} = \partial R_{i,j} + (1 - \partial)r_{i,j} \quad (1)$$

Where, the parameter ∂ is a history factor, $\partial \in (0,1)$ shows the proportion of the history trust degree score in the computation of the local trust degree. If ∂ approaches 1, it indicates that the influence of the history trust degree score on the computation of the new trust degree is great, while when ∂ approaches 0, it indicates that the latest trust degree score play the center role.

(2) Computation of the globe trust degree

Let td be the globe trust degree of the service j , through combining the local trust degree scores other services gives to the service j , we can compute the globe trust degree of the service j . The formula is as follows:

$$td = \sum_{k=1}^N R_{k,j} / N. \quad (2)$$

Through the above steps, we can obtain the trust degree of a service and the value of the trust degree is sure to be between 0 and 1.

3 Trust-Based Service Composition Privacy Authorization Model

In order to fine-grained protection the privacy of consumers, we set a trust level for every privacy privilege. Only when the trust degree of a service is higher than this trust level, can it access the privilege. Trust level actually limits the qualification of the services to access privilege.

When the services receive or send a message, they may request to access not only one privilege, but a set of privileges. Hence, we use the authorization policies to specify a set of access privileges granted to the services. Each authorization policy has a corresponding authorization policy matrix, for short, authorization matrix, where the row vectors of the matrix denotes a set of privacy privilege types, the column vectors of the matrix denotes a set of private data objects, the elements of the matrix denote the trust level required by granting a type of privilege of a data object to the services. Fig. 1 is the description about an authorization matrix.

| | | | | |
|----------|-----------|-----------|----------|-----------|
| | o_0 | o_1 | \dots | o_n |
| t_0 | tl_{00} | tl_{01} | \dots | tl_{0n} |
| t_1 | tl_{10} | tl_{11} | \dots | tl_{1n} |
| \vdots | \vdots | \vdots | \ddots | \vdots |
| t_n | tl_{n0} | tl_{n1} | \dots | tl_{nn} |

Fig. 1. The description of authorization matrix

Definition 2: Authorization Policy. Let S be a finite set of services, M be a finite set of messages, OP be a finite set of operations, O be a finite set of private data objects and T be a finite set of privacy privilege types, a authorization policy is defined as $plcy := \langle s, op, am \rangle$, where:

- $s \in S$ is a service;
- $op \in OP$ is an operation of s , it may be a sending operation or a receiving operation of the message msg , which are defined as $type(op) = msg!$ or $type(op) = msg?$ respectively, where $msg \in M$;
- am is an authorization matrix, it is formed by $|T|$ rows and $|O|$ columns. The elements of the matrix $am[i, j]$ ($0 \leq i < |T|$, $0 \leq j < |O|$) denote the trust level required by granting the privilege with the data object as j and the privilege type as i to service s , $am[i, j] \in [0, 1]$, where 0 indicates to authorize this kind of privilege and 1 indicates not.

Based on the definition of authorization policy, we can present the authorization specification of a service composition as follows.

Definition 3: Authorization specification. A authorization specification is defined as $as := \langle S, OP, AM, PLCY, F_{am} \rangle$, Where:

- S is a finite set of services;
- OP is a finite set of operations;
- AM is a finite set of authorization matrices;
- $PLCY \subseteq S \times OP \times AM$ is a finite set of authorization policies. For any authorization policy $(s, op, am) \in PLCY$, it denotes that the service s has the privileges granted by authorization matrix am for operation op , where $s \in S$, $op \in OP$, $am \in AM$;
- $F_{am}: (S, OP) \rightarrow AM$ is a function that takes as input a $s \in S$ and a $op \in OP$ and returns the authorization matrix am , where $am \in AM$.

4 Privacy Authorization Enforcement Process

Based on the above privacy authorization model, now we can describe the architecture of the privacy authorization enforcement system. This system involves four components: authorization control manager (ACM), authorization decision maker (ADM), authorization policy repository (APR) and trust repository (TR). When a role requests for the privacy privilege, it firstly sends the request to the ACM, which includes the id of service and the name of operation. After the ACM receives the request, it asks the ADM to make the decision. The ADM firstly finds the APR according to the service id and the operation name and gets the correspondent privacy authorization matrix. Then it finds the TR and gets the correspondent trust degree of the service. Finally it makes the authorization decision according to the authorization matrix and the trust degree of service, and sends the authorization result to the ACM. The ACM sends the service about the result that whether the request is accepted or denied. The privacy authorization enforcement system is as follows:

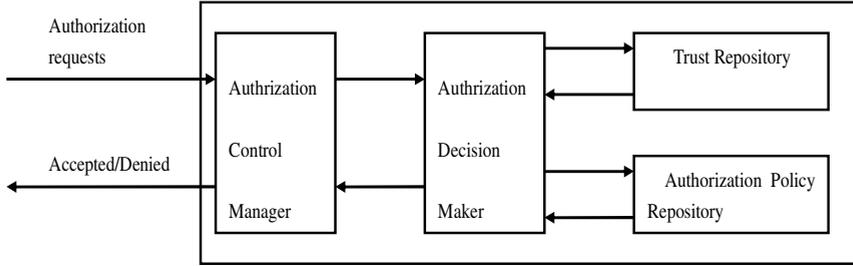


Fig. 2. The architecture of privacy authorization enforcement system

In the above authorization process, the authorization decision-maker makes the authorization decision according to the authorization matrix and the trust degree. The specific steps are as follows:

- 1) Let op is an operation of service s , set privacy-satisfied mark $Psatisfied$ as 0, where $Psatisfied$ is an integer variable.
- 2) According to the op and s , we are obtains the corresponding authorization matrix am and iteratively takes the trust threshold of every privacy privilege out of am ; Then compares the trust threshold with the trust degree of s , if the trust threshold is lower than or equal to the trust degree, then $Psatisfied$ is decreased by 1.
- 3) Finally it checks whether the value of $Psatisfied$ is 0, if yes, then the authorization request is accepted, otherwise it is denied.

The following is the specific description of the privacy authorization decision algorithm.

Algorithm 1: The privacy authorization decision

Input: s ; op ;

Output: True; False;

$d = f_{tl}(s)$; //get the trust degree of the service

$am = f_{am}(s, op)$; //get the authorization matrix

For(each t in $am.T$) { //make the authorization decision

For(each o in $am.O$) {

If($pm[t,o] \leq d$) { $Psatisfied := Psatisfied - 1$ }

}}

If($Psatisfied == 0$) {Return true;} Else {Return false;}

5 Related Work

In recent years, increasing attention has been paid to privacy-aware technologies and mechanisms to protect data privacy in the Web Services. The World Wide Web Consortium is advocating P3P (Platform for Privacy Preferences) [7] as the standard format for expressing data-collection and data-use practices on the web. The XACML (Extensible Access Control Markup Language) [8] is an XML-based access control

language. It provides an application-independent policy language which enables the use of arbitrary attributes in different types of policies, including privacy policies.

Many other proposals can be found in current literature. For instance, Byun et al. [2] proposed a purpose-based access control framework extending RBAC along the lines given by Agrawal et al. [1] in their Hippocratic database systems. The aim of this framework is to enforce privacy promises encoded in privacy policy languages in database management systems. The framework is based on the notion of intended purpose and the notion of access purpose. The problem of those methods is that they just focus on the privacy within single organizations rather than in a distributed system. Another question of these models is that they use the “all-or-nothing” authorization decision mechanism, which may cause the unexpected privacy leakage.

Different from previous models, the novelty of our access control approach mainly includes two aspects: Firstly, our approach is constructed according to the collaborative environment of Web services. It is not limited in the inside of the organizations. Secondly, our approach has introduced the trust relationships between services and made the fine-grained authorization based on the trust relationships. In our approach, the trust level of every piece of privacy privilege has been set. Only when the trust degree of the service is higher this level, can it access the privilege.

6 Conclusions

Based on the demand for fine-grained privacy authorization, this paper proposes a trust-based privacy authorization model. First, it uses privacy authorization policies to specify the privacy privileges of service composition. Second, it proposes the method of computing the trust degree of services and utilizes the trust degree of services to make the authorization decisions. The trust-based authorization mechanism provides a feasible and effective method to protect the privacy privileges with fine granularity.

The protection of data privacy in Web services involves research in other branches. Several improvements to our current work can be foreseen. First, this paper considers the collection, usage and disclosure of private data. It has not yet considered the time property of privacy data. Therefore, more work can be done to investigate the time factor relevant to the retention time of private data. Second, in order to meet the personalization needs of service consumers for the protection of private data, the privacy policies may have to change dynamically. It requires the service composition to possess an active adapting and adjusting capability that provides suitable service composition in a timely and dynamic manner.

References

1. Agrawal, R., Kiernan, J., Srikant, R., Xu, Y.: Hippocratic databases. In: Proceeding of the 28th International Conference on Very Large Data Based, Hong Kong, China, pp. 143–154 (2002)
2. Byun, J.W., Bertino, E., Li, N.: Purpose based access control of complex data for privacy protection. In: Proceedings of 10th ACM Symposium on Access Control Models and Technologies, pp. 102–110. ACM Press (2006)

3. Guarda, P., Zannone, N.: Towards the development of privacy-aware systems. *Information and Software Technology* 51(2), 337–350 (2009)
4. Karjoth, G., Schunter, M., Herreweghen, E.V.: Translating privacy practices into privacy promises -how to promise what you can keep. In: *Proceeding of the 4th IEEE International Workshop on Policies for Distributed Systems and Networks*, pp. 135–146 (2003)
5. Liu, L., Huang, Z., Zhu, H.: Role-based consistency verification for privacy-aware Web services. In: *Proceeding of the 7th Int. Symposium on Collaborative Technologies and Systems*, Irvine, CA, USA, pp. 399–407 (2009)
6. Rezgui, A., Bouguettaya, A., Eltoweissy, M.: Privacy on the Web: facts, challenges, and solutions. *IEEE Security & Privacy* 1(6), 40–49 (2003)
7. W3C. The Platform for Privacy Preferences 1.0 Specification, P3P (2002), <http://www.w3.org/P3P>
8. OASIS Standard. Extensible access control markup language version 2.0 (XACML), OASIS Standard (2005), http://docs.oasis-open.org/xacml/2.0/access_control-xacml-2.0-core-spec-os.pdf
9. Xu, W., Venkatakrishnan, V.N., Sekar, R., Ramakrishnan, I.V.: A Framework for building privacy-conscious composite Web services. In: *Proceeding of the International Conference on Web Services*, pp. 655–662 (2006)
10. Yee, G., Korba, L.: Privacy policy compliance for Web services. In: *Proceedings of 2004 IEEE International Conference on Web Services*, pp. 158–165 (2004)
11. Zhang, J., Chang, C.K., Zhang, L.J., Hung, P.C.K.: Toward a service-oriented development through a case study. *IEEE Transaction on Systems, Man, Cybernetics, Part A* 37(6), 955–969 (2007)

Talent Training of Software Testing Based on University-Enterprise Cooperation

Zhang Shanwen and Zhou Zhengguang

Department of Engineering and Technology, Xijing University, Xi'an 710123, China
gguai1012@163.com

Abstract. The university-enterprise cooperation is one of the win-win patterns of the universities and enterprises, which not only makes both schools and enterprises play their respective advantages, but also trains talents with social and market need. In this paper, we introduce the important of school-enterprise cooperation on software testing, and propose some measures of school-enterprise cooperation, point out the questions of university-enterprise cooperation.

Keywords: Software testing, University-enterprise cooperation, Talent training.

1 Introduction

Compared with the developed countries, the strength of domestic software testing is still relatively weak which matches the overall economic development of the domestic currently, because the current ratio of demand and supply of domestic software testing is lower than foreign countries, people's service demand to satisfy themselves is not so strong, they are lack of certain knowledge of new things and new technologies. Software testing is taken as a very important project in many software companies in Europe and America, but more software testing in domestic enterprises is done some simple test by the developer or the customer only after the development when engaged in software development, a professional team is seldom used for testing, which is one of the reasons the majority of China's software quality level is relatively low. Ten years ago, enterprises engaged in software development at home was rare, the competition among similar software products was not fierce, so software quality had not been regarded as importance by software companies. However, with the development of China's software industry, software companies have shot up, competition becomes more and more fierce. At the same time, user consciousness of the software product quality gradually raise, product failures, business loss caused by software quality problems often happened in many companies. In order to win the competition, software companies must improve software quality, and the means to achieve this goal is to test software. Software testing is an emerging field of IT industry in recent years with rapid development. Many software companies have set up software testing department, and software outsourcing companies engaged in testing and third party testing organizations appear ceaselessly.

2 Measure of University-Enterprise Cooperation

Software testing profession is as a "prick, i.e., pick the wrong" role in the software industry. What kind of role is really software testing profession in the software industry? During the 1970s, due to lack of software testing, software project's mortality rate was more than 70% in America, 90% projects exceeded cost goals or overran schedule. Like other products with quality problems, software inevitably has all kinds of loopholes, known as the "bug". If the software has too much "bug", it will cause the computer frequently "crash" and influence the user's normal use. Therefore, in order to ensure the normal function of the software, we need to keep testing the software during the development process. Our colleges and universities have begun to focus on training of software testing profession for its importance. Currently software test Engineer sources consist of three parts: first, former software development professionals; Second, recruited from the university undergraduate or graduate students; Third, recruitment of professional students from training institutions. It is hard for company to recruit Software Test Engineer, for this reason, many colleges and universities have offered courses related to the test in the past years, and promoted the training mode of the university-enterprise cooperation to satisfy the employment of IT business needs. Based on the different cooperation models, Schools and enterprise can choose different operating mechanism, their common goal is to make the laboratory to better serve the teaching, research, production services for greater social benefits. University-enterprise cooperation system fully simulates the actual working scenario of software companies, achieves the role of interactive exercises between department manager and staff, it enables students to work in schools and to obtain practical ability, rather than the second training of practice courses after graduation with high prices. The main operating mechanism consists of the following aspects.

Training the graduating students of professional computer and software as a potential capability of the "software testing engineer," they will thus become comprehensive software talent, not only understand the software programming, but also understand software testing. Students will attend examination after studying practical courses, the National Software Center will give "Software Quality Engineer" certificate to those who pass the examination, their name can be published in the "China Software Quality Network", member companies regularly are informed of the interview and the enroll. In this way, the enterprise can find the right people.

Training the teachers of Engineering Training Center for "Colleges and Universities Software Training system" freely, until the teachers completely grasp the courses and may lead students to practice independently, at the same time, the teachers will be visited regularly.

Inviting software programming and testing experts from well-known enterprise to explain actual project, partner institutions may send less than three teachers to study free of charge for 10 days. Thus, study can fundamentally improve teaching standards of partner universities, which make the teaching level of university teachers enhance synchronization following the practical ability of enterprises, so that university teachers' project combat capability can achieve enterprise engineer level.

Company provides equipment and funding, school provides sites, the two sides persons set up laboratory together to complete the development of research projects, sharing resources and results. Laboratory not only provides for the advanced

undergraduate of experimental teaching spaces, such as the "Principles of Microwave Technology", "Antenna" and other courses, but also for graduate students and teachers, provides technical support for research activities. Once students enter the laboratory, they can understand corporate culture, know well the business regulations, which will strengthen the corporate identity of students, train student sense of responsibility and professional dedication, and enhance collaboration awareness of student, improve their organizational management and decision-making capacity. For teachers, through participating in laboratory research, they will contact with business frequently, understand the dynamic development of the industry, which will enhance the level of school teachers and teaching quality, improve research efficiency, and expand opportunities for innovation. Meanwhile, coordinating the implementation of enterprises with the advantages of school personnel and technology, the scientific achievements of university can be quickly converted.

Enterprise provides laboratory equipment and chip, sends engineers to train the students in software testing. Training takes the form of intensive classes in laboratory, students volunteer, and finally through the company's assessment, the company issues the technology certification. The model, Suzhou University and America joint laboratory of Microchip "SCM", is generally welcomed by the students. Students not only learn the latest SCM knowledge, exercise the practical ability and creative ability, but also get the technology certificate which increased the chance for employment after graduation. Enterprise regards the social effects as important, through the joint of college and enterprise, enterprise has good publicity to increase his own reputation and value of intangible asset. Enterprise can attract the best students to his own businesses through training. Because students have accepted their own training philosophy, they are more fitted for enterprise. In this way university turns into a powerful talent training base for enterprise.

University-enterprise joint laboratories may adopt the open mode in order to give full paying the resource advantage. After completing the professional experiment of normal teaching program, the student is arranged to manage the joint laboratory, keep it open. The different levels students may take different open forms. For junior students, they can not use the experimental equipment as they haven't done many experiments. Exercising will save much time to familiar instruments in the future experiment, and ensure the experimental teaching effect. In normal experimental teaching program, the ability of students varies, different levels of the design program, within the prescribed time, although most of the students have completed the experiment, there are still a few students did not complete. Some students completed the experiment, but they do not understand the situation of experimental results. For the above all, make the laboratory open to students when it isn't occupied, thus students can do experiments when they are not in experimental class. For interested high school students with strong ability, to carry out large-scale integrated experiments or difficult specialized topics. Teachers design a number of innovative projects for students to choose each year, students use their spare time to complete the task in the open laboratory. For these innovative projects, enterprises and schools will give to certain funds or material support. The projects will be supervised and inspected by teacher finally until complete. In addition, to participate in various forms of electronic design competition, all kinds of scientific research projects, graduate design, students can be trained in the laboratory, which will fully obtain the advantages of new laboratory equipment and high technology.

Multi-level open Laboratory can stimulate enthusiasm and creativity of students, then students should be taught by different levels, it is conducive to discovering and developing highly qualified personnel with the spirit of hard work.

To participate the various competitions organized by enterprise. Many famous enterprises home and abroad with the University Planning Department, take the form of gifts and training in colleges and universities to promote their products. The United States TI and XILINX attach great importance to cooperation with the university, they have cooperated with the centre, then set up the joint laboratories. Experimental conditions have achieved modernization, openness and sharing. Insist on paying equal attention to the theory and practice, ability-training parallels quality-improving, combining curriculum experimental and innovative design with scientific research, adhering to knowledge transmitting, ability improving, and quality training has always been run throughout the experiment. The American TI organizes the various types of electronic design competition in college students each year to inspire students to electronic design innovation and creative passion. "TI - DSP Innovation Design Competition", "TI Cup Analog Electronics Design Competition" etc., the competitions cause great controversy among the students. Joint laboratory takes this type of competition as an opportunity, guiding students with the experience teachers, makes full use of modern laboratory equipment laboratory, takes the open form, enables students have sufficient opportunity to exercise capacity. These competitions can stimulate the students learning interest, greatly improve the overall quality of students. The practice shows, enterprises and universities cooperate and return to the society, is a significant public welfare, and is a strategy to revitalize the country through science and education. This strategy is an useful important measure for building a harmonious society. From the academy perspective, they hope to strengthen cooperation with enterprises as a link to the cooperative relationship of industries, universities and research, make a breakthrough on the cooperation and achieve "win-win" between schools and employers. In addition to "primary level" cooperation relations of constructing practice base, in future they will gradually extend the cooperation relationship to set up specialized laboratories, reform the existing training model, establish education fund with enterprise together, which will promote accelerated growth of universities and enterprises.

3 Problems in University-Enterprise Cooperation

Although university-enterprise cooperation has a lot of advantages, it inevitably exists some problems. Some laboratories are not linked closely with the subject, teachers and students involved the laboratory equipment less, which make laboratory equipment use ratio lower, experimental resources is wasted because the laboratory is idle for long time. Some laboratories are occupied by enterprise for long time, the school does not dominate running in the laboratory, the result is teachers or students only use the lab in a small period of time in the case of an appointment, making the use of the experimental resources unfair. In addition, there have been some problems in the process of opening the joint lab. First of all, the strength equipments lack maintenance. After lab opening, the students' time greatly increased access to laboratory, equipment utilization is improved. However, due to varying levels of abilities of students, it inevitably causes

damage to laboratory equipment, increases maintenance workload. Secondly, the students lack teachers guiding force. In Joint laboratory, the experiment content is flexible and open, and some content is closely with the enterprise production, students will ask all kinds of questions at any time, which propose higher requirements to teachers, only these teachers with theory and practical ability will qualify open laboratory work. Finally, students lack enthusiasm. Open laboratory aims to improve the practical ability of students, in practice, due to lack self-awareness, and sense of participation, only a few students enter the laboratory. For these problems, it is necessary to establish some essential rules and regulations to norm open laboratory, enable a variety of modes carried out well results.

4 Conclusion

Now, software testing is in the start-up best period, how to make software development and testing enterprise cooperate with the post-secondary colleges, make up for their current shortage becomes an urgent and realistic issue. Colleges and enterprises set up laboratory together, the movement can use social resources to run schools, realize the complementary advantages and mutual benefits between schools and enterprises. The current popular "production-education cooperation" and "university-enterprise cooperation" models, which not only make both schools and enterprises play their respective advantages, but also train talents with social and market need, which is one of the win-win patterns of the universities and enterprises (social). Strengthening the cooperation between schools and enterprises, the combination of teaching and production, school and enterprise support each other, mutual penetration, two-way intervention, complementary advantages, resource interoperability, interests sharing, is an important way to realize modernization of higher education and corporate management, and promotes the development of productivity, accelerates the pace of own talents academic education and make education and production for sustainable development. University-enterprise cooperation or production-education combining is effective means to promote technology, economic and enterprise development. It is an important way to run higher education well, promote the vitality of enterprises and train expertise of production, construction, management service line. Based on "sincere cooperation and mutually beneficial win-win" cooperation model, it will greatly enhance the quality of teaching practice in schools, speeds up construction of related courses, improves the student overall quality. It will be the new model of setting up university laboratory and training personnel.

References

1. Binder, R.V.: Testing Object-Oriented Systems: Objects, Patterns, and Tools, p. 45. Addison-Wesley Professional (1999) ISBN 0-201-80938-9
2. Beizer, B.: Software Testing Techniques, 2nd edn., p. 21, 430. Van Nostrand Reinhold, New York (1990) ISBN 0-442-20672-0
3. IEEE, IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries. IEEE, New York (1990) ISBN 1559370793

4. van Veenendaal, E.: Standard glossary of terms used in Software Testing, <http://www.astqb.org/educational-resources/glossary> (retrieved June 17, 2010)
5. Myers, G.J.: The Art of Software Testing, pp. 145–146. John Wiley and Sons (1979) ISBN 0-471-04328-1
6. Dustin, E.: Effective Software Testing, p. 3. Addison Wesley (2002) ISBN 0-20179-429-2
7. Zhang, G.-P., Wang, Y.-C.: Predicament and Outlets of the Current University-Enterprise Cooperation in Personnel Training. Journal of East China Institute of Technology, Social Science (March 2009)
8. Zheng, X.: Enhancing university-enterprise cooperation to promote the development of competence education. Journal of Technology College Education 23(2) (2004)
9. Van der Beken: Interuniversity cooperation and university-enterprise partnership in the framework of the programmes of the European Community. EC (2006)

Using Surface Electromyography Signals to Control an Exoskeleton Arm Driven by Pneumatic Artificial Muscles

Jie Liu¹, Yu Wang^{2,*}, Heting Tong², and Ray P.S. Han^{1,*}

¹ College of Automation Engineering, Qingdao University, Qingdao, China 266071

² College of Mechanical Engineering, Qingdao University, Qingdao, China 266071
liujie.hz@gmail.com, ywang@qdu.edu.cn, tongheting@163.com,
ray-han@pku.edu.cn

Abstract. This paper is concerned with the use of surface electromyography (sEMG) signals to control an exoskeleton arm driven by pneumatic artificial muscles (PAM). The acquisition and processing of the sEMG signals are first discussed and then used in the controls algorithm for driving the PAM-activated exoskeleton arm. The sEMG signals are generated by repeated cycles of the human arm in flexion-extension movements, causing the exoskeleton arm to move repeatedly in unison in our experimental setup.

Keywords: Surface Electromyography (sEMG), Pneumatic Artificial Muscles (PAM), Exoskeleton Arm, Flexion-Extension Motions.

1 Introduction

Surface electromyography (sEMG) signals generated by human skeletal muscles are weak but detectable using non-invasive electrodes attached to the skin surface. They capture the state of muscle function and are widely used for the diagnosis of muscle injuries, determination of muscle fatigue and other aspects of rehabilitation medicine involving human motions. sEMG is an active area of research with feature extraction and pattern classification methods being popular techniques for use in data processing and interpretation [1, 2, 3, 4]. For a man-machine system such as an exoskeleton robot, the collected sEMG signals can be rapidly processed for feedback controls of the motors for driving the device motion. Japan's HAL-3 is a good example [5]; Bitzer at London University is another, where he employed sEMG signals to control a robot's four-fingers hand movement [6]. A motor-driven motion is not only unnatural compared to a muscle-driven motion, it also, produces varying degrees of discomfort for the human user. In our work, we are proposing to use sEMG signals for driving PAMs in an exoskeleton arm. However, this approach can be challenging as there are other issues, for example, air compressibility and viscoelasticity of the PAM element that also, need to be considered [7].

* Corresponding authors.

2 Acquisition and Processing of sEMG Signals

sEMG signals are physiological in nature, with a minor amplitude in the 100-500 μ V range, the peak at between 0-6mV, RMS in the 0-1.5mV range, the useful signal frequency component varying from 0-500Hz and the main energy concentrated in the 10-150Hz range [6].

sEMG Signal Collection. The bicep sEMG signals of 11 healthy Chinese men are collected in an indoor environment of between 18-24 degrees Celsius as the source data. Signal acquisition is achieved using the MyoTrac Infinity dual-channel device that allows for the measurement and real-time display of high-resolution sEMG signals (sampling frequency set at 2048 Hz). The device is connected to the computer via a USB data cable and the sEMG data can be processed using the software package supplied by Thought Technology Inc [8]. To collect the signal, medical ethanol is used to clean the skin surface of a human subject and a pair of positive and negative electrodes is placed on the cleaned area of the bicep with a 5-cm spacing. Next, the reference electrodes are placed in a position away from the detection electrodes at a location near the human shoulder (at a minimal distance of 10 cm). The human subject is then asked to perform repeated cycles of arm flexion-extension exercise.

sEMG Signal Denoising. Due to electromagnetic and other interferences in the lab environment, it is common for external influences to set-in during the collection process. This paper describes how to denoise the collected signals.

A Low-Pass Filtering Method. Considering the sEMG frequency characteristics, the raw signals are passed into an IIR low-pass filter with a cut-off frequency of 500 Hz. This filters out some of the high-frequency noise and the result is depicted in Fig. 1

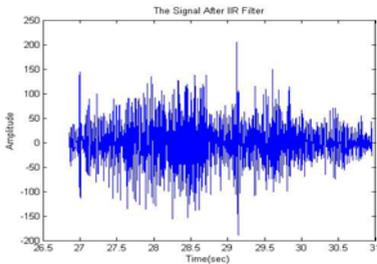


Fig. 1. IIR low-pass filtered EMG signals from a 1-cycle arm flexion-extension motion

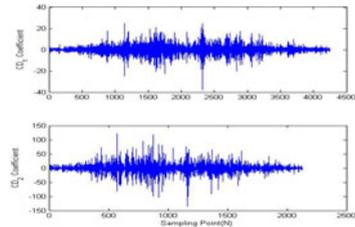


Fig. 2. Determining the CD_1 and CD_2 wavelet coefficients from wavelet decomposition

Wavelet Denoising. With the removal of some high frequency noise a second filtering of the signals is attempted for removing not only the remaining noise but also, for facilitating later post-processing of the collected data. The wavelet denoising method is employed for this purpose. The Daubechies wavelet is an orthogonal wavelet that can effectively eliminate non-orthogonal signal interferences and also, to extract the characteristics of the ideal waveform. The db4 wavelet base is used here and after two runs of wavelet decomposition, the CD_1 , CD_2 wavelet coefficients CA_1 and CA_2 are

obtained (see Fig. 2). Our work employs the high-frequency CD_2 , which contains the intensity change information as the main data source of the sEMG signals for the absolute values envelope curve.

Amplitude Characteristics of the sEMG Absolute Value Curve. To better understand the changes in sEMG signals, the envelope curve in Fig. 3 is generated by computing the absolute value of CD_2 .

Observe that one cycle of the arm flexion-extension movement is depicted in the envelope curve (Fig. 3) with the 2 phases; flexion and extension clearly depicted. The flexion phase corresponds to muscle contractions. As the muscles contracts, the speed of elbow joint angle varies from an increasing value to a decreasing one. The result is a release of the muscular energy that varies from a high to a low and this variation produces a strong-to-weak sEMG signals. During the extension phase, the released energy and thus, the sEMG signals are smaller than those encountered during flexion.

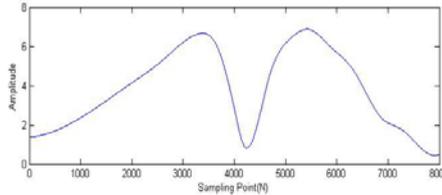


Fig. 3. Wavelet decomposition of the absolute value of the CD_2 envelope curve

3 Controls Algorithm

The controls algorithm can be explained as follows. The voltage signal that controls the arm flexion and extension movements is proportional to the ratio of the area under the amplitude envelop curve (see Fig. 3) under consideration to the largest area in the plot. Observe that the envelope curve possesses 2 peaks; the max area (s_1) of the first peak that corresponds to the flexion phase and the max area (s_2) of the second peak that corresponds to the extension phase. Therefore, at any given time, the current voltage of the control signal is proportional to the ratio of the current area under the envelope to the max area of the appropriate current peak. The specific voltage value of the control signal is determined by 2 coefficients \bar{A}_1 and \bar{A}_2 that can be obtained experimentally. The area under the amplitude envelope can be calculated by an absolute integral algorithm in time domain.

Flexion Phase. This phase corresponds to a complete bending movement of arm, with the biceps as the main muscles for the energy output. Using Eqns. (1), (2) below, we obtain the controls signals from the sEMG signals.

$$s_1(k) = \sum_{t_1}^{t_k} |e(t)| dt, \quad k = 1, 2, 3, \dots \quad (1)$$

where $e(t)$ can be determined from the pre-processing of sEMG signals, $s_1(k)$ is the sEMG energy release function, t_1 is the start time of the bending movement and k is sampling point during the flexion phase.

In order to obtain the control signals employed in this work, the absolute value of the integral function $s_1(t)$ in time domain is normalized after computations, and then, followed by a scaling. Here we introduce a parameter A_1 to account for the effect of the external environment on the human muscles; for example, the human body at different times and different temperatures, different reactions to situations, as well as the differences between people and people. By suitably adjusting the magnitude of A_1 we can manually adjust the PAM's movement velocity to arrive at:

$$C_1(k) = \frac{A_1}{\max(s_1)} s_1(k), \quad k = 1, 2, 3, \dots \quad (2)$$

After transformation, the function $C_1(k)$ constitutes a time series for k and our work utilizes it as an exoskeleton arm control signal during the flexion phase to control the PAM movement.

Extension Phase. The modeling is the same as with the flexion phase, except that now, the biceps are stretching and the triceps are the main muscles for energy output. A new parameter A_2 is needed to handle the magnitude of the barometric pressure in the PAM. Further, since the energy of the biceps during extension is opposite to that during flexion, the sEMG signals are opposing to the desired PAM movement controls. Hence, making Eqns (1), (2) need to be modified to reflect the new situation, that is:

$$s_2(k) = \sum_{t_2}^{t_k} |e(t)| dt, \quad k = 1, 2, 3, \dots \quad (3)$$

$$C_2(k) = A_2 \left(1 - \frac{s_2(k)}{\max(s_2)} \right), \quad k = 1, 2, 3, \dots \quad (4)$$

where $e(t)$ can be determined from the pre-processing of the sEMG signals, $s_2(k)$ is the sEMG signal energy release function, t_2 is the start time for extension exercise and k is the sampling point in extension phase. After transformation, the function $C_2(k)$ constitutes a time series for k and our work utilizes it as an exoskeleton arm control signal during the extension phase to control the PAM movement.

Our work employs the time series $C_1(k)$ and $C_2(k)$ for feeding the controller that activates the PAM for driving the exoskeleton arm under a 1-kg load.

4 Experiment and Results

Experiment Platform Introduction. The experimental platform (see Fig. 4) consists of an exoskeleton arm the length of which is that of a typical human arm size, together with the PAM-driven motion equipment controlled by sEMG signals. We used the MXAM-20-AA PAM made by Festo Inc. with the GMS412 RE10PB encoder manufactured by Shanghai GEMPLE Inc. and carrying a 1-kg external load.

Acquisition of sEMG signals is performed via the Advantech data collecting card through a real-time monitoring of the pressure data using the SMC ISE30-01-28- M pressure sensor. To regulate the air pressure feed to the PAM we used the SMC ITV0050-3BL proportional pressure valve.

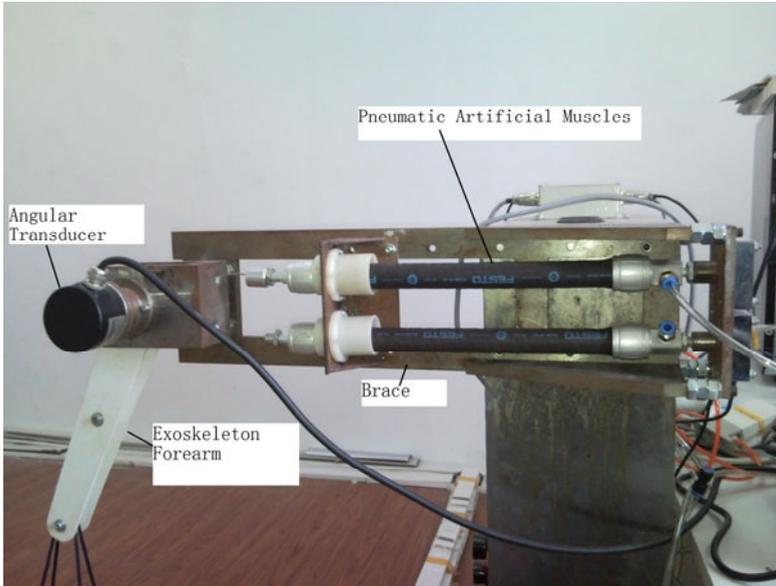


Fig. 4. The experimental platform consisting of a PAM-driven exoskeleton arm

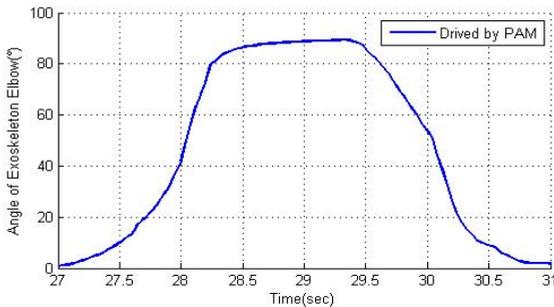


Fig. 5. The PAM-driven exoskeleton arm rotation motion from 27 to 31 seconds

Experimental Results Demonstration and Analysis. The experimental results of the exoskeleton arm joint rotation θ driven by PAM from 27 to 31 seconds are depicted in Fig. 5. During the 4-second duration, the human arm completed a cycle of flexion-extension movement; the exoskeleton arm also simultaneously completed a cycle of flexion-extension movement controlled by sEMG signals extracted from the human arm. In total, both the human subjects and the exoskeleton arm completed a total of 50 flexion-extension motions.

5 Conclusions

This paper demonstrates the use of sEMG signals to control a PAM-driven exoskeleton arm. Repeated runs of the experiment clearly show that the exoskeleton arm completed the flexion-extension motions as the human arm did. However, because only 1-pair of PAM was used for driving, the exoskeleton arm during the flexion motion moved at about the same speed as the human arm, but this was not the case during the extension motion – it moved at a relatively slower speed.

References

1. O'Sullivan, L.W., Gallwey, T.J.: Electromyorg. Kines 12, 275–285 (2002)
2. Kallenberg, L.A.C., Hermens, H.J.: Electromyorg. Kines 18, 780–788 (2008)
3. Han, J.S., Song, W.K., Kim, J.S., Bang, W.C., Lee, H., Bien, Z.: 6th International Conference on Soft Computing, pp. 890–897 (2000)
4. Nandedkar, S.D., Stalberg, E.V., Sanders, D.B.: IEEE. T. Bio-Med. Eng. 32, 775–785 (1985)
5. Kawamoto, H., Lee, S., Kanbe, S., Sankai, Y.: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, pp. 1648–1653 (2003)
6. Bitzer, S., van der Smagt, P.: Proceedings of the 2006 IEEE International Conference on Robotics and Automation, pp. 2819–2823 (2006)
7. Li, F., Wang, Y., Tong, H., Han, R.P.S.: Adv. Mater. Res. 186, 31–35 (2011)
8. Information on, <http://www.thoughttechnology.com>

A Study on Regression Analysis of Body of 4-6 Years Old Children in Henan Area*

Wang Shi-zhu and Zhou Ping**

Henan Institute of Science and Technology, College of Art, Xinxiang, Henan,
453003, China
zhouping8628@163.com

Abstract. Children's body to produce more in line with the characteristics of children's clothing, children's clothing to improve the relevance of this body and comfort the different counties of Henan Province selected children aged 4-6 to garment production head to the body measurements, measure the item 22, by data processing, access to 531 children aged 4-6 effective body measurement data. Statistical analysis of data size characteristics of children, selected 11 children can be summarized basic characteristics of body condition measure; Through regression analysis has been the main measurement site and height of children, bust and hip regression relationship. Regression relationship for the establishment of children's clothing designed to provide data based on the structure, thereby enhancing the physical fitness of children's clothing and comfort.

Keywords: Regression Analysis, 4-6 Years Old Children, Anthropometric, Clothing.

With the improvement of socio-economic conditions and the development of clothing, children's clothing fit the requirements increases. More than 4-6 years old children in the 80-130cm tall in China, between National standard type of 80-130cm tall with 10cm of the smallest children in sub-file units, and after that the actual measurement data, the height difference of 10cm length of children, circumference and width of the indicators vary widely; The other hand, our children's clothing standard number type used number system is bust arithmetic progression [1], is often used in the initial standardization simpler classification. Some countries [3-7] such as Germany, Japan, have to change specifications point arithmetic progression is to use a ladder, which can reduce small degree of density unevenness quantity and specifications. In view of children's physical fitness requirements and the changing shape of children, by children aged 4-6 Henan body measurements, for children aged 4-6 Henan body research.

* Henan Province Natural Science Research Education Program in 2010(2010A540005).

** Corresponding author.



1 4-6 Year-Olds Body Measurements

Anthropometric [2] is the body shape analysis and research foundation, this study used a static measurement of human contact in manual measurement method.

1.1 Anthropometric Object

4-6 years old children in Henan, the main measurement area Xinxiang, Zhengzhou, Anyang, Nanyang, Puyang, Kaifeng and so on.

1.2 Measurement Sample Selection

1.2.1 Samples with Sample Size Sampling Method

Sample contains the number of individuals known as the sample size [9], that sample size. Determine the number of sampling units is calculated as: $N=(t \times s / \Delta)^2$: N-number of sampling units; t-degree of probability sampling error; s-standard deviation; Δ - tolerance. In the experiment, t is 1.96, confidence F (t) is 0.9500 Reference to previous studies [10-13] the main part of the maximum allowable measurement error and standard deviation of data in table 1.

Table 1. Sample size calculator

| Site | the maximum permissible error | refer to standard deviation | s/ Δ | sample size calculation |
|-----------|-------------------------------|-----------------------------|-------------|-------------------------|
| Height | 1.0 | 6.2 | 6.2 | 148 |
| Bust | 1.5 | 5.5 | 3.67 | 52 |
| Waistline | 1.0 | 6.7 | 6.7 | 173 |
| Hip | 1.5 | 5.2 | 3.45 | 46 |

As can be seen from Table 1, sample 173 can meet the test requirements. According to law of large numbers, a large number of random phenomena with the stability, the sample size of this sample is 531.

1.3 Measurements of Children Aged 4-6

1.3.1 Measurements

The measurement is based on number of children's clothing clothing production and research-based specifications for the main purpose of the measure selected 22 projects, mainly related to children's body height, circumference, length, width, angle and weight. Specific measurements [14] as follows: height, waist high, arm circumference, head circumference, neck root circumference, chest circumference, waist circumference, abdominal circumference, Root arm circumference, wrist circumference, hip circumference, total shoulder, right shoulder width, front axillary points wide, before the axillary point width, back length, full-arm length, stocks on the long, long threatened, long the seventh cervical vertebra to the hip, shoulder angle, body weight.

1.3.2 Measuring Tools

This measurement tool used in the main instrument for the height and weight, soft feet, set square, protractor, angle meter made.

2 Data Processing

2.1 Data Preprocessing

2.1.1 Data Screening

This actual total of 607 children had body measurements, after data screening, the final total of 531 children in the full measuring volume data.

2.1.2 Normal Test Sample

According to the theory of mathematical statistics, the sample is large, the sample measurements approximate normal distribution. For not normally distributed random variables, if the sample size is quite large, and the standard normal random variable random variable on the difference is small. The measurement test of the sample size of 531, fully meet the requirements, so that all the measurements tests measure both normal distribution.

2.2 Principal Component Analysis

For 4-6 year olds from the 22 body measurements, select the most representative characteristics of the body measurements of children's human data principal component analysis. The results shown in Table 2.

Table 2. 4-6 year olds body data principal component analysis

| | Eigenvalue | Difference | Proportion | Cumulative |
|----|------------|------------|------------|------------|
| 1 | 10.1600554 | 8.0550411 | 0.4618 | 0.4618 |
| 2 | 2.1050143 | 0.6714255 | 0.0957 | 0.5575 |
| 3 | 1.4335888 | 0.2933662 | 0.0652 | 0.6227 |
| 4 | 1.1402226 | 0.1229137 | 0.0518 | 0.6745 |
| 5 | 1.0173089 | 0.1805821 | 0.0462 | 0.7207 |
| 6 | 0.8367268 | 0.1013321 | 0.0380 | 0.7588 |
| 7 | 0.7353947 | 0.1412722 | 0.0334 | 0.7922 |
| 8 | 0.5941224 | 0.0378977 | 0.0270 | 0.8192 |
| 9 | 0.5562247 | 0.0240632 | 0.0253 | 0.8445 |
| 10 | 0.5321615 | 0.0651011 | 0.0242 | 0.8687 |

Table 2. (Continued)

| | | | | |
|----|-----------|-----------|--------|--------|
| 11 | 0.4670604 | 0.0345272 | 0.0212 | 0.8899 |
| 12 | 0.4325333 | 0.0449647 | 0.0197 | 0.9096 |
| 13 | 0.3875686 | 0.0750361 | 0.0176 | 0.9272 |
| 14 | 0.3125325 | 0.0175609 | 0.0142 | 0.9414 |
| 15 | 0.2949716 | 0.0611867 | 0.0134 | 0.9548 |
| 16 | 0.2337849 | 0.0285435 | 0.0106 | 0.9654 |
| 17 | 0.2052415 | 0.0245731 | 0.0093 | 0.9748 |
| 18 | 0.1806683 | 0.0590418 | 0.0082 | 0.983 |

As can be seen from Table 2, the first five principal components that can provide information on the original index 72.07%; The first five principal components can be selected. Principal component analysis, the first five principal components of correlation matrix in Table 3.

Table 3. Principal component analysis of variable correlation matrix

| Measurements | Prin1 | Prin2 | Prin3 | Prin4 | Prin5 |
|----------------------------|----------|----------|----------|----------|----------|
| Height | 0.288616 | -.057336 | -.161115 | 0.096470 | -.073331 |
| Waist High | 0.269283 | -.166220 | -.215606 | 0.025805 | -.001233 |
| Head circumference | 0.230444 | 0.065405 | 0.146407 | 0.161056 | 0.075760 |
| Neck rhizosphere | 0.171975 | 0.251517 | -.164211 | -.054883 | 0.475684 |
| Bust | 0.255964 | 0.243067 | 0.015782 | -.165582 | 0.018922 |
| Waistline | 0.150196 | 0.465449 | 0.141231 | -.290611 | 0.057935 |
| Abdominal circumference | 0.212252 | 0.357986 | -.023049 | -.110677 | -.031410 |
| Hip | 0.261862 | 0.197466 | 0.068966 | -.072707 | -.156692 |
| Arm rhizosphere | 0.213416 | -.011422 | 0.220597 | 0.195123 | -.141421 |
| Wrist | 0.166222 | -.114469 | 0.183338 | 0.099899 | 0.140534 |
| Total Shoulder | 0.190173 | -.133521 | 0.208142 | 0.109408 | -.136838 |
| Before axillary-point wide | 0.195083 | -.321893 | 0.376248 | -.141655 | 0.209194 |
| After axillary-point wide | 0.191425 | -.348762 | 0.276332 | -.089758 | 0.351255 |
| Back Length | 0.137516 | 0.238838 | 0.435916 | 0.284621 | -.162462 |
| Full arm length | 0.263663 | -.141105 | -.130024 | -.040067 | -.114514 |
| crotch length | 0.220348 | -.227958 | -.192319 | 0.011693 | -.168257 |
| Arm circumference | 0.174821 | 0.106681 | 0.096152 | -.015475 | 0.252663 |
| flank length | 0.265357 | -.195537 | -.101562 | -.061168 | -.141593 |
| Shoulder angle | 0.017700 | 0.131420 | -.136453 | 0.805778 | 0.246000 |
| Body weight | 0.212808 | 0.056764 | -.296859 | 0.046730 | -.141954 |

Table 3 shows that:

(1) The first principal component of the index coefficients of standardized variables, representing the height, waist high, chest, full arm length, flank length variable coefficient: 0.288616, 0.269283, 0.255964, 0.261862, 0.263663, 0.265357, these value is distributed more evenly, and are positive, indicating that the human body circumference and length of the main components of the index plays the same direction, a considerable role.

(2) The second principal component coefficients in waist circumference and waist circumference than other items large and positive, while other indicators coefficient is small, mainly reflect the child's body circumference of the central information basis that the second principal component mainly reflected the children around the central body degree of information.

(3) The third main component in body length and width before the axillary points than any other factor items large and positive, while other indicators coefficient is small, mainly reflect the child's upper body circumference information.

(4) Shoulder angle in the fourth principal component factor than other large items, and a positive number, and other indicators are factors too small, mainly reflect the child's body in the shoulder angle information.

(5) Fifth principal component in the neck and right shoulder width rhizosphere coefficient, and is positive, the other indicators are factors too small, mainly reflect the child's neck and body of information, whereby the fifth principal component that mainly reflects the children's body shoulder and neck of the information.

Five main components of comprehensive information, with height, waist circumference, height, chest circumference, arm circumference, full arm length, flank length, waist circumference, hip circumference, the root neck circumference, body length, shoulder angle of a few indicators can be broadly used to describe the characteristics of children's body.

2.3 4-6 Year Olds Body Data Regression Model

Children 4-6 years old in order to further determine the quantitative relationship between the human data, the measurement data of the regression analysis.

After principal component analysis, the first principal component in the height, waist high, full-arm length, flank length, chest, hip coefficient. We build upper body and chest height to the regression model [15]; To lower body height, waist and hip as the independent variable regression model, dependent variable selection with the first five principal components in principal component analysis of the information, taking into account the national standard in child control parts and apparel design a key dimension. In table 4.

Table 4 shows children aged 4-6 the quantitative relationship between the measured data, the regression equation can be 4-6 years old children type of clothes making to provide the data basis.

Table 4. 4-6 year olds body data regression relationship

| Dependent variable | independent variables | the regression relationship |
|------------------------------------|-----------------------|--|
| Arm rhizosphere | Height bust | Arm rhizosphere = $-6.33078 + 0.16445 \times \text{Height} + 0.21939 \times \text{bust}$ |
| Neck rhizosphere shoulder width | bust Height bust | Neck rhizosphere = $3.12234 + 0.45530 \times \text{bust}$ shoulder width = $14.19510 + 0.09987 \times \text{Height} + 0.06079 \times \text{bust}$ |
| Back Length | Height bust | Back Length = $11.36011 + 0.04553 \times \text{Height} + 0.20509 \times \text{bust}$ |
| Arm length | Height bust | Arm length = $-5.04023 + 0.30519 \times \text{Height} + 0.11047 \times \text{bust}$ |
| crotch length | Height | Crotch length = $-5.05537 + 0.20827 \times \text{Height} + 0.03037 \times \text{Hip}$ |
| flank length | Height | flank length = $-2.41959 + 0.56672 \times \text{Height}$ |

3 Conclusion

Through the body of data analysis, the following main conclusions:

(1) principal component analysis extracted five principal components. Five main components of comprehensive information, with height, waist high, full-arm length, threatening long, bust, waist, hip, neck rhizosphere, back length, shoulder width, shoulder angle of 11 indicators will be basically summed up the children's body characteristics of the situation.

(2) Through regression analysis has been the main measurement site and height of children, bust and hip regression relationship.

References

- Li, Y.-M.: Study on improving the standards of apparel type and size in our country. *Shanghai Textile Science & Technology* 6, 70–75 (2006)
- Zhang, W., Fang, F.: *Clothing ergonomics*, vol. 9, pp. 15–55. Donghua University Press, Shanghai (2008)
- Xie, H.: *Woman Body Recognition and Basic Block Designing*, vol. 7, pp. 8–12. Doctoral Dissertation of Donghua University (2002)
- Jarosz, E.: Anthropometry of elderly women in Poland: dimensions for design. *International Journal of Industrial Ergonomics* (25), 225–209 (1999)
- Kothiyal, K., Tettey, S.: Anthropometric data of elderly People in Australia. *Applied Ergonomics* (31), 329–330 (2000)
- Connell, L.J., et al.: Body scan analysis for fit models based body shape and posture analysis. *National Textile Center Ann Report* 11 (2002)
- Barroso, M.P., Arezes, P.M., et al.: Anthropometric study of portuguese workers. *International Journal of Industrial Ergonomics* (35), 401–410 (2005)
- Guan, S.-Y.: *Anthropometry 4-6 years old children*, vol. 1, pp. 1–3. Tianjin University (2007)

9. Wang, H.-E.: Research on Body Shape of Women and Basic Block Founding on MTM. Donghua University, Master Thesis, vol. 12, p. 10 (2003)
10. Ray, G.G., Ghosh, S., Atreya, V.: An anthropometric survey of Indian schoolchildren aged 3-5 years. *Applied Ergonomics* 26(1), 67-72 (1995)
11. Carter, J.E.L., Mirwald, R.L., et al.: Somatotypes of 7- to 16-Year-old boys in Saskatchewan, Canada. *American Journal of Human Biology* 9, 257-272 (1997)
12. Prado-Leon, L.R., Avila-Chaurand, R., et al.: Anthropometric study of Mexican primary school children. *Applied Ergonomics* (32), 339-345 (2001)
13. Ventrella, A.R., Semproli, S., et al.: Somatotype in 6-11-year-old Italian and Estonian school children. *Homo—Journal of Comparative Human Biology* (59), 383-396 (2008)
14. Wu, S.-G.: 6-9 year-old boy in Liaoning size classification and prototyping. Suzhou University. Master Thesis 10, pp. 68-69 (2006)
15. Hu, X.-Y.: SAS statistical analysis, vol. 08, p. 18. Northwest A & F University Press (2007)

Recognition and Evaluation of Information Security Based on Non-optimum Analysis

Lihua Duan

Department of Information, Liaoning Police Academy,
Dalian, 116036, P.R. China
dldlh@sina.com

Abstract. This paper discusses the problems of non-optimum recognition, and establishes the system of non-optimum evaluation in the field of information security, based on non-optimum theory of systems. According to researches, the recognition and evaluation of non-optimum information directly decide the attributes of information security. It is an effective method to evaluate the information security that the non-optimum recognition about information systems is in the past, the present and future tense.

Keywords: Non-optimum analysis, information security, recognition algorithm, evaluation system.

1 Introduction

System non-optimum theory (SNOT) is an embryonic system science. Since it was put forward by He Ping in 1986 [1], this thought way has been a hotspot in the many fields, and now it is developing rapidly. For example, Professor Zoglarke of La-Salle University in American proposed thoughtfulness about limiting factors of systems based on the thought of SNOT [2]. And it has been achieved satisfactory results with the successful application in cost-benefit analysis. Brussels party, lead by Prigogine professor in Belgium, has been researched order-disorder evaluation law of systems [3]. The scholars in economics have been putting up economic development and decision-making by using SNOT in Austria. The studies showed that the development of SNOT is perfect, and it can be increasingly known by human.

System non-optimum analysis is confirmed two categories, that is “optimum” and “non-optimum”, based on processes and results of human knowledge and practice, and satisfied the demand of subjectivity and objectivity, as well as fitted objective reason standards. Hereinto “optimum” category is concluded the optimal and excellent, that is to say, it includes processes and results of success; “non-optimum” is concluded not good processes and results, maybe it is failed and acceptable. The classic non-optimum is infeasible and irrational. Even if it is feasible in certain degree, but it often belongs to non-optimum category. Quantitative descriptions of the bound between optimum and non-optimum category are one of the important study items of system non-optimum analysis [4]. As far as the dependent relationships of two categories are concerned, the form of non-optimum category is the foundations of

optimum category. Generally, only when people have jumped out non-optimum category, will they be possible to go in for optimum models or optimum processes.

Because of the happened and happening events (Social and National) in human life, like non-optimum events, states, knowledge and behaviors, it has been rich in original resources of system power. We can excavate it, transfer it, store it, and process it. According this, the databases of non-optimum information system could be building. The height form is that it could be brought into effect by hardware and software in computer. The initial form is composed by datum, documents, and charts and so on. That the information stored in information storehouse is as the initial “energy source” of system power. It broken through the limitation of single transformation from material energy to power, this “energy source” builds up system power accordingly.

Based on the power analysis of system non-optimum information, we show quantitative description by mathematical form, in order to establish the basis non-optimum constraint of knowledge and behaviors in human. This paper introduced the theory simply, then focused on the non-optimum recognition for information security and evaluation system of information system, then discussed how to build up the system.

2 Non-optimum Recognition of Information Security

Whether or not it has a non-optimum recognition system, this is the security problem of an information system S . if it has it, then the system S is called security system. The above called recognition system is composed of three tense parameters systems offered by information system S . As shown in fig. 1:

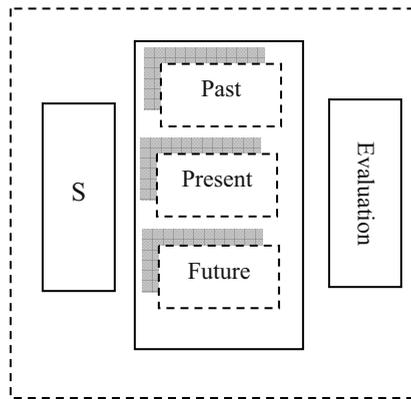


Fig. 1. Information security system

So here said the parameters system about past tense, means information system S has the various parameters with non-optimum factors during its growth. The parameters system about present tense is that the identifying of the running system about non-optimum factors in all types of parameters. What is called the future tense of the system S means that the information system has the forecast capability of

non-optimum factors, affected information security in the future. And it is expressed by a group of forecasting factors.

Because security problems of information system in operation is uncertainty and the general recognition and evaluation system is difficult to establish. In other words, it is difficult to assign a yardstick to measure information systems degree of security. Therefore, we can seek the recognized standards of information safety by the principle of non-optimum inversion. Let S^0 express information security system under the state of optimum, we will conform the conditions x^0 of optimum states. That is to say, if and only if what the value of x is? S^0 is a optimum system. Suppose f be a recognition, S^0 can be mapping to S^N by f , according to the experience model of the past tense; the distinguished model of the present tense; the predictions model of the future tense. Where it includes S^N , and the generated conditions x^N . Because x^N is easy to certain, then x^0 can be curtained by x^N . This is called non-optimum inversion principle, it shown as the following figure.

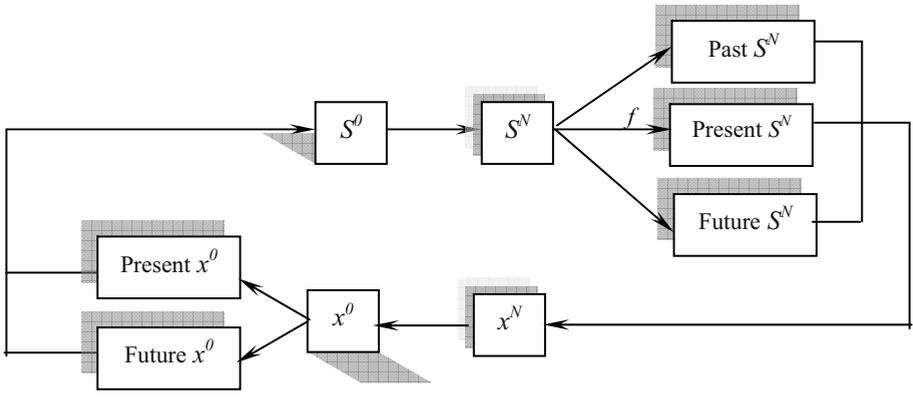


Fig. 2. Non-optimum inversion recognition system

"Three tenses" in recognition and evaluation systems are perhaps key roles in the identifying "non-optimum" information security. Analyzing the past tense of systems can be distinguished that if there is a non-optimum influence in the past. Analyzing present tense can be attained random non-optimum influence of system. Then we can attain the non-optimum omens, may be happened in the future, through the past and present tense. Therefore, "three tenses" composed three steps of non-optimum recognition, and decided the content of non-optimum recognition.

That the first step of recognition about non-optimum system S^N of information security, is to make analysis of S^N , happened in the past. The value of state variable could be produced by experience of statistics when non-optimum state came into being. The principles of recognition are depended on how much information content non-optimum system provided.



Let $P(x_j | N_i) = \frac{Q_{xj}}{Q_{Ni}}$ be a conditional probability of each element changed under produced conditions of non-optimum N_i . Where Q_{Ni} the arisen number of non-optimum is N_i , Q_{xj} is the changed number of system element x_j under the conditions of non-optimum N_i . And $P(N_i) = \{P(N_1), P(N_2) \dots P(N_{m+r})\}$ is the probability distribution of non-optimum N .

$$P(N_i) = \frac{Q_{Ni}}{Q_N}$$

Where Q_N is the total arisen number of all non-optimum (for system S). The above mentioned Q_{xj} , Q_{Ni} , Q_N can be obtained by experience of statistics. Thus entropy function according with $P(Z | N_i)$ is

$$H(Z | N) = - \sum_{j=1}^n P(x_j | N_i) \log_2 P \bullet (x_j | N_i)$$

Correspondingly mean entropy is

$$H(Z | N) = - \sum_{j=1}^{n+r} P(N_i) H(Z | N_i)$$

Select recognition level β , if

$$\frac{H(Z | N_i)}{H(Z | N)} \leq \beta,$$

correspondingly non-optimum N_i is main non-optimum as needed. That is multibreath non-optimum states for the changing factor of recognition system x_j as needed.

Thinks to the different attitude of system, lead to the different probability distribution of non-optimum N . Combining with multi-factor analysis, we attained variation degree of system S when non-optimum emerged.

Namely $dx_i = \frac{G - G^0}{G}$, where G practical action of system is, G^0 is an action of

non-optimum then $x_i = \frac{G - G^0}{G} dt$, where x_i is non-optimum factor of system.



3 Evaluation of Information Security Based on Non-optimum Recognition

It could be attained that the relations between non-optimum factors and eigenvectors of non-optimum behaviors using by regression analysis. This is non-optimum analysis for the past tense in the information system. The key for us is to attain the present behaviors of non-optimum, and how does the different factor change, expressed the behaviors of information system. Thus it is clear that these factors play the function in the information security. Finally, we could find which factors brought non-optimum behaviors for systems, and also find interrelation of non-optimum factor. A series of non-optimum factors could be afforded for the present information security.

In the present information system, it always exists a certain probability (visible or invisible) of what non-optimum happened. The happened of non-optimum is a random event. The non-optimum probability is not only decided by the forms and results of transformation between system and outside. Therefore, suppose system S has m non-optimum characters. Namely non-optimum character set is $N = \{N_1, \dots, N_m\}$. This set has two kinds of non-optimum character about inside and outside of systems. Each of non-optimum character will affect the system. Because of the affecting of non-optimum character, system S is transferred into non-optimum system S^N . We could attain non-optimum effect set $\mu = \{\mu_1, \dots, \mu_m^*\}$,

It's important to note that m is not equal to m^* . This is because if there are two characteristics of non-optimum only when they combine to influence on the system, such a non-optimum feature is called an association features of non-optimum. If every non-optimum feature might be influence the system independently, then $m = m^*$.

If μ_1, \dots, μ_m^* is influence degree of non-optimum features respectively, then the total influence degree is

$$\mu = \bigvee_{i=1}^H \left(\bigwedge_{j=1}^{m^*} \mu_{ji} \right)$$

When $\mu = \theta$, system S turned into S^N , wherefore, using the value of θ to measure the safety of property information system is optimum or non-optimum, then

$$S = \begin{cases} S^0 & \mu < \theta \\ S^N & \mu \geq \theta \end{cases}$$

The value of θ is a standard of measures. Any information security systems have a value of θ , nothing but the θ is given a different value according with the evolution of system. If the system is evolutionary, the value of θ will decrease, conversely the value of θ will increase. Otherwise the system would have some random non-optimum influence except it has the value of θ . This random influence must be have

a standard of measure, denoted as $\tilde{\theta}$, here the $\tilde{\theta}$ does not include the existing random effects but the new appear with the system functions changed. At the same time, in order to reach purposes of the system running, it must have a systemic forecast about non-optimum influence. That is to say, the system might be having non-optimum influence at a moment in the future. Suppose $\hat{\theta}$ be the standard measure of non-optimum influence, the standard measure of the system about non-optimum influence is composed by $\{\theta, \tilde{\theta}, \hat{\theta}\}$, then we have

$$S = \begin{cases} S^0 & \mu < \theta + \tilde{\theta} + \hat{\theta} \\ S^N & \mu \geq \theta + \tilde{\theta} + \hat{\theta} \end{cases}$$

When $\mu > \theta + \tilde{\theta} + \hat{\theta}$, the system is S^N , and each state is $x_i^N \in x^N$, that is to say the system is under non-optimum conditions. If and only if $x_i^N \in x^N$ changes a certain value, meanwhile $\mu < \theta + \tilde{\theta} + \hat{\theta}$, $x_i^N \in x^N$ is translated into $x_i^0 \in x^0$, namely the system is in the state of optimum. So there is

$$S^N \rightarrow x^N \searrow \mu < \theta + \tilde{\theta} + \hat{\theta} \nearrow x^0 \rightarrow S^0$$

Information security assessment system is to establish the standard measure of non-optimum influence about the past, present and future. Thought anglicizing statistical data of information systems in the past, we can attain parameters series $\theta_1, \theta_2, \dots, \theta_p$, where $\theta_1 = [\alpha_1, \beta_1]$, $\theta_2 = [\alpha_2, \beta_2]$, $\theta_p = [\alpha_p, \beta_p]$, $\alpha_i, \beta_i (i = 1, 2, \dots, p)$ is the upper and lower limitation of parameters, then

$$\theta = \bigwedge_{i=1}^p [\alpha_i, \beta_i] = [\alpha, \beta]$$

Recognize the present information security system, and attain $\tilde{\theta} = [\tilde{\alpha}, \tilde{\beta}]$, in the meantime $\hat{\theta} = [\hat{\alpha}, \hat{\beta}]$ is attained by forecasting. Well then, every system S has $\theta = [\alpha + \tilde{\alpha} + \hat{\alpha}, \beta + \tilde{\beta} + \hat{\beta}]$. Afterwards, compare the influence degree among non-optimum state of the system, and obtain the value of x^N . This x^N is called non-optimum constraint conditions. The constraint conditions of optimum x^0 is ensured by x^N .

A information security system is an adaptive system. That measure targets and conditions of system is optimum or non-optimum mainly depend on changing the influence degree, and adjust the parameter θ . Therefore, the recognition and decision-making are decided by adjustment of μ, θ in the system. The basic program of non-optimum evaluation about information security system is:

(1) Various non-optimum features should be collected as comprehensive and systematic as possible, and also collected most information in the development process of various non-optimum states. Then that would list all the distribution of non-optimum variables, and form the set of non-optimum features, then build system storehouse of non-optimum information;

(2) Based on systems analysis, to statistically process the collected information to order the features of non-optimum, analyze major and minor non-optimum features;

(3) Seek influence degrees of non-optimum features for system then attain the matrix of influence degree finally find synthetically influence in the base of the matrix of influence degree and influence weight;

(4) Make sure the parameter θ , then attain the constraint conditions of non-optimum x^N on the conditions of μ, θ ;

(5) Build up non-optimum parameters storehouse of information security systems.

4 Conclusions

The key issues of information system security are effective recognition and evaluation for non-optimum of the system. Depending on all above studies, we can identify security attributes of information systems by using non-optimum analysis methods of systems. Thereby we can control the security of information system. Because there are various uncertainty attributes of information security. For example, the random of non-optimum occurrence, the fuzzy of behaviors judgments, the unascertained of security attributes.

Due to the limit of space, the detailed algorithm and computer program, which is about recognition and evaluation system, will be introduced in another paper.

Acknowledgments. I gratefully acknowledge and thank Professor He Ping and Dr. He Jiantong for their guidance throughout this project, along with He Ping and He Jiantong for their reviews. This work was completed with the support of a research grant from the Ministry of Public Security.

References

1. Ping, H.: Theories and Methods of Non-optimum Analysis on Systems. *Journal of Engineering Science* 1(1), 73–80 (2004)
2. Ping, H.: Non-optimum system theory and methods. In: *First Joint IFSA-EC Workshop* (1986)
3. Cai, W.: Extension management engineering and applications. *International Journal of Operations and Quantitative Management* 5(2), 59–72 (1999)
4. He, P.: System Non-optimum Analysis and Extension Optimum Theory. In: Chen, G. (ed.) *Proc. of the Int'l Conf. on Systems Science and Systems Engineering*, pp. 131–137 (2003)
5. Linstone, H.: *Multiple Perspectives for Decision Making*, pp. 10–17. North Holland Publishing, New York (1984)
6. Linstone, H.: *Multiple Perspectives: Concepts, Applications, And User Guidelines*. *Systems Practice* 2(2), 307–331 (1989)

7. Ping, H.: Method of System Non-optimum Analysis in Crisis Management. In: Proc. of Second International Conference of Information System for Crisis Response and Management, pp. 640–645 (2007)
8. Ping, H.: The Method of Non-optimum Analysis on Risk Management System. In: Van de Walle, B. (ed.) Proc. of the Int'l Conf. ISCRAM 2008, pp. 604–609 (2008)
9. Ping, H.: Methods of Systems Non-optimum Analysis Based on Intervenient Optimum. In: Wang, Q. (ed.) Proc. of the Int'l Conf. on System Science, Management Science and System Dynamic, pp. 661–670 (2007)
10. Qu, Z., He, P.: Method of Non-optimum Analysis on Risk Control System. Journal of Software 4(4), 374–381 (2009)
11. He, P., Qu, Z.: Comparison Computing Based on Sub-optimum Analysis:A Guide System of Network Security. Journal of Networks 4(8), 779–786 (2009)
12. Qu, Z., He, P.: Intelligence Analysis Based on Intervenient Optimum learning Guide System. In: International Conference on Computational Intelligence and Natural Computing, pp. 336–339. IEEE Computer Society (2009)
13. He, P., Qu, Z.: Theories and Methods of Sub-optimum Based on Non-optimum Analysis. ICIC Express Letters 4(3), 237–242
14. He, J., He, P.: A New Intelligence Analysis Method Based on Sub-optimum Learning Model. In: Proc. of the Int'l Conf. International Conference on Future Computer and Communication, pp. 658–661. IEEE Computer Society (2009)

The Analysis of Two Dimensional Resource Allocation Procedures for IEEE802.16m of Mobile WiMAX Systems

Young-il Kim, Otgonbayar Bataa, Bat-Enkh Oyunbileg,
and Khishigjargal Gonchigsumlaa

Springer-Verlag, Computer Science Editorial, Tiergartenstr. 17,
69121 Heidelberg, Germany

Abstract. Multi user resource allocation is one of the key features towards high speed wireless network based on Orthogonal Frequency Division Multiplexing Access (OFDMA). According to IEEE802.16m (Mobile WiMAX) standard resource allocation problem has to be performed on a frequency and time two-dimensional space with the Physical and logical resource units (PRU and LRU) including Distributed logical resource unit (DRU and CRU) and Contiguous logical resource unit. In this paper we analyse the WiMAX frame structure in IEEE802.16m based on the Mobile WiMAX Standard. We apply novel resource allocation algorithm are used for managing two dimensional resources (time and frequency) to maximizing system capacity depending on mobile user's data rate.

Keywords: IEEE802.16m, Mobile WiMAX, Resource Allocation, PRU, DRU, CRU and LRU.

1 Introduction

The most suitable physical layer for the future deployment of Broadband Wireless Access networks is OFDMA. OFDMA has a few benefits large data rates, robustness against multi-path, flexible bandwidth allocation, and the possibility of exploiting multi-user diversity through an intelligent scheduling and resource allocation [1]. One of the most important advantages of OFDMA systems is included flexible resource allocation in two dimensional spaces.

The purpose of this paper is to provide resource allocation procedure and to give comparative analysis that need to be considered in developing distributed resource allocation for IEEE 802.16m. In OFDMA, the available spectrum (frequency domain) is divided into orthogonal sub-carriers, which are combined into groups, usually referred as sub-channels. Furthermore, the time domain structure is segmented into consecutive frames, each one containing multiple OFDM symbols. A slot delimited by one sub-channel and by a given number of OFDM symbols is the smallest resource unit that a Base Station (BS) can allocate to the users. OFDMA systems can provide multi-user access assigning different sub-channels and OFDM symbols to different users. Moreover, the system can exploit multi-user diversity, allocating in each frequency and time region (two dimensions) the user with the best channel conditions. This feature makes OFDMA systems, such as Long Term Evolution (LTE) or

WiMAX, more flexible than 3G solutions as High Speed Packet Access (HSPA) networks. However, these degrees of freedom present new challenges for the resource assignment and they require more sophisticated algorithms, capable of handling efficiently two-dimensional allocations.

In IEEE 802.16m for Mobile WiMAX [2], the case study in this paper, two constraints affect the allocation strategies. For OFDM, each subchannel can be modulated differently, but it only in the time domain. Because in OFDMA is considered both time and frequency domains. The OFDMA scheduler is the one of most complex problems because each MS can receive some portions of the allocation for the combination of time and frequency so that the channel capacity is efficiently utilized.

2 WiMAX 16m Frame Structure

The DL/UL subcarrier to resource unit mapping process, which is used to our research, is defined as following and illustrated in the Fig. 1.

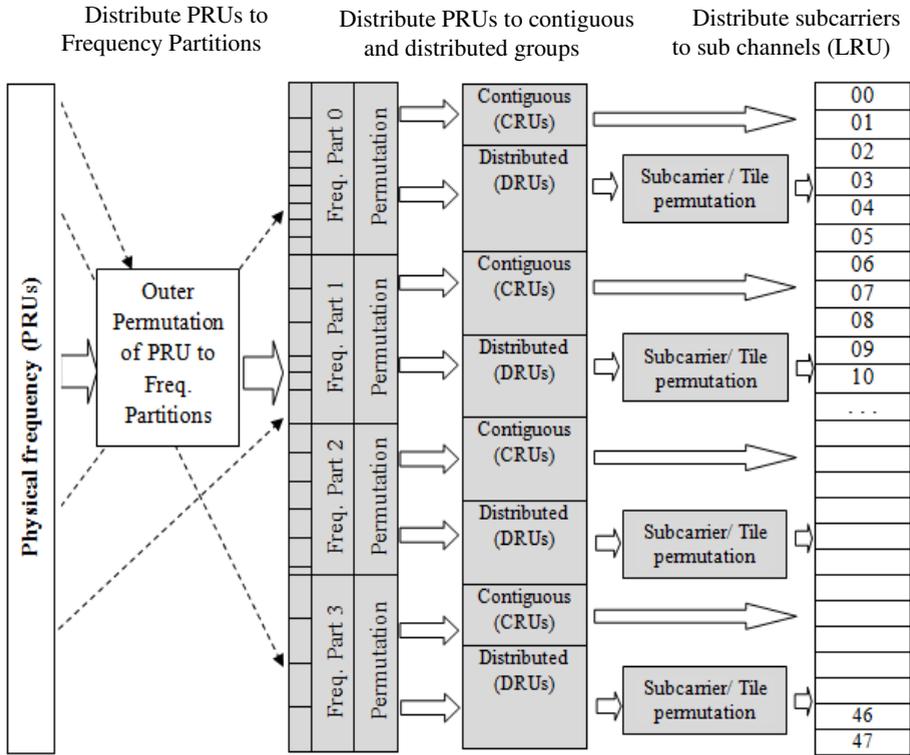


Fig. 1. DL and UL subcarrier to resource unit mapping

Table 1. OFDMA primitive and derived parameters for TDD mode

| Parameters | Values |
|--|---------|
| The nominal channel bandwidth, BW (MHz) | 10 |
| Sampling factor, n | 28/25 |
| Sampling frequency, F_s (MHz) | 11.2 |
| FFT size, N_{FFT} | 1024 |
| Subcarrier spacing, (kHz) | 10.94 |
| Useful symbol time, T_b (μ s) | 91.4 |
| CP ratio, G | 1/8 |
| OFDMA symbol time, T_s (μ s) | 102.857 |
| Number of OFDMA symbols per 5ms frame | 47 |
| TTG + RTG (μ s) | 165.714 |
| Number of guard sub-carriers (Left) | 80 |
| Number of guard sub-carriers (Right) | 79 |
| Number of used sub-carriers | 865 |
| Number of physical resource unit $-N_{PRU}$ (18x6) in a type-1 AAI subframe. | 48 |

The Advanced Air Interface supports TDD and FDD duplex modes, including H-FDD AMS operation. Unless otherwise specified, the frame structure attributes and baseband processing are common for all duplex modes. The Advanced Air Interface uses OFDMA as the multiple access schemes in the downlink and uplink.

2.1 TDD Frame Configuration

In the advanced air interface, each 20 ms superframe is divided into four equally-sized 5 ms radio frames. When using the same OFDMA parameters with the channel bandwidth of 5 MHz, 10 MHz, each 5 ms radio frame further consists of eight AAI subframes for $G = 1/8$. The basic frame structure is applied to FDD and TDD duplex schemes, including H-FDD AMS operation. In our case, TDD frame structure with DL to UL ratio of 5:3. Assuming OFDMA symbol 15 duration of 102.857μ s and a CP length of $1/8 T_s$, the lengths of type-1 subframe and type-3 subframe are 0.617, 16 ms and 0.514 ms, respectively. In our case, each downlink AAI subframe is divided into 4 frequency partitions; each partition consists of a set of physical resource units across the total number of OFDMA symbols available in the AAI subframe. Each frequency partition can include contiguous (localized) and/or non-contiguous (distributed) physical resource units. A physical resource unit (PRU) is the basic physical unit for resource allocation that comprises P_{sc} consecutive subcarriers by N_{sym} consecutive OFDMA symbols. P_{sc} is 18 subcarriers and N_{sym} is 6 OFDMA symbols for type-1 subframe, respectively. A logical resource unit (LRU) is the basic logical unit for distributed and localized resource allocations.

3 Multi-cell Resource Mapping

The PRUs are first subdivided into subbands and minibands. The number of subbands is denoted by K_{SB} . The number of minibands is denoted by K_{MB} . Since the maximum valid value of K_{SB} is 12 or N_{PRU}/N_l , we choose 10 MHz bandwidth with K_{BS} equal to 7. In case of $K_{BS} = 7$, values of parameters used multi-cell resource mapping are

specified Table 2. PRUs are partitioned and reordered into two groups: subband PRUs and miniband PRUs, denoted by PRU_{SB} and PRU_{MB} .

Table 2. Parameters for subband and miniband partition

| Parameters | Values |
|------------|--------|
| N_1 | 4 |
| N_2 | 1 |
| K_{SB} | 7 |
| L_{SB} | 28 |
| L_{MB} | 20 |
| $DSAC$ | 7 |
| N_{PRU} | 48 |

Equation (1) defines the mapping of PRUs to PRU_{SB} s. Equation (2) defines the mapping of PRUs to PRU_{MB} s.

$$PRU_{SB}[j] = PRU[i]; \quad j=0, 1, \dots, L_{SB} - 1. \tag{1}$$

where

$$i = N_1 \cdot \left\{ \left\lfloor \frac{N_{sub}}{N_{sub} - K_{SB}} \right\rfloor \cdot \left\lfloor \frac{j + L_{MB}}{N_1} \right\rfloor + \left\lfloor \frac{j + L_{MB}}{N_1} \right\rfloor \cdot \frac{GCD(N_{sub}, \lfloor \frac{N_{sub}}{N_{sub} - K_{SB}} \rfloor)}{N_{sub}} \right\} \bmod \{N_{sub}\} + \{j + L_{MB}\} \bmod \{N_1\}.$$

where $\{x\} \bmod \{y\}$ is modulus when dividing x by y , and $GCD(x, y)$ is the greatest common divisor of x and y .

$$PRU_{MB}[k] = PRU[i]; \quad k=0, 1, \dots, L_{MB} - 1. \tag{2}$$

where

$$i = \begin{cases} N_1 \cdot \left\{ \left\lfloor \frac{N_{sub}}{N_{sub} - K_{SB}} \right\rfloor \cdot \left\lfloor \frac{k}{N_1} \right\rfloor + \left\lfloor \frac{k}{N_1} \right\rfloor \cdot \frac{GCD(N_{sub}, \lfloor \frac{N_{sub}}{N_{sub} - K_{SB}} \rfloor)}{N_{sub}} \right\} \\ \cdot \bmod \{N_{sub}\} + \{k\} \cdot \bmod \{N_1\} & K_{SB} > 0 \\ k & K_{SB} = 0. \end{cases} \tag{3}$$

The miniband permutation maps the PRUMBs to Permuted PRU_{MB} s ($PPRU_{MB}$ s) to ensure frequency diverse PRUs are allocated to each frequency partition. Equation (3) describes the mapping from PRU_{MB} to $PPRU_{MB}$ s:

$$PPRU_{MB}[j] = PRU_{MB}[i]; \quad j=0, 1, \dots, L_{MB} - 1.$$

where

$$i = (q(j) \bmod (D)) \cdot \min(K_{MB}, N_1 / N_2) + \left\lfloor \frac{q(j)}{D} \right\rfloor. \tag{4}$$

$$r(j) = \max(j - (K_{MB} \bmod (\min(K_{MB}, N_1 / N_2)) \cdot D), 0). \tag{5}$$

$$q(j) = j + \left\lfloor \frac{r(j)}{D-1} \right\rfloor. \tag{6}$$



$$D = \left\lfloor \frac{K_{MB}}{P} + 1 \right\rfloor. \quad (7)$$

The number of subbands in i th frequency partition is denoted by K_{SB,FP_i} . The number of minibands is denoted by K_{MB,FP_i} , which is determined by the FPS_i and $DFPSC$ fields. The number of subband PRUs in each frequency partition is denoted by L_{SB,FP_i} , which is given by $L_{SB,FP_i} = N_1 \cdot K_{SB,FP_i}$.

$$K_{SB,FP_i} = \begin{cases} K_{SB} - (FPCT - 1) \cdot DFPSC & i=0, FPCT=4 \\ DFPSC & i>0, FPCT=4 \\ DFPSC & i>0, FPCT=3, DFPC=1 \\ K_{SB} - (FPCT - 1) \cdot DFPSC & i=0, FPCT=3, DFPC \neq 1 \\ DFPSC & i=1,2, FPCT=2 \\ K_{SB} & i=0, FPCT=1 \end{cases} \quad (8)$$

The number of miniband PRUs in each frequency partition is denoted by L_{MB,FP_i} , which is given by $L_{MB,FP_i} = N_2 \cdot K_{MB,FP_i}$. The number of subbands for each frequency partition is given by Equation (8). The mapping of subband PRUs and miniband PRUs to the frequency partition is given by Equation (9).

$$PRU_{FP_i}(j) = \begin{cases} PRU_{SB}(k_1) & \text{for } 0 \leq j < L_{SB,FP_i} \\ PPRUMB(k_2) & \text{for } L_{SB,FP_i} \leq j < L_{SB,FP_i} + L_{MB,FP_i} \end{cases} \quad (9)$$

$$\text{where } k_1 = \sum_{m=0}^{i-1} L_{SB,FP_m} + j \quad k_2 = \sum_{m=0}^{i-1} L_{MB,FP_m} + j - L_{SB,FP_i}$$

4 Cell-Specific Resource Mapping

PRU_{FP_i} s are mapped to LRUs. All further PRU and subcarrier permutation are constrained to the PRUs of a frequency partition. The number of subband-based CRUs for FP_0 is given by the Equation (10).

$$L_{SB-CRU,FP_0} = N_1 \cdot DCAS_{SB,0}. \quad (10)$$

The number of subband CRUs (L_{SB-CRU,FP_i}) and miniband CRUs (L_{MB-CRU,FP_i}) are derived using Equation (11) and Equation (12) respectively.

$$L_{SB-CRU,FP_i} = N_1 \cdot \min \{ DCAS_i, K_{SB,FP_i} \}. \quad (11)$$

$$L_{MB-CRU,FP_i} = \begin{cases} 0, & DCAS_i \leq K_{SB,FP_i} \\ (DCAS_i - K_{SB,FP_i}) \cdot N_1 & DCAS_i > K_{SB,FP_i}. \end{cases} \quad (12)$$

The total number of CRUs and DRUs in frequency partition FP_i is denoted by L_{CRU,FP_i} and L_{DRU,FP_i} .

$$\text{where } L_{CRU,FP_i} = L_{SB-CRU,FP_i} + L_{MB-CRU,FP_i}. \quad (13)$$

$$L_{DRU,FP_i} = FPS_i - L_{CRU,FP_i}. \quad (14)$$

5 Modeling Resource Allocation

Fig. 2 (a) shows the flowchart of the subband and miniband partition and Fig. 2 (b) shows the flowchart of the miniband permutation, which is used to our research.

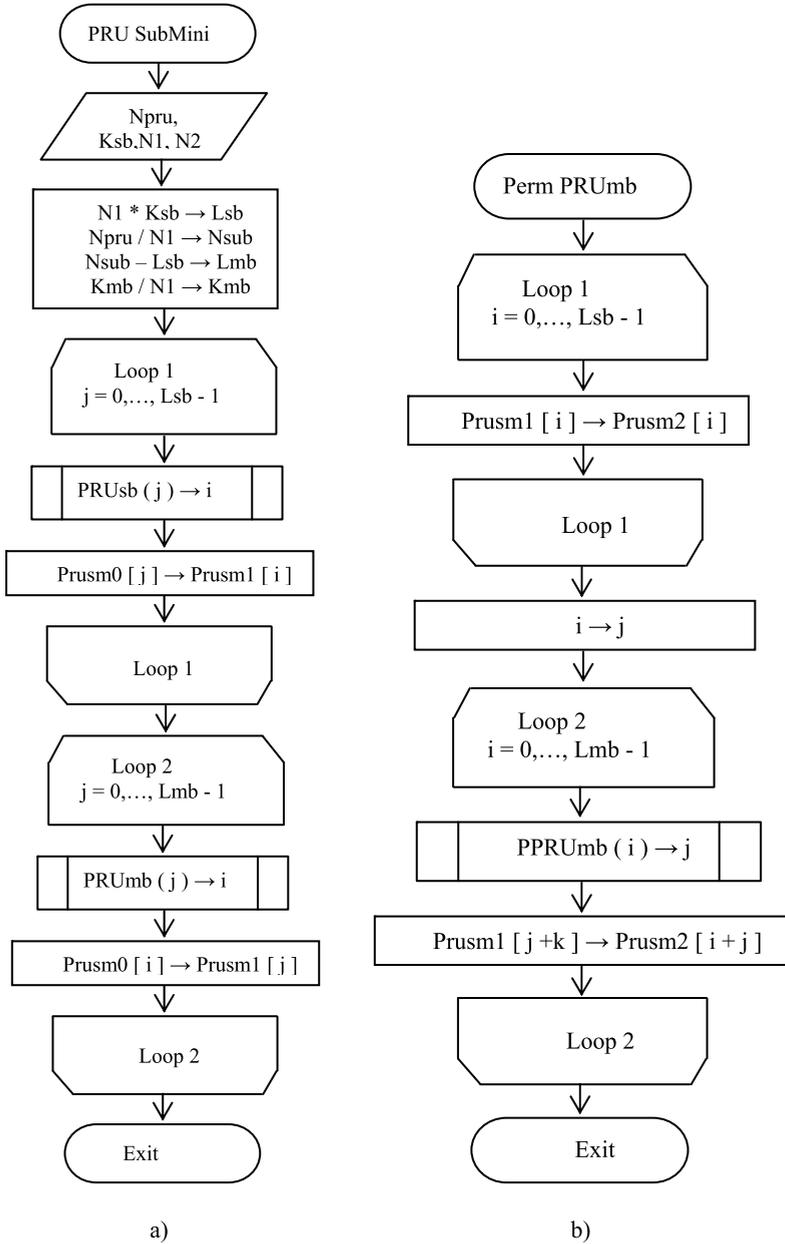


Fig. 2. a) Subband and Miniband partition's flowchart b) Miniband permutation's flowchart

Fig. 3 presents an example to illustrate the various steps of subband partitioning, miniband permutation, frequency partitioning, and CRU/DRU allocation for the case of 10 MHz system bandwidth.

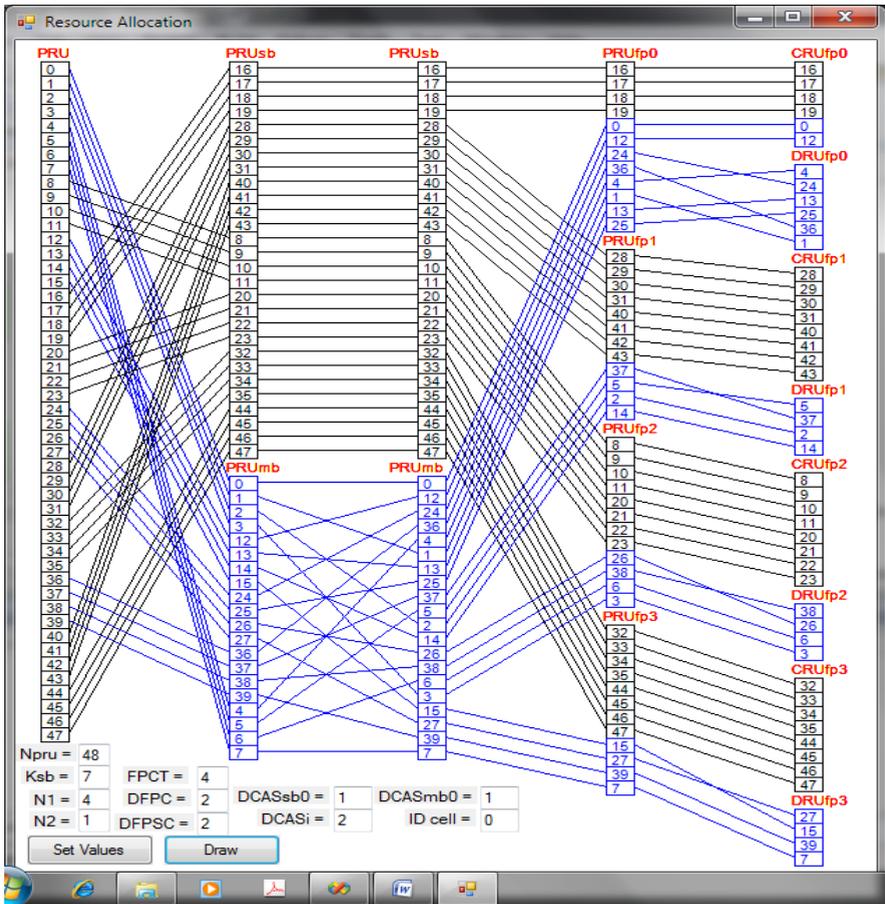


Fig. 3. Frequency partitioning for BW=10MHz, $PS_0=FPS_i=12$, $DFPSC=2$, $K_{SB}=7$, $FPCT=4$, $DCAS_{SB,0} = 1$, $DCAS_{MB,0} = 1$, $DCAS_i = 2$ and $ID_{cell}=0$

6 Conclusion

The novel resource allocation algorithms in uplink and downlink is proposed to optimize resource allocation procedure for use in 802.16m system. A downlink resource mapping algorithms supports both localized and distributed sub carriers, subcarrier permutation, the mapping between LRU and PRU, supports different fractional frequency reuse (FFR) group allocations. In a future work to developing that algorithm we can possible to reduce redundancy of transmitted control bit field in super frame header.

Acknowledgments. This work was supported by the IT R&D program of KCC/MKE [2010-10035206, The Development of IMT-Advanced Mobile IPTV core technology], Electronics and Telecommunications Research Institute of Korea and Mobicom corporation of Mongolia.

References

1. IEEE P802.16m/D7 July 2010, DRAFT Amendment to IEEE Standard for Local and metropolitan area networks, Part 16: Air Interface for Broadband Wireless Access Systems (July 29, 2010)
2. IEEE 802.16m System Description Document. IEEE 802.16m-08/003r9a (2009)
3. IEEE 802.16m Evaluation Methodology Document (EMD) (January 15, 2009)

DAIM: A Distributed Algorithm for Isolating Malfunctioning Nodes in Wireless Sensor Networks

Mohamed K. Watfa¹ and Rawad Abu Assi²

¹ University of Wollongong in Dubai, Computer Science Department, Dubai, UAE

² American University of Beirut, Computer Science Department, Beirut, Lebanon
sne.mohammad@gmail.com

Abstract. It has been identified that as complexity of computing and communication devices increases, fault-tolerance will gain more and more importance. Wireless sensor networks (WSNs) are exceptionally complex distributed systems where a variety of components interact in a complex way and should therefore help narrow down failures and diagnose their causes, as much as possible, with minimal physical access and interactivity. In this paper, we present an algorithm for isolating malfunctioning nodes in WSNs and provide two parallel variants of it: Naïve and Greedy. The algorithm is based on the idea that a covered node can be turned off and that turning off a malfunctioning node causes the WSN to function properly. The experiments we conducted show that the *Naïve Approach* is very precise in locating malfunctioning nodes whereas the *Greedy Approach* is very fast in finding a cover free of such nodes.

Keywords: Sensor Networks, Coverage, Malfunctioning nodes, Debugging, Sensor Networks, Greedy.

1 Introduction

Sensor networks introduce new challenges for fault-tolerance. Sensor networks are inherently fault-prone due to the shared wireless communication medium: message losses and corruptions (due to fading, collision, and hidden-node effect) are the norm rather than the exception. Moreover, node failures due to crash and energy exhaustion are commonplace. Thus, sensor nodes can lose synchrony and their programs can reach arbitrary states. Applications that impose an excessive communication burden on nodes are not acceptable since they drain the battery power quickly. Failures that often occur in wireless sensor networks can be attributed to many causes such as node failures, link failures, errors in the design, implementation errors, etc. Locating the causes of such failures is crucial to insure the reliability of the network but it is usually a challenging task due to several reasons such as the distributed nature of most protocols and applications, the energy constraints imposed on any technique, and the wide variety of faults in such networks that range from node crashes to bugs in the code running on the nodes. Since the complexity of software grows drastically with respect to its size, large scale software systems are extremely error-prone and fail frequently especially for sensor network applications, that are inherently distributed,

reasoning about the system and verification of correctness are more difficult due to the lack of a centralized controller and the lack of a globally shared memory. For example, Whitebox approaches for designing fault-tolerance, such as exception handling, forward recovery, recovery blocks, and application specific fault-tolerance methods assume that the implementation is fully available, and study the source code for designing fault-tolerance which are not applicable for large scale software systems because the task of studying the implementation and designing a fault-tolerant version becomes unbearable as the size of the implementation grows.

In this paper, we target failures that are caused by malfunctioning nodes that we define as the nodes suffering from a local defect i.e. the defect is not characterized as a general defect found in all nodes such as a software bug in the application code. Our approach is based on the idea that a node can be turned off if it is covered and consequently turning off a malfunctioning node will cause the network to behave properly. We propose DAIM, an algorithm which attempts to turn off such nodes and then tries to locate them. We provide two different versions of DAIM and analyse their performance and capabilities on a Java-based simulator we developed especially for this research. The rest of the paper is organized as follows: Section 2 elaborates on DAIM and its two versions, Section 3 presents the performance analysis we conducted, Section 4 discusses some related work, and Section 5 concludes the paper with some future work and in-sight.

2 Related Work

The full coverage problem, which verifies if every point in the region of interest is covered by at least one active sensor, has been studied in a variety of contexts. Our previous work [1-3] focused on the full coverage problem in 2D and 3D regions and provided algorithms to locate redundant sensor nodes in the region and deactivate them using simple geometric techniques. In [8], Gupta and Das design and analyze algorithms for self organization of a sensor network to reduce energy consumption. The work in [4] considers a large population of sensor nodes, deployed randomly for area monitoring to achieve an energy-efficient design that maintains area coverage. S. Meguerdichian, et al. in [5] consider an unreliable sensor grid-network and derive necessary and sufficient conditions for the coverage of the region and connectivity of the network in terms of the transmission radius, sensing radius, and failure rate of the sensor nodes. In [6], Lieska et al. formulate coverage problems to address the quality of service (surveillance) provided by a sensor network. The coverage concept with regard to the robot systems was introduced by Gage [7].

Concerning fault localization in WSNs, several approaches were proposed but as far as we know none of them followed our approach in searching for malfunctioning nodes. One of the proposed paradigms is to fix the faults before deployment by debugging a simulated version of the application. This approach doesn't take into consideration the physical effects of the environment on the behavior of the network and thus might overlook a wide range of faults. Another family of approaches is based on the idea of providing the developer with efficient tools to debug the network after being deployed [9-12]. Clairvoyant [10] was proposed as a source-level debugger for wireless sensor networks. It enables the developer to execute debugging commands

on each node as well as on the network as a whole. Commands such as inserting breakpoints and inspecting data values help the developer monitor the execution and infer the cause of the fault. Clairvoyant doesn't require any additional hardware and no modifications need to be made to the application's source code. On the other hand, some techniques try to target specific kinds of faults based on some heuristically algorithm such as Sympathy [9] that was proposed as a tool for detecting failures that are revealed by shortage in data flows, i.e. those that result in missing data. The intuition is based on the observation that there is a direct relationship between the amount of data collected at a sink and the existence of failures in the system. The main shortcoming of Sympathy is that it is limited to specific types of faults: those that are revealed by gaps in the data flow.

3 DAIM: A Distributed Algorithm for Isolating Malfunctioning Nodes

3.1 The Definitions and Assumptions

Definition 1: We refer by Cover to any subset of the nodes that can sense the monitored phenomenon at every point in the region.

Definition 2: A passing cover is a cover with no mal-functioning nodes. In contrast, a failing cover is a cover with at least one malfunctioning node.

Assumption 1: The basic assumption we make is that all malfunctioning nodes are covered i.e. which implied that the network is dense enough. In addition, we assume that each node knows the distances separating it from each of its 1-hop neighbours and thus can know if it is covered or not using pre-existing coverage algorithms in our previous work [1-3].

Assumption 2: Another basic assumption we adopt in our experiments is that the WSN is 2-Dimensional and that all nodes have the same sensing radius. In this case, and in a similar approach to that adopted in [1], we consider that a node is covered if the intersections of all neighbouring open disks that lie inside the node's open disk are covered.

3.2 Algorithm Description

Initially, we assume that all nodes are active. The state transition diagram of the algorithm is depicted in Figure 1. At the beginning of each iteration of the algorithm, each node checks if it is covered; if so, it waits for a random time (to avoid contention between neighbouring nodes) and then decides to sleep with a sleeping probability P which will be a subject of experimental evaluation in section 3. If it decides to sleep, then it selects a set of neighbours that cover it and sends to them an invitation message (a message asking these neighbours to be active in the next iteration so that the node will be covered). We propose two ways for selecting this set in the next section which will lead to two versions of the algorithm. Then the node waits for a certain period (we call it the negotiation period). When this period expires, if it

received an advertisement message from all of its invited neighbours it enters the sleeping mode. Nodes use advertisement messages to tell their 1-hop neighbours that they will be active in the next iteration. When a node enters the sleeping mode, it remains in that mode for a period of time we call application period which is the same period nodes remain in for the active mode. The idea behind this period is to allow the base station to get a feedback from the network (as part of the usual functioning of the application) so that it would be able to decide whether the current cover is passing or not. When the application period expires, a sleeping node enters the negotiation phase and waits for invitations from its neighbours. It is important to note that such nodes do not attempt to get active. They get active only if some neighbour invites them. In such case, they advertise that they will be active and enter the active mode.

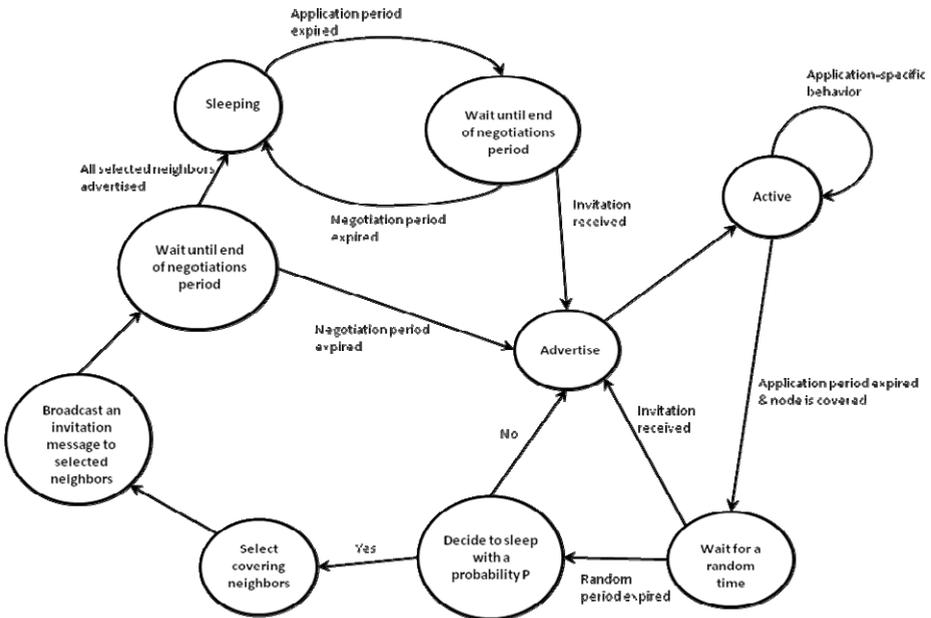


Fig. 1. The state transition diagram of DAIM

- After the application period expires, if an already active node received an invitation from one of its neighbours, then it will not attempt to sleep but rather it will advertise its presence and enter the active mode again.
- If the sleeping decision yielded a negative result, a node will advertise its presence and stay active in the next iteration.
- If a node that already broadcasted an invitation message didn't receive an advertisement message from all of the neighbours it invited, then it will advertise its presence and stay active in the next iteration. The algorithm terminates at the iteration in which the set of active nodes form a passing cover.

We present two different variations of the proposed technique. The two versions only differ in the way the covering nodes are selected. The naïve version (we will

refer to it by Naïve) simply chooses the set of all neighbours whereas the greedy version (we will refer to it by Greedy) attempts to select a subset of the neighbours that didn't participate much in previous iterations. It begins by selecting one neighbour at a time in a greedy fashion (selects the one with minimum participation times) and checks if the updated selected neighbours cover it. The pseudo code of Greedy is presented in Figure 2. Isolating the malfunctioning nodes can be implemented easily at the base station by maintaining a list of the covers that were selected through the algorithm (one cover for each iteration). The algorithm terminates when a passing cover Pass is found. Then, this cover can be compared with the failing covers of the previous iterations to find the one that resembles it the most (call it similarFail). Then the malfunctioning nodes will be isolated in the set Pass - similarFail. We will refer to such set by SUSPICIOUS.

```

-----
1. TEMP ← NEIGHBORS(i)
2. SELECTED ← ∅
3. while TEMP != ∅
4.     j ← NeighborWithMinParticipation(TEMP)
5.     SELECTED ← SELECTED ∪ {j}
6.     TEMP ← TEMP - {j}
7.     if isCovered(i, SELECTED)
8.         return SELECTED
9. return SELECTED
-----

```

Fig. 2. The Greedy Version Pseudo code

4 Performance Analysis

To evaluate the effectiveness of DAIM in isolating malfunctioning nodes, we define two metrics: convergence delay and isolation precision. The first is the average number of iterations taken to find a passing cover and the second is simply the size of the SUSPICIOUS set (the smaller the SUSPICIOUS the higher the precision). The first metric would be a beneficial indicator if the priority is to ensure reliability of the network (i.e. we need to make sure that the network is always supported by a passing cover) whereas the second relates to the situations in which locating the malfunctioning nodes is the top priority. We conducted several experiments to analyse the performance of the two versions of DAIM (Naïve and Greedy) under different settings and the results are presented accordingly.

We developed our own simulator in Java to analyse the performance of DAIM. Each node is associated with a thread and all communication is held through a synchronizer process. Our simulator assumes that no collisions occur by making sure that no two messages from two different nodes are sent at the same time. Invitation messages are handled using a FIFO protocol. That is, nodes that initiate an invitation message first are served first. Besides, an invitation message is cancelled if the inviter was a target of an earlier invitation. Before simulation begins, the simulator randomly generates the locations of the nodes and randomly selects the malfunctioning ones. At

each iteration, the synchronizer checks the attempts of all nodes and decides the action to be taken by each of them (i.e. sleep or be active). We define three types of attempts:

- Type 0: The node might be already sleeping or active but is ready to do what the synchronizer decides.
- Type 1: The node decided to sleep. That is, it attempted to send an invitation message to a set of neighbors (targets).
- Type 2: The node can't be sleeping because it is not covered.

The synchronizer maintains two lists: Active and Sleeping which represent the nodes that will be active (resp. sleeping in the next iteration). It first handles nodes of type 2 and adds them to the Active nodes list. Then it handles nodes of type 1 based on FIFO as follows: if the node to be handled has one of its invited neighbours in the Sleeping list then it has to be active and thus added to the Active list. Otherwise, it is added to the Sleeping list. Finally, if a node of type 0 is a target of another node in the Sleeping list then it is added to the Active list, otherwise it is added to the Sleeping list. In the next iteration, nodes that are in the Active list will be active and those that are in the Sleeping list will be sleeping. The simulation variables are: *nbNodes*, *areaWidth*, *Rc/ Rs*, *sleepingProbability*, *maxIterations*, *nbMal*, *selectionAlgorithm*.

The first set of experiments was conducted to compare the performance of the two versions with respect to each other. After setting *nbNodes* to 20, *sleepingProbability* to 0.8, and *nbMal* to 1, we performed 16 experiments for each version and determined at the end of each experiment the values of the two metrics. The results showed that Greedy is much faster in finding a passing cover whereas Naïve is much more precise in locating the malfunctioning node. Figure 3(a) shows the performance of both algorithms in terms of convergence delay. The chart can be read in the following way: a point (x,y) on the Greedy (resp. Naïve) curve means that in (y*100)% of the experiments Greedy (resp. Naïve) finds a passing cover in no more than x iterations. For example, in 100% of the experiments Greedy didn't need more than 5 iterations to find a passing cover.

The result that Greedy is much faster than Naïve in finding passing covers can be attributed to the fact that the former tends to find new covers which consist of nodes that didn't participate much in previous iterations. On the other hand, we found out that Naïve is much more precise in locating the malfunctioning node. We can see that in 100% of the experiments Naïve isolated the malfunctioning node in a set of at most 4 nodes. The result that Naïve is much more precise than Greedy in locating the malfunctioning nodes is reasonable since the former results in covers that have large overlaps with previous ones and thus the difference between the passing cover and the previous failing covers will be small.

We also used Greedy to study the impact of node density as well as sleeping probability on the convergence delay. For each value of density (40 till 65 with step 5) we conducted 5 experiments, determined the convergence delay for each and computed the average. A similar approach was done to study the impact of sleeping probability (the values ranged from 0.2 till 0.8 with step 0.2). Concerning node density, experiments show that low densities as well as high densities degrade the performance of Greedy. In terms of sleeping probability, we found out that it performs better at lower probabilities. (Fig4(a) and Fig4(b)).

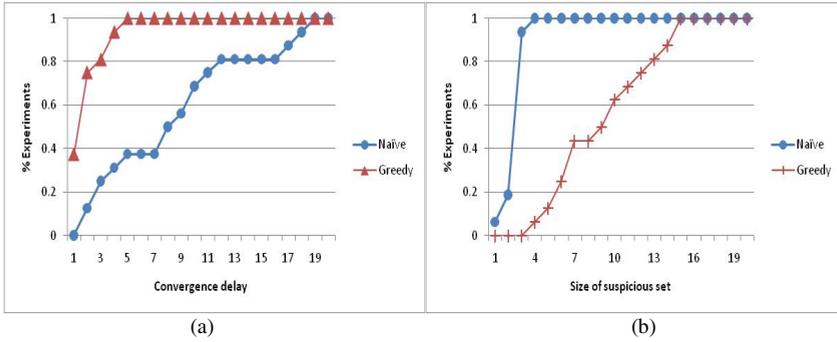


Fig. 3. (a) The convergence Delay of Naïve vs. Greedy. (b)The size of suspicious set of Naïve vs. Greedy.

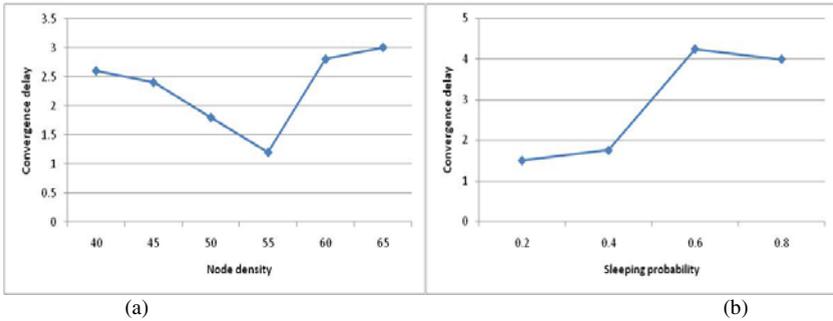


Fig. 4. (a) Convergence delay of Greedy as a function of density. (b) Delay of Greedy as a function of probability.

5 Conclusions and Future Work

In this paper, we presented an algorithm to isolate malfunctioning nodes in WSNs and provided two variants of this algorithm: Naïve and Greedy. The two variants differ in the way covering nodes are selected. We found out that Naïve is more precise in locating the malfunctioning nodes whereas Greedy is much faster in finding a passing cover. On the other hand, we noticed that its precision is slightly affected by the sleeping probability. Another important observation is that when the number of malfunctioning nodes is greater than one, both Naïve and Greedy fail to isolate them. This is left for future investigation and might be solved by forcing nodes that participate in a large number of covers to sleep. In this way, we tend to create new covers that have minimum overlap with previous covers and thus we would increase the chances of isolating the nodes simultaneously (i.e. finding a cover that doesn't contain any of the malfunctioning nodes). Another future work would be analysing the energy cost of DAIM especially that we weren't able to do it in this work because of the current limitations of our simulator.



References

1. Watfa, M., Commuri, S.: Coverage strategies in 3D wireless sensor networks. *Int. J. Distributed Sensor Netw.* 2(4), 333–353 (2006)
2. Watfa, M., Commuri, S.: An energy efficient and self-healing 3- dimensional sensor cover. *Int. J. Ad Hoc and Ubiq. Computing* 2, 121–138 (2007)
3. Watfa, M., Commuri, S.: The three dimensional coverage problem. *J. Netw.* 1(4), 10–20 (2006)
4. Slijepcevic, S., Potkonjak, M.: Power efficient organization of wireless sensor networks. In: *Proc. IEEE ICC*, pp. 472–476 (June 2001)
5. Meguerdichian, S., Koushanfar, F., Potkonjak, M., Srivastava, M.B.: Coverage problems in wireless ad-hoc sensor networks. In: *INFOCOM*, pp. 1380–1387 (April 2001)
6. Lieska, K., Laitinen, E., Lahteenmaki, J.: Radio coverage optimization with genetic algorithms. In: *Proc. IEEE PIMRC*, pp. 318–322 (September 1998)
7. Gage: Command control for many-robot systems. In: *Proc. the 19th Annual AUVS Technical Symp.*, pp. 22–24 (June 2002)
8. Gupta, H., Das, S., Gu, Q.: Connected sensor cover: Self-organization of sensor networks for efficient query execution. In: *Proc. MobiHoc*, pp. 189–200 (June 2003)
9. Ramanathan, N., Chang, K., Kapur, R., Girod, L., Kohler, E., Estrin, D.: Sympathy for the Sensor Network Debugger. In: *3rd International Conference on Embedded Networked Sensor Systems (Sensys 2005)*, pp. 255–267 (November 2005)
10. Yang, J., Soffa, M.L., Selavo, L., Whitehouse, K.: Clair-voyant: A Comprehensive Source-Level Debugger for Wireless Sensor Networks. In: *5th International Conference on Embedded Networked Sensor Systems*, pp. 189–203 (November 2007)
11. Girod, L., Ramanathan, N., Elson, J., Stathopoulos, T., Lukac, M., Estrin, D.: Emstar: A software environment for developing and deploying heterogeneous sensor-actuator networks. In: *ACM TOSN*, vol. 3(3) (August 2007)
12. Whitehouse, K., Tolle, G., Taneja, J., Sharp, C., Kim, S., Jeong, J., Hui, J., Dutta, P., Culler, D.: Marionette: Using RPC for Interactive Development and Debugging of Wireless Embedded Networks. In: *5th International Conference on Information Processing in Sensor Networks (IPSN 2006)*, pp. 416–423 (April 2006)

IPurse: An Innovative RFID Application

Mohamed K. Watfa, Manprabhjot Kaur, and Rashida Firoz Daruwala

University of Wollongong in Dubai, Computer Science Department, Dubai, UAE
sne.mohammad@gmail.com

Abstract. RFID tags are miniscule microchips, which continue to get smaller and cheaper every day. This paper introduces IPURSE, an intelligent system built on a mobile platform which keeps track of items a user carries in their purse/handbag and also alerts the user when any item is missing from the purse/handbag. IPURSE is a unique project that brings the RFID and NFC technologies together into a single system by ‘cleverly’ monitoring, reminding and alerting users about their missing items based on the current weather, reminder messages and daily items usually carried by the user. We discuss the methods, equipments and technologies used to successfully implement such an application. It includes the design evolution and technical design of the system, followed by a working prototype and experimentation. The pros and cons of the application are also discussed towards the end of this paper.

Keywords: Intelligent system, Near Field Communication, NFC, RFID, Purse.

1 Introduction

This research project is committed to the forgotten items and to an apparatus and method for reminding forgotten items. This project is a computerized system that will assist in monitoring all the items carried in a purse/handbag. Humans cannot multitask – the human brain can focus only on one thing at a time; even if the focus is for a very short period (like a nanosecond). It is virtually impossible for the human brain to pay quality attention to all tasks – it eventually forgets everything that is not committed to long term memory. As a result of which, people often tend to forget and loose things. A purse/handbag is used to contain all the personal items that one needs to carry while travelling. By keeping all items in a single container, the chances of loosing things are lessened. However, it is equally difficult to monitor all the items in the carrier at all times, especially when people are engaged in other activities. At present, people set reminders to remember to carry items and repeatedly check the items they are carrying (especially in case of important documents like passports or expensive objects) at regular intervals to ensure they are not missing or stolen.

To resolve this issue, a software application was implemented which would automate the monitoring process and also incorporate additional ‘intelligent’ features like the alert system, real-time weather check and reminders, which would help enhance the functionalities of the overall system. The system utilizes Radio Frequency Identification (RFID) and Near Field Communication (NFC) technologies. RFID technology comprises two components: tag and reader. The RFID tag is an IC

chip that can store and send information over radio frequency signals and can be applied onto different items. The RFID reader acts as an antenna for receiving and transmitting signals. NFC technology is a short range wireless connectivity technology that enables simple and safe two way interaction among electronic devices. NFC establishes faster connections compared to Bluetooth and it also provides a degree of security due to its short range communication. An NFC reader is also capable of reading RFID tags. The RFID and NFC technologies are incorporated in this software application to make it act as an intelligent system which will help in monitoring all items in the purse/handbag and also in alerting users when any item is forgotten or misplaced.

2 Method and Apparatus

For the successful working of this software application, it was important to jot down the specific requirements that the application would need to meet. The main functional requirements included providing a monitoring system, providing a web interface, providing an intelligent weather check system and an intelligent reminder system, sending alert/reminder messages at the right time to the mobile phone. Whereas, the non-functional requirements from the project included user friendliness, reliability in terms of retrieving information and sending alert messages, fast response time, error free application to offer best results and accessibility with regard to be able to create/view/modify the item list(s) at any point of time and from anywhere.

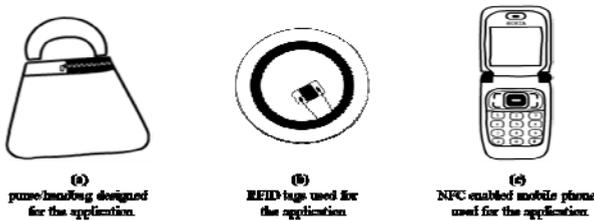


Fig. 1. Apparatus used for the IPURSE research project

IPURSE makes use of the RFID and NFC technologies as mentioned earlier in this paper. A special purse/handbag (refer Fig. 1(a)) was made that would help the users in carrying their item(s) and mobile phone and at the same time enable easy interaction between the tags and NFC mobile phone. The entry/exit point of the purse has been made in such a way that it is half-closed and half-open so as to restrict the reading space and ensure communication between the tagged item and mobile phone. The purse also has a pocket to hold the mobile phone at the front of the purse/handbag which allows the users to have fast access to the mobile phone and also easy communication with the tags. Mirfare 1K (S50 type) RFID tags (refer Fig. 1(b)) were purchased and also the Nokia 6131 NFC enabled mobile phone (refer Fig. 1(c)) which was the only available phone in the developing market for the application. Apart from the hardware, software and support was also looked at – Nokia 6131 SDK Kit and Nokia 6131 emulator were available for download at the Nokia site. JRE 1.6 and

J2ME Wireless Toolkit 2.1 were downloaded as well along with Eclipse 3.1.1 and MySQL. A few compatibility issues were faced in the beginning few weeks as the NFC based mobile phones and support for them was discontinued in the Middle East and many other countries since 2004.

When designing the system, all the requirements and functionalities of the system were taken into consideration. Different classes were defined to handle the different functionalities of the system. However, a few difficulties were faced while designing the system: One of the difficulties was to decide how frequent should the automatic scan be. The user generally would need to be reminded about the missing items before they left their home. Hence, it was decided that the user’s entry and exit points of the house would be monitored. Another difficulty faced was how often the anti-losing alerts be sent. Questions like: would the users be alerted as soon as an item was taken out of the purse; what if the users remove an item from the bag, use it for some time and then forget to keep it back in purse/handbag, then after how long would an alert be sent; how often would the alerts be sent if the item was outside the purse/handbag – once or more than once; would the alerts annoy the users, etc. were considered. Finally, after enough research, a design was sought wherein every item in the carrier would be associated with a priority and the priority would define a timer for each time, so when the time went off, the users would be alerted. Three types of priority – high, medium and low were discussed, each type associated to a timer of 15, 30 and 60 minutes respectively. The design of the system is modular and function-based. It is designed in such a way that the system can easily be extended to include new functionalities. As mentioned earlier, different classes were constructed to handle different functionalities of the system. Fig. 2 shows how the classes are linked with one another. Class Asset keeps a check at the items required according to the item list set by the user.

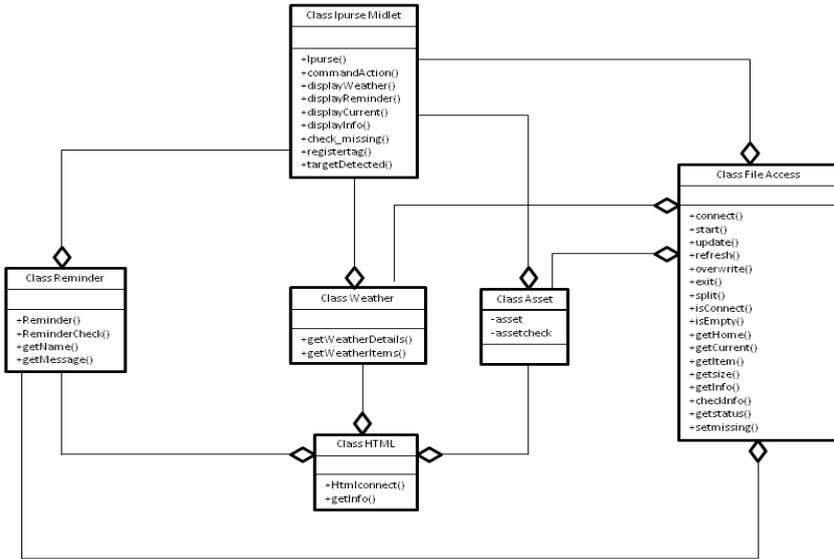


Fig. 2. Class diagram representing relationship between the classes

Class Reminder keeps in store the reminder message, the name of the person who wrote the message and also for whom it is written. Class Weather gets updated weather information for the city name specified by the user. Each of the above mentioned classes are linked to class HTML from which all the details and constraints regarding items specified by the user are received and then manipulated by the Asset, Weather and Reminder classes. Class IPURSE Midlet handles all operations made by the user via the mobile phone. All the classes, except class HTML, are linked to the class FileAccess to receive information about the items that are either present/absent from the user's purse/handbag.

3 Working of Research Project

This project needs the user, tags and mobile phone to interact together for its smooth working. At the onset, the users need to write the item's identification data, that is, name of the item onto the RFID tags and then apply these RFID tags onto the respective item(s). Once the users have tagged the item(s), they need to specify to the system which item(s) they would want to be reminded about to carry and also on which date(s) and/or day(s) of the week. The users would need to visit the website of the application in order to make use of the application. Then, the users need to create an asset list – specify item names along with their priority level and the date(s) and/or day(s) for which they want to carry them. The system has two different scenarios that it needs to deal with – one is when the users just leave home and the other is when they are away from home. The user needs to place every item that should be carried into their carrier before leaving house. When an item is placed inside the purse/handbag, the status of the item changes from 'out purse' to 'in purse' and vice versa when taken out. An RFID tag named 'home' must also be applied to the user's main door at home so that the system can know when the user leaves or comes home. When the user passes through the door, the NFC mobile phone will interact with the 'home' tag, and the status of the user will change from 'in home' to 'out home', thereby notifying the system that the user has left home. If the status of the user is set to 'out home', then the application performs different checks – asset check, weather check and reminder check one at a time. In the asset check module, the application checks the list of the items for that specific date or day as set by the user and cross checks the item(s) with the item(s) that are present in the carrier at that time. If any item is absent from the purse/handbag and needs to be carried according to the user's asset list for that date or day; then an alert message is immediately sent to the user on their mobile phone. The user can either ignore the message or collect the item from their home and place it in the purse/handbag. In the other scenario, when the user is away from home, all the items that the user placed into their carrier are monitored at all times. Since all the item(s) are inside the purse/handbag, their status is set to 'in' and when an item is taken out of the purse/handbag, the status is changed to 'out'. The application then instantly checks the priority time that the user had set for the item, for instance, the user took out their wallet for which they had set high priority (15 minutes). The application keeps track of the time and waits for the user to place the item (wallet in this case) inside the purse/handbag before the priority time that is 15 minutes runs out. But if the user has not placed back the item in the carrier within

that time, then the system sends an alert message to the user on their mobile phone informing them about the missing item. However, if the user places the item back into the purse/handbag in time, then no alert message is sent to the user at all.

4 Results

Testing was carried out for the application in two different scenarios – one when the user is at home and the other when the user is away from home. 5 participants were selected of different age groups, two being males and three females. For each scenario, same participants were called in and were briefed about the project and how it worked, after which there were given tasks to perform for application testing. At the end of the session, the participants were asked also to comment on the innovation of the application. They were asked to fill in a questionnaire based on the tasks they performed in their scenario. The answers from all the participants were collected and analyzed to get the following results:

a) When user leaves home

The participants were asked to write identification data to the tags provided to them and apply them to the respective items. Then they needed to create an asset list and specify the requirements as mentioned on the website. Later, they were asked to place their tagged items into their carrier and leave home, once by forgetting an item at home and the other time taking all items along with them. At the end of the session, the participants were asked to fill in a questionnaire which included rating (on a scale from 1-10) questions (refer Fig. 4 – x-axis for questions asked in questionnaire). The results showed that on an average all participants found the application to be easy to use including the writing data to tags. The receiving time of alert messages scored the highest on the questionnaire, while the web interface scored the least. When asked as to what they found missing in the web interface, the participants simply said that they were not accustomed to using a website as a reminder source for their required items. While the experiment was in progress, the time each participant took to put the items in to their carrier was noted, to give an average time of 20 seconds for placing 5 items in their first attempt.

b) When user is away from home

The same participants were told to purchase something at any shop/restaurant and leave the place, once by forgetting behind an item (each person left behind an item of different priority type) at the cashier's and the last time not forgetting anything. At the end of the session, the participants were asked to fill in the same questionnaire as in the scenario A with a few added questions like use of priority for items, and on the default priority time. The results for the questions repeated in this scenario were almost in sync with those in scenario A. For the additional questions, it was found from the scores that participants were pleased to have a priority timer that would allow them to assign importance levels for their items. However, the default time set for the each priority type was not liked by almost all participants. When asked for their reasons of low scoring the last priority type questions, the participants said that the time set was too long a wait and should be narrowed down for faster alerts.

c) User leaves home VS. User away from home

For this experiment, the results received from both the scenarios were compared to find which participants were more comfortable at using the application and also where. From Fig. 3, it can be seen that the scores plotted in scenario a. and scenario b. are similar for most of the questions. The major difference in the scores among the two scenarios can be seen for questions related to web interface and response time. This variance in response time could be due to the difference in the network speed as the venues were different for both the scenarios. Also, it was found that as the participants used the apparatus more, they felt more comfortable while using the application every next time, as a result of which, the average time of putting items in carrier improved over time. Additionally, the participants seemed to be pleased with the results of scenario b. than of scenario a. as shown in Fig.3. The reason deduced for this could be that people felt vulnerable of loosing items when away from home than when they are at home.

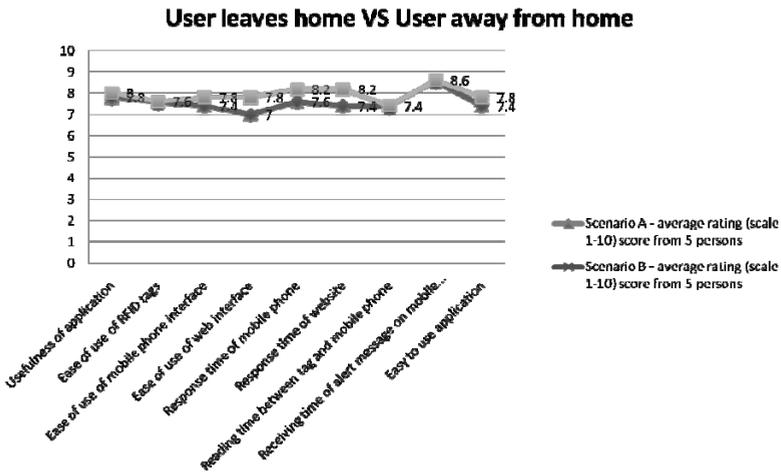


Fig. 3. Line graph showing the comparison between the two scenarios used for the evaluation of the IPURSE application

Overall, from the results it was known that the participants found this project to be very useful and user-friendly in terms of daily use and people got more comfortable with using the application every next time. The participants were also asked to suggest functionalities that they would like the application to have. All the participants who took part in the evaluation of the application, found it to be a very innovative idea. Most participants felt that the application was very useful, but said that it failed to find a solution to tackle a situation in which the user has misplaced or lost the item. In such a situation, the application cannot tell the user where the item is or at least where it was last placed.



5 Related Work and Future Directions

Some existing work with similar concept and functionalities that were found during the research phase of this project are as follows:

- Lost Item Locator [1]: this system helps in locating lost items by using a passive RFID tag and a portable reader transceiver.
- Method and System for tracking objects using GPS receiver built in active RFID [2]: this system tracks object using Global Positioning system receiver that is built in the RFID receiver.
- Electronic tether for portable objects [3]: this system issues an alarm when the distance between the module and the alert module increases by a certain value.
- LadyBag [4]: this system uses RFID technology to detect items in the bag and sends out alert with the help of LED lights if an item goes missing.

From the above stated works, it is clear that many applications have been developed but all revolving around a similar idea. Although most projects list the use of RFID technology, IPURSE is a unique project that brings the RFID and NFC technologies together into a single system. The most important advantage of this application is the intelligence part of it which ‘cleverly’ monitors, reminds and alerts users about their missing items on the basis of the weather, reminder messages received by the user and daily items carried by the user. Moreover, this application is cost-effective and users will definitely feel a sense of security while carrying their items away from home. Along with advantages, IPURSE also has a few shortcomings – most people feel that the web interface of the application is not needed. The participants at the application testing felt it was ‘too much of work and roundabout’. The participants argued that they would not want to switch between mobile phone and computers and would prefer everything to be centralized onto a single platform. Additionally, the participants were also concerned that the reading distance between the tag and mobile phone was very small. Although, this is a valid point, it is not of much importance as the reason for using NFC in this application is to provide a cheap, handy and secure communication between the RFID tag and the NFC mobile phone. At present, there are a limited number of mobile phones which comprise NFC technology. However, in the near future, most mobile phone companies are releasing handsets that will support NFC technology. Mobile manufactures namely Nokia and Apple informed in the year 2010 that they would integrate the technology into their future handsets, with NFC-enabled smart phones expected to be available by 2011 [5]. The co-CEO of RIM, had said in the Web 2.0 conference held in November 2010, that all Blackberry phones will soon be getting NFC chips in the future [6]. Alternatively, if users do not want to purchase an NFC enabled mobile phone separately, then they can purchase NFC stickers for the same. The leading suppliers of NFC mobile phone chips – Inside Contactless and NXP, Android NFC mobile phones will arrive in the first half of the year 2011 as well [7]. MyMax, the mobile NFC sticker created by TWINLINX says that any Bluetooth phone can be turned into an NFC terminal by simply placing the thin NFC sticker on its back [8].

Along with all the reviews obtained from the participants during the evaluation about the procedure and usability of the application, the participants were also asked to suggest their opinions and additional features they would want to be a part of the application. Participants wanted the application to include a tracking feature that would enable the users to track their missing or lost items. Other features that could also be added into this application would be – linking the users' social networking sites with the application's system so that reminder messages posted by the user's friends and family and birthday dates (of course with their permission) could be retrieved and intelligent reminder messages could be sent to the user in case they have missed to see the posts. As a final idea, the user's emails can be linked to the application so that the application can retrieve reminder messages or meeting details from the users' mailbox and intelligently remind the users to carry the required item(s) for the day. With the promising advent of the RFID and NFC technologies around the world and also the extensive use of mobile phones and smart phones, IPURSE will have a great impact on the community.

References

1. Stump, L.E., Ferri, L.: Lost Item Locator (2006), <http://www.freshpatents.com/Lost-item-locator-dt20060622ptan20060132308.php>
2. Chan, K.W.: Method and system for tracking objects using global positioning system (GPS) receiver built in the active radio frequency id (rfid) receiver (2009), <http://www.freshpatents.com/-dt20091217ptan20090309731.php>
3. Light, E.D., Boesch, B., Roberts, J.L.: Electronic tether for portable objects (2006), <http://www.freshpatents.com/Electronic-tether-for-portable-objects-dt20060907ptan20060197658.php>
4. LadyBag, <http://www.talk2myshirt.com/blog/archives/447>
5. Ghee, R.: NFC-enabled mobile phone – the future of the check-in process? (2010), <http://www.futuretravelexperience.com/2010/08/nfc-enabled-mobile-phones-the-future-of-the-check-in-process/#>
6. Rosoff, M., Yarrow, J.: RIM CEO Touts the Blackberry Playbook NFC for phones (2010), <http://www.businessinsider.com/live-rim-ceo-jim-balsillie-at-web-20-2010-11>
7. Clark, S.: Multiple NFC-enabled Android handsets to arrive from late, Near Field Communications World (2010) <http://www.nearfieldcommunicationsworld.com/2010/04/23/33506/multiple-nfc-enabled-android-handsets-to-arrive-from-late-2010/> (2010)
8. An NFC terminal in every pocket (2010), <http://www.twinlinx.com/>

An Epistemic Logic Based on Change and the Paradox of Unknowability

Lei Liu^{1,2}

¹ Department of Computer Science Tongji University
Shanghai, China

² Department of Computer Science and Technology Dezhou University
Dezhou, China
leiliusid@gmail.com

Abstract. Unknowable paradox threatens any theory that entails the principle of knowability. In this paper, I argue that the propositional attitude "knowing" is a state that is undergoing change when we know something that was unknown before. And the unknowable paradox is in fact a denial of this change by contradictions. There are consistent and inconsistent theories for modeling change. In a view of physicalism, we propose to relate this propositional attitude to the present time of possible worlds based on the fact that the tense of its definition in order to avoid a inconsistent theory of change. It shows that theorem 5 of Fitch turns out to be that if there is an unknown truth now then it is unknown now that it is unknown truth now. It is compatible with the principle of knowability because speaking specially, it is still possible that it is known in a future time that it is an unknown now.

Keywords: Paradox of knowability, Modal logic, epistemic logic.

1 Introduction

Verificationist hold that a sentence A is true if and only if it is possible to verify (prove) A . From this anti-realist point of view, it seems plausible that if A is true, then A is knowable: if it is possible to prove A , it is possible to know that A . Since its first published presentation in [9] and its rediscovery by Hart[12] and McGin[13], the paradox of knowability has received considerable attention, see, for example, articles [1][3][5][6][7][8][10][11][14][15][16][17][18][19][20][22]. We present the proof here for further discussions.

Let K be the epistemic operator "it is known by someone at some time that." Let \diamond be the modal operator "it is possible that".[2]

Suppose the knowability principle (KP), more carefully, all truths are knowable by somebody at some time:

$$(KP)\forall p(p \rightarrow \diamond Kp).$$

And suppose that we are non-omniscient, that there is an unknown truth:

$$(NonO)\exists(p \wedge \neg Kp)$$

If this existential is true, then so is an instance of it:

- (1) $p \wedge \neg Kp$
 (2) $(p \wedge \neg Kp) \rightarrow \diamond(K(p \wedge \neg Kp))$ from(1) and (kP)
 (3) $\diamond(K(p \wedge \neg Kp))$ from(2) and (NonO)

The problem is that it can be shown, independently, that it is impossible to know (1), thus (3) is false.

The independent result presupposes two very modest epistemic inferences: first, a conjunction is known, only if the conjuncts are known. Second, a statement is known, only if it true. Respectively,

- (A) $K(p \wedge q) \vdash Kp \wedge Kq$
 (B) $Kp \vdash p$

Also presupposed is the validity of two modest modal inferences: first, all theorems are necessarily true (the rule of necessitation). Second, if it is necessary that p, then it is impossible that p (the definition of \square). Respectively,

- (C) $If \vdash p, then \square p$
 (D) $\square p \vdash \neg \diamond p$

we can get the independent result:

- (4) $K(p \wedge \neg Kp)$ Assumption.
 (5) $Kp \wedge K(\neg Kp)$ from (4), by (A)
 (6) $Kp \wedge \neg Kp$ from (5), applying (B) to the right conjunct
 (7) $\neg K(p \wedge \neg Kp)$ from (4)-(6), by discarding the assumption (4)
 (8) $\square \neg K(p \wedge \neg Kp)$ from (7), by (C)
 (9) $\neg \diamond K(p \wedge \neg Kp)$ from (8), by (D)

Notice that (9) contradicts to (3). So a contradiction follows from (KP) and (NonO). The advocate of the view that all truths are knowable must deny that we are non-omniscient:

- (10) $\neg \exists p(p \wedge \neg Kp)$
 (11) $\forall p(p \rightarrow Kp)$ from (10).

2 An Solution Based on a View of Change

Now we consider (8). Combined with the assumption (7), it can be read as that if it is known that it is an unknown truth, then it is impossible to be known that it is an unknown truth. That is to say, the propositional attitude towards that it is an unknown truth can't be changed from unknown to known. This is exactly a denial of the possibility of change of the state of the propositional attitude.

Though it is a change, it is a metaphysical change [4][21] because it could be a change in the monadic or internal or intrinsic properties of the thing. But the fact is that this change does have relation to time as exemplified by the definition of the epistemic operator K . $Kp \wedge \neg Kp$ is not permitted in epistemic logic. Thus, combined with the view of physicalism, we treat the propositional attitude similar to an object-language concept. This change is not cyclic according our definition. A logic which

could characterize acyclic change character of the propositional attitude would be a little complex.

Now we take a careful look at why the proof of the paradox of Knowability is denial to that propositional attitude change.

Recall that K be the epistemic operator "it is known by someone at some time that." The tense of the drill "it is known by someone at some time that" is present. So if there is someone who knows or knew that at some time t , then t must not be later than the present time. After all, this is more clear when we express $\neg K$, which is read as "Nobody ever knows that." Hence, let us add this constraint explicitly at first.

We view time as an assemblage of point-instants to avoid the contradiction of change in this paper. At first we now relates the propositional attitude K with the time index of it.

Let $K_{t_s} p$ iff $\exists x \exists t (t \leq t_s) K_t p$, where t_s represents a given point-instant. and $t \leq t_s$ means that t is not later than t_s .

It reads that it is known by somebody at some time not later than the given point-instant t_s that p . It can be also be read as (the propositional attitude towards p) at t_s is "known".

From this definition, we could induce two rules which characterize the nature of our knowledge, which is that our knowledge increases in a monotonic way:

(*INK*) $\forall p (K_{t_{past}} p \rightarrow K_{t_p} p)$, where $t_{past} < t_p$.

(*NonDecK*) $\forall p (\neg K_{t_p} p \rightarrow \neg K_{t_{past}} p)$, where $t_{past} < t_p$. Derived from (*INK*) directly.

And

(*B'*) $K_{t_p} p \vdash p$ this is true clearly.

(*A'*) $K_{t_s} (p \wedge q) \vdash K_{t_s} p \wedge K_{t_s} q$

The first can be understood as that if it was known that p at an earlier time, it is known that p now. The second can be understood as that if it is not known that p now, it was not known that p at an earlier time too. Of course, the above two formulas is a generalization of these kind of cases.

And an obvious observation is:

(*NonC*) $\forall p (\neg K_{t_1} p \wedge K_{t_2} p) \rightarrow t_1 \neq t_2$.

Notice that these conclusions are based on the assumption that the whole of actual beings (possible beings in possible worlds) can continue to exist and share all the knowledge, which is endorsed by previous epistemic logic too.

Fact: We notice that we do know that we were Non-omniscient before. It can be written as

(12) $K_{t_p} (p \wedge \neg K_{t_{past}} p)$, where $t_{past} < t_p$, t_{past} represents a past time, t_p represents the present time, or to say, now, and $t_{past} < t_p$.

According to similar arguments with (4)–(7) we could get the following arguments:

(4') $K_{t_{past}} (p \wedge \neg K_{t_{past}} p)$ Assumption.

(5') $K_{t_{past}} p \wedge K_{t_{past}} (\neg K_{t_{past}} p)$ from (4'), by (A')

(6') $K_{t_{past}} p \wedge \neg K_{t_{past}} p$ from (5'), applying (B') to the right conjunct

(7.1) $\neg K_{t_{past}} (p \wedge \neg K_{t_{past}} p)$ from (4')–(6') by discarding the assumption (4')

That is to say, From (7.1), which we accepted, to (12), with the elapse of time from t_{past} to t_p , our propositional attitude towards the truth $p \wedge \neg K_{t_{past}} p$ did changed from $\neg K$ to K . It can be interpreted like this. p was not known by any actual being at the instant t_{past} . It is known by an actual being at the instant t_p which is now. it reads that an actual being knows now that p and it was not known ever by any actual being. Normally, this is the case that now we know a truth we didn't know before.

There are four cases of (7) $\neg K(p \wedge \neg Kp)$ after we index Kp with the time t_s :

(7.1) $\neg K_{t_{past}} (p \wedge \neg K_{t_{past}} p)$

(7.2) $\neg K_{t_{past}} (p \wedge \neg K_{t_p} p)$

(7.3) $\neg K_{t_p} (p \wedge \neg K_{t_{past}} p)$

(7.4) $\neg K_{t_p} (p \wedge \neg K_{t_p} p)$

As we proved, (7.1) is (necessarily, which will be proved after) true. Similarly, (7.4) is (necessarily) true too. (7.2) can proved to be (necessarily) true by applying (*NonDecK*) to (7⁴).

(7.3) is false as the fact (12) and (*NonC*) tells us. Thus we can infer that at time t_{past} , (12) is possible. But according the proof of the paradox of knowability, (12) is impossible too. This is because the arguments from (4)–(7) is based on the neglecting the time when the two propositional attitude towards p and $p \wedge \neg Kp$ respectively are related. Thus mistakenly take (7.3) and (7.2) the same as (7.1) and (7.4). Recall that Recall that K be the epistemic operator "it is known by someone at some time that", neglecting the time index to K , it means that for all time, if there exist a time it is known by somebody that. As we assume time is universal, it is inevitably that we know all the truths we could know. Thus combining with the principle of knowability, all truths are known.

We now relate the epistemic operator with the present time of possible worlds in order to avoid to make this mistake.

Let $K_w p$ iff $\exists x \exists t (t \leq t_w) Kp$ where t_w represents the present time of the possible world w . It reads that it is known by somebody in world w at some time not later than present that p or there exists or existed somebody whose propositional attitude to p at a time t_w is K . Specially, $K_{t_a} p$ reads that it is known by an actual being and at an actual time that p .

The (KP) rule is changed into $(KP') \forall p (p \rightarrow \diamond K_w p)$.

and (B) is changed into $K_{t_a} p \vdash p$.

Then (3) is changed into:

(3') $\diamond (K_{t_a} (p \wedge \neg K_{t_a} p))$

(*NonO*) is changed into (*NonO'*) $\exists (p \wedge \neg K_{t_a} p)$

Then (3) is changed into:

(3') $\diamond (K_{t_a} (p \wedge \neg K_{t_a} p))$

Then there is no contradiction in supposing that some possible being at some possible time knows that p and it had never been known by an actual being at an actual time.

Now take a look at the proof of the paradox of unknowability:

(4^{*}) $K_{t_w}(p \wedge \neg K_{t_w} p)$ Assumption.

(5^{*}) $k_{t_w} p \wedge K_{t_w}(\neg K_{t_w} p)$ from (4^{*}), by (A')

(6^{*}) $K_{t_w} p \wedge \neg K_{t_w} p$ from (5^{*}), applying (B') to the right conjunct

As the fact (12) justified, (6^{*}) is not necessarily a contradiction when $t_a < t_w$ if they are interpreted in the possible world w to have the relation of sequence, thus the proof of (7) fails.

Concluding (7) from (6^{*}) resembles Zeno's arguments about motion directly because if we interpret motion in relation with point-instants of time: At each instant, the object would be at a place identical to itself, but motion is still possible because at different instants, it can be at different places.

But for (7.1), (7.2), and (7.4), they are true, thus the respective version of them to (8) are:

(8.1) $\Box \neg K_{t_{past}}(p \wedge \neg K_{t_{past}} p)$ from (7.1), by (C)

(8.2) $\Box \neg K_{t_{past}}(p \wedge \neg K_{t_p} p)$ from (7.2), by (C)

(8.4) $\Box \neg K_{t_p}(p \wedge \neg K_{t_p} p)$ from (7.4), by (C)

From (KP') and (NonO'), we get $\Diamond(K_{t_p}(p \wedge \neg K_{t_{past}} p))$. It can be read as it is possible that it is known at a time t_p that p and it was not known that p at a time t_{past} which is before the time t_p .

Combine (8.1) and (8.4), we get a interesting observation.

$$\Box \neg K_{t_w}(p \wedge \neg K_{t_w} p) \quad (EU)$$

Clearly, from (EU), we could infer that (8.2) is true too.

It can be interpreted as if there is an unknown truth now, then it is impossible("not the case" would be more welcome because normally we assume that there is no change at the same time for the same being) that it is known now that it is an unknown truth now. Combining with (3'), it can be interpreted as this:

If there is an unknown truth p now then that it is an unknown truth now is itself unknowable (unknown would be more welcome because normally we assume that there is no change at the same time for the same being) now. But it is possible that it is known at a possible time that p and it is not known now (the present time of the actual world). Specially, it is possible that it is known in a future time that p and it is not known now.

Replace K_{t_a} with K_{t_w} , we could get a more general observation:

$$\Box \neg K_{t_w}(p \wedge \neg K_{t_w} p) \quad (EU')$$

It reads that in any possible world, if there is an unknown truth p at present time (in this possible world) then that it is an unknown truth at present is itself unknowable (unknown would be more welcome because normally we assume that there is no change at the same time for the same being) at present.

In conclusion, in any world, it is impossible that it is known at present that subjects in this world are omniscient at present. Specially, if (*NonO*) is true now (at present in the actual world), it is possible that it is known in the future (at some possible world by some possible being that it is true now (at present in the actual world)).

Nevertheless, this is based on a consistent view of the change of the state of propositional attitude. It is also reasonable to interpret epistemic logic from a inconsistent view of it.

3 Discussion

It emerges that the connection between change and inconsistency is deep. If we take a inconsistent view of change, like to take time as composed of nested intervals, it seems that paraconsistent logic is more reasonable and will be more complicated. The nature of change is clearly not a easy issue.

The results of the article[23] resembles some previous results.

What is true may come to be known (VT)

But it has a restriction that is CK:

Kp is consistent for any proposition p (CK).

By bringing to bear this analogue and resources from dynamic epistemic logic, he suggested that the failure of the knowability principle only teaches us more about the subtleties of knowledge and communication.

References

1. Beall, J.C.: Fitch's Proof, Verificationism, and the Knower Paradox. *Australas. J. Philos.* 78, 241–247 (2000)
2. Brogaard, B., Salerno, J.: Fitch's Paradox of Knowability (2002), <http://plato.stanford.edu/entries/fitch-paradox/>
3. Brogaard, B., Salerno, J.: Clues to the Paradoxes of Knowability: Reply to Dummett and Tennant. *Analysis* 62, 143–150 (2002)
4. Chris, M.: Change (2006), <http://plato.stanford.edu/entries/change/>
5. Cozzo, C.: What We Can Learn from the Paradox of Knowability. *Topoi* 13, 71–78 (1994)
6. DeVidi, D., Kenyon, T.: Analogues of Knowability. *Australas. J. Philos.* 81(4), 481–495 (2003)
7. DeVidi, D., Solomon, G.: Knowability and Intuitionistic Logic. *Philosophia* 28, 319–334 (2001)
8. Dummett, M.: What is a Theory of Meaning? In: Evans, G., McDowell, J. (eds.) *Truth and Meaning*, ch. 4. Clarendon Press (1976)
9. Fitch, F.: A Logical Analysis of Some Value Concepts. *J. Symbolic. Logic* 28, 135–142 (1963)
10. Hand, M.: Knowability and Epistemic Truth. *Australas. J. Philos.* 81(2), 216–228 (2003)
11. Hand, M., Kvanvig, J.: Tennant on Knowability. *Australas. J. Philos.* 77, 422–428 (1999)
12. Hart, W.D.: The Epistemology of Abstract Objects: Access and Inference. *Pro. Aristot. Soc.* 53, 153–165 (1979)
13. Hart, W.D., McGinn, C.: Knowledge and Necessit. *J. Philoso. Logic* 5, 205–208 (1976)

14. Hegel, G.: 1812, *Wissenschaft der Logik*, see A. Miller (tr) *Hegel's Science of Logic*. Allen and Unwin, London (1969)
15. Kvanvig, J.: The Knowability Paradox and the Prospects for Anti-Reality. *Nous* 29, 481–499 (1995)
16. Mackie, J.L.: Truth and Knowability. *Analysis* 40, 90–92 (1980)
17. Melia, J.: Anti-Realism Untouched. *Mind* 100, 341–342 (1991)
18. Nozick, R.: *Philosophical Explanations*, ch. 3, Cambridge (1981)
19. Percival, P.: Fitch and Intuitionistic Knowability. *Analysis* 50, 182–187 (1990)
20. Percival, P.: Knowability, Actuality and the Metaphysics of Context- Dependence. *Australas. J. Philos.* 69, 82–97 (1991)
21. van Inwagen, P.: *Metaphysics* (2007), <http://plato.stanford.edu/entries/metaphysics/>
22. Rabinowicz, W., Segerberg, K.: Actual Truth, Possible Knowledge. *Topoi* 13, 101–115 (1994)
23. van Benthem, J.: What One Come to Know. *Analysis* 64(2) (2004)

Criminal Investigation DSS Based on Trust Intuition Analysis Model

Lihua Duan

Department of Information
Liaoning Police Academy, Dalian, 116036, P.R. China
lnpolice@126.com

Abstract. On purpose of improving the research in trust intelligence systems when the knowledge in hand is not sufficient, an trust intuition model (EIM) based on human-computer cooperative is presented. From the initial intuition process space defined by the primitive experience, a series of interactive experience learning systems (IELS) with various reductive levels are created. For, each IELS, the rule sets with respective belief degree are induced and saved. The paper introduces experience mapping intuitive inversion (EMII) method to the criminal investigation, and poses a skeleton of cooperative reasoning. Through the trusted intuition relationship construction of practical crime model and on-the-spot model, it sets up a couple of mapping-inversion models in information acquisition of human-computer interaction.

Keywords: IDSS, criminal investigation, intuition learning, intuition-formal models, cooperative reasoning, human-computer interaction.

1 Introduction

To mimic the problem solving capacity of human being is one of the most basic and important task of artificial intelligence (AI). Such capacity in case solving is neither merely a pure reasoning algorithm, not completely relies on some formatters. The establishment of cooperative relationship can be regarded to be the identification and evaluation of fact inversion and evolution.

From the research achievements of IDSS in recent years [6], the original intention of researchers is that computers can substitute for the intelligence of human beings, thus acquire the decision-making capacity of experts and also overcome the limitation of experts in the field, so that to reach the level of true experts. However, to study the IDSS as an issue in computer science has hampered the system development. No matter in the aspects of knowledge acquisition and expression, or in uncertain reasoning, though great research achievements have been obtained (especially the introduction of artificial neural network and fuzzy system provides many new tools for development of IDSS), few successful IS are available [1, 2]. Many scholars believe that the key to build IDSS is the selection and effective use of knowledge [3, 4, and 5]. The “effective use” means whether the rule in the system synchronize with the thinking of the actual users, which is also the difficulty in IDSS.

At present, most IDSS is designed manually based on past experience of their intuition. Since the number of possible intuition is very large for realistic applications of reasonable complexity, heuristics designed manually may not work well when applied in new problem instances. Further, there is no systematic method to evaluate the effectiveness of IDSS designed manually. For these reasons, an automated method for discovering the proper IDSS for a particular application is very desirable. This leads to the development of our system for automated learning of intuition [1, 2, 3, and 4].

The rest of this paper is organized as follows: In Section 2 we give an outline of the design of learning experiments methodology. In Section 3 we describe the overall structure of the interface. We also discuss in rather more detail the key modules and routines contained in the interface. In Section 4 we demonstrate use of the interface in conducting the interactive learning system of a particular queuing model. The final section 5 concludes the paper and points out further work.

2 Literature Survey

2.1 CIS Based on Knowledge

Research results indicate the need for wide application of computer-based techniques, e.g., databases, neuron networks, spreadsheets, and expert systems to problems in the criminal investigation system (CIS) [5, 8]. However, computer-based methodologies have not been a focus or researched extensively, and specifically, in the area of CIS.

Traditional decision support systems (DSSs) for crime investigation are difficult to construct because of the almost infinite variation of plausible crime scenarios [3, 4]. Thus existing approaches avoid any explicit reasoning about crime scenarios. They focus on problems such as intelligence analysis and profiling [6]. propose a novel model based reasoning technique that takes reasoning about crime scenarios to the very heart of the system, by enabling the DSS to automatically construct representations of crime scenarios. It achieves this by using the notion that unique scenarios consist of more regularly recurring component events that are combined in a unique way. It works by selecting and instantiating generic formal descriptions of such component events, called scenario fragments, from a knowledge base, based on a given set of available evidence, and composing them into plausible scenarios. This approach addresses the robustness issue because it does not require a formal representation of all or a subset of the possible scenarios that the system can encounter. Instead, only a formal representation of the possible component events is required. Because a set of events can be composed in an exponentially large number of combinations to form a scenario, it should be much easier to construct a knowledge base of relevant component events instead of one describing all relevant scenarios.

[Ref. 3, 2008] shows that there are three attribute of the recognition to the criminal investigation DSS, experience, knowledge and intuition. The attribute of experience reflects the recognition to the characteristics of the basic behavior. The attribute of knowledge reflects the learning recognition to the characteristics of the intelligent behavior. The attribute of intuition reflects the fuzzy dynamic recognition to the characteristics of the intelligent behavior.

People's experience provides basic intelligence for solve many problems. When the recognitions are different, the basic intelligence is different as well. The tracing to the problem's conditions of the past can propose an experience set. In an artificial system, different people have different behaviors and stories, thus different experiences. Sometimes experiences are called a kind of recognitions; but as the level of recognition is different, the experience of the human is also different. The intelligence of the human is selected and decided by the experience of the human, and the reasonability of the experience's selection is also a meaningful question for discussion.

2.2 Overview of Reasoning Technique

[Ref. 7, 8] use a novel model based reasoning technique, derived from the existing technology of compositional modelling, to automatically generate crime scenarios from the available evidence. Consistent with existing work on reasoning about evidence the method presented herein employs adductive reasoning. That is, the scenarios are modelled as the causes of evidence and they are inferred based on the evidence they may have produced. The goal of the DSS described in this paper is to find the set of hypotheses that follow from scenarios that support the entire set of available evidence. This set of hypotheses can be defined as:

$$H_E = \{h \in H : \exists s \in S, (\forall e \in E, (S \mapsto e)) \wedge (S \mapsto h)\}$$

where H is the set of all hypotheses (e.g. accident or murder, or any other important property of a crime scenario) S is the set of all consistent crime scenarios, our mini-stories in the example E is the set of all collected pieces of evidence.

Figure 1 shows the basic architecture of the proposed model based reasoning DSS. The central component of this architecture is an assumption based truth maintenance system (ATMS). An ATMS is an inference engine that enables a problem solver to reason about multiple possible worlds or situations. Each possible world describes a specific set of circumstances, a crime scenario in this particular application, under which certain events and states are true and other events and states are false. What is true in one possible world may be false in another. The task of the ATMS is to maintain what is true in each possible world.

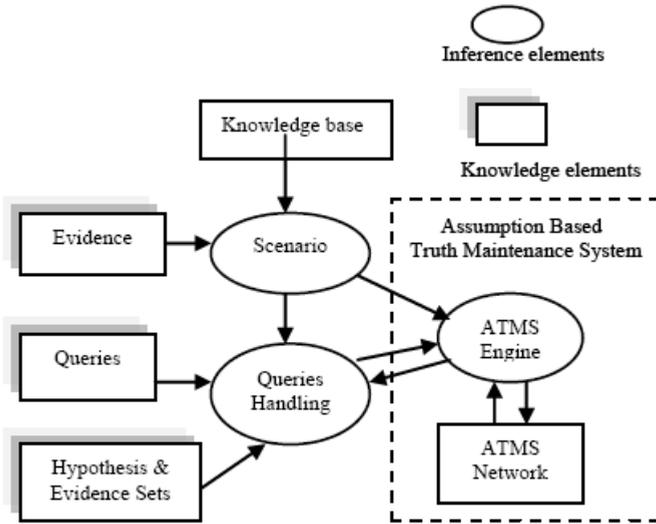


Fig. 1. Basic Architecture of the model based reasoning for crime scenarios. (See [7])

The ATMS is employed by two separate problem solvers. First, the scenario instantiate constructs the space of possible worlds. Given a knowledge base that contains a set of generic reusable components of a crime scenario (think of the locked door, the jealous partner etc) and a set of pieces of evidence (Peter’s fingerprints, John’s DNA etc), the scenario instantiate builds a space of all the plausible crime scenarios, called the scenario space, that may have produced the complete set of pieces of evidence. This scenario space contains all the alternative explanations to the preferred investigative theory. Once the scenario space is constructed, it can be analysed by the query handler. The query handler can provide answers to the following questions: [8].

3 Trust Intuition Analysis

3.1 From Basic Intelligence to Trust Intelligence

The purpose of this section is to discuss an “Extended Intelligence (EI)” concept to establishing the new ideas and approach of IDSSs. This approach will include basic intelligence, learning-intelligence and trust intelligence.

There are many different definition of intelligence in IDSSs, but none of them give the answer acceptable by a scientific community. First of all, intelligence is a fuzzy term. In some cases it is very difficult to draw a line between intelligent and non-intelligent natural and artificial systems. For example, biological adaptation or any kind of evolution can be presented as learning intelligent ability or non-intelligent process. It is difficult to determine when DSSs became an AI system. All intellectual activities are triggered by the goal. A DSS can be intelligent only in relation to a defined goal [3].



In fact, there are two components of intelligence: experience-based intelligence (basic-intelligence) that is inherited at past activities process, and knowledge-based intelligence that can be improved by learning (learning-intelligence).

All kinds of intellectual activities in the specific area are based on knowledge, but intelligence is not knowledge. Knowledge is a “tool” of intelligence. If you don’t understand a goal, you are not capable to reach it. An ability to learn is an important intellectual ability that can improve knowledge. Knowledge reinforces intellectual activities.

There are three attributes of the recognition to the intelligent system, experience and knowledge. The attribute of experience reflects the recognition to the characteristics of the basic behavior. The attribute of knowledge reflects the learning recognition to the characteristics of the intelligent behavior. The attribute based on experience and knowledge is called “trust” attribute of intelligence, on the other word, the intelligence with trust attribute is called “Trust Intelligence” (EI).

3.2 Measurement of Experience-Based Intuition

In this section we are mainly, but not exclusively situated in the descriptive perspective of analysis of the deciding process and we’re therefore trying to accredit the following ideas:

- Experience, intuition and empirical Knowledge can have a positive influence in practice, in successfully completing a deciding process (although those entire three elements cannot be formalized and included in a normative model);
- Reaching maximum efficiency by the decider may be the outcome of an apparently limited irrationality based on experience, intuition and imagination;
- The promoters of the deciding process formalizing models and procedures (the normative perspective) should have a more flexible attitude towards the deciding act (not everything excessively formalized automatically leads to maximum efficiency).

At present, most Intuition learning (IL) [6] is designed manually based on past experience of their intuition. Since the number of possible Intuition is very large for realistic applications of reasonable complexity, learning system designed manually may not work well when applied in new problem instances. Further, there is no systematic method to evaluate the effectiveness of IL designed manually. For these reasons, an automated method for discovering the proper IL for a particular application is very desirable. This leads to the development of our system for automated learning of intuition.

4 Application of Extended Knowledge Management

In order to evaluate the usefulness and the usability of the extended intelligence in a crime case setting, we conducted a study to determine the feasibility of this intelligence analysis tool in real criminal investigations.

First, The general area of knowledge management (KM) has attracted an enormous amount of attention in recent years. Although it has been variously defined, it is evident that knowledge management exists at the enterprise level and is quite distinct from mere information. Also apparent in this area are the challenges that knowledge

management poses to an organization. In addition to being difficult to manage, knowledge traditionally has been stored on paper or in the minds of people. The KM problems facing many firms stem from barriers to access and utilization resulting from the content and format of information. These problems make knowledge management acquisition and interpretation a complex and daunting process. Nevertheless, knowledge management information technologies have been developed for a number of different applications, such as virtual enterprising, joint ventures, and aerospace engineering. The same problems of knowledge management exist at the specialized organizations of police department. Many record management systems for crime control agencies contain a large amount of data for each case or incident, but although data may be available, they are not available in a form that makes them useful for higher level processing. For example, the ideal knowledge management system should be able to provide information about problems that have not been identified previously, and thus be able to give innovative and creative support for new investigations. The conversion of information to knowledge is an important concern for police department agencies. Information is a product that is designed with a purpose in mind, while data serve as the ingredients in this product. Furthermore, addressing the conversion of information to useful and easily understandable knowledge is a powerful aspect of knowledge management that has thus far been missing from most law enforcement information systems.

A basic task for detectives and crime analysts at Dalian Police Department (DPD) is to create knowledge from information. In this case, information is made up of approximately 1.5 million criminal case reports, containing details from criminal events dating back to 2003. Tacit knowledge has also been described as the means through which new knowledge is generated as well as the practical knowledge used to perform a task. It is tacit knowledge that is used as investigators try to tie together information to solve cases and crimes. This ability to combine information to create knowledge is often hampered by the amount of information that exists.

The purpose of this paper is to explore the development of a system of knowledge management and intuition learning that can provide the functionality of trust intelligence analysis that currently do not exist in the RMS system. This system is designed to serve as a type of knowledge detectives and has been evaluated in a real life context. Its findings also are discussed.

5 Conclusions

From this pilot study, we conclude that the use of intuition learning as a crime knowledge management and intelligence analysis tool in a law enforcement environment is quite promising. This paper discusses the critical issues in establishment of interactive intuition learning system that should be paid attention to through practice of criminal investigation work. The development of the interactive intelligent analysis system must be grounded on identification, otherwise this work is of little significance or value. Simultaneously, the intuitive reasoning should be distinguished from experiential and intuition reasoning. For different cases, experiential reasoning is variable. Only by combining the two together with intuition to reach cooperative reasoning can they possibly play their roles in reality. As a matter of

fact, the key of founding the system of trust automatic crime detects reasoning is to make good use of the EPIP. However, it requires good man-computer functions. Meanwhile, the statistical information function is also needed. The studies and practices show that, the using of IRMI principle combining intuitionistic automatic reasoning surely has a promising future.

In future work, the method presented here will be trusted upon. Firstly, the representation formalisms employed to describe states and events in intuitive process of criminal investigation will be elaborated. As described earlier, the intuitionistic fuzzy set of states and events that constitute a scenario are restricted by the consistency requirements. This paper introduced a generic means to represent when inconsistencies occur and to prevent inconsistent experience and knowledge from being considered when hypotheses are generated and evidence collection strategies are constructed. When reasoning about related events that take place over experience and intuition, the experience process of the intuition are an important source of such inconsistencies. To avoid overcomplicating this paper, the important issues of knowledge and intuition reasoning were not considered, but will be addressed in future work. Secondly, methods are under development to assess the relative likelihoods of alternative learning system. Several methods to expand the intuition entropy based decision making techniques employed by model based intuitionistic fuzzy diagnosis techniques have been presented in other papers.

Acknowledgments. I gratefully acknowledge and thank Professor He Ping and Dr. He Jiantong for their guidance throughout this project, along with He Ping and He Jiantong for their reviews. This work was completed with the support of a research grant from the Ministry of Public Security.

References

1. Dijkstra, J.J.: The influence of an expert system on the user's view: How to fool a lawyer. *New Review of Applied Expert Systems* 1, 123–138 (2006)
2. Keppens, J., Schafer, B.: Knowledge based crime scenario modelling. *Expert Systems with Applications* 30(2), 203–222 (2006)
3. Adderley, R.W., Musgrove, P.: Police crime recording and investigation systems – A user's view. *An International Journal of Police Strategies & Management* 24(1), 100–114 (2001)
4. Buchanan, B.G., Shortliffe, E.H. (eds.): Shortliffe, Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project. Addison-Wesley, Reading (1984)
5. Oatley, G.C., Ewart, B.W.: Crimes analysis software: pins in maps, clustering and Bayes net prediction. *Expert Systems with Applications* 25(4), 569–588 (2003)
6. Donalds, C.M., Osei-Bryson, K.-M.: Criminal Investigation Knowledge System: CRIKS. In: *Proceedings of the 39th Hawaii International Conference on System Sciences*, pp. 1–9 (2006)
7. He, P.: Fuzzy relationship mapping inversion and the automatic reasoning of crime detective. *Pattern Recognition and Artificial intelligence*, 58–62 (2003) (Chinese)
8. Bex, F.J., Prakken, H., Reed, C., Walton, D.N.: Towards a formal account of reasoning about evidence: argumentation schemes and generalisations. *Artificial Intelligence and Law* 12, 125–165 (2003)

9. He, P.: Fuzzy relationship mapping inversion and automatic reasoning of crime detective. In: Artificial Intelligence Application and Innovations (AIAI 2005), pp. 691–700. Springer, Heidelberg (2005)
10. Li, J., He, P.: Extended automatic reasoning of criminal investigation. In: International Conference on Industrial Mechatronics and Automation, pp. 546–649. IEEE Computer Society (2009)
11. Bex, F.J., Prakken, H., Verheij, B.: Anchored narratives in reasoning about evidence. In: van Engers, T.M. (ed.) Legal Knowledge and Information Systems, JURIX: The Nineteenth Annual Conference, pp. 11–20. IOS Press (2006)
12. van den Braak, S.W., Vreeswijk, G.: AVER: Argument visualization for evidential reasoning. In: van Engers, T.M. (ed.) Legal Knowledge and Information Systems, JURIX 2006: The Nineteenth Annual Conference, pp. 151–156. IOS Press, Amsterdam (2006)
13. Crombag, H.F.M., van Koppen, P.J., Wagenaar, W.A.: Dubieuze zaken: De psychologie van strafrechtelijk bewijs, 2nd edn. Uitgeverij Contact, Amsterdam (1994)
14. Loui, R.P.: Defeat among arguments: a system of defeasible inference. *Computational Intelligence* 2, 100–106 (1987)
15. Lucas, P.: Symbolic diagnosis and its formalisation. *The Knowledge Engineering Review* 12, 109–146 (1997)

Crime Data Mining Based on Extension Classification

Weidong Tao

Department of Information
Liaoning Police Academy
Dalian, China
lnpolice@126.com

Abstract. The problem of knowledge management in the knowledge intensive and time critical environment of police department has posed an interesting problem for information technology professionals in the field. In this paper, we propose an extension formal analysis algorithm for maintaining temporal association rules with cooperative numerical attributes by using the intuition learning method. The new algorithm has been implemented to support the discoveries of crime patterns in a district of Dalian of China. We have also experimented with another real life database of courier records of a criminal investigation. Future development efforts for data analysis and knowledge discovery project will also be discussed.

Keywords: extension data mining, computing data mapping, intuition learning, intuitive data inversion, crime pattern discovery.

1 Introduction

For the recent years, data mining and knowledge discovery plays an important role in crime analysis and has been becoming applicable in many areas of police information technology. But, conventional environment of information conducting is “man round mathematical model”. It is difficult to deal with the information of being of brain thinking procedure. It make person have definite dependence that computer bring into play great power while dealing with information. Whether, when the result from computer is different from reality significantly, person will doubt it. It is the reason that occur incompatible question when we applicant routine data mining ways on solving the conducting of crime information. The development of computer technologies during the past few years has enabled many organizations to improve both the understanding and the dissemination of data mining methods. The development of powerful data mining and knowledge discovery allows information to be organized in a manner that improves access to it, increases speed of retrieval, and expands searching flexibility.

Crime analysis agencies across the China have begun to adopt innovative knowledge management technologies to aid in the management of policeman department [1, 2]. Such technologies can serve as intelligence tools to combat crime activity by aiding in case investigation or even predicting criminal activity. In this research project we developed and evaluated one such crime information management

tool in the context of real criminal investigation by real law enforcers, fully acknowledging that many issues as well as obstacles must be addressed to ensure the successful deployment of this and similar data analysis technologies[3].

We propose an approach that incorporates an extension data mining module to identify relevant system inputs and to analyze the way these inputs interact within the system. This extension data mining model uses traditional methods to analyzed relation between data and intuition. Its principal contribution makes use of a cooperative formal concept procedure that is activated at certain intervals during the intuitive learning process in order to make use of information obtained during that process, with the goal of speeding the search for extension data analysis.

The paper is organized as follows. In the next section, we discuss our proposed extension data mining models, and the manner in which the dynamic aspects of data mining interact with the data-intuition process. The section on an integrated approach to simulation data-intuition illustrates our proposed approach in detail, focusing on an example using extension-based simulation in a crime data mining system analysis application. The final section presents our concluding remarks and avenues for future research.

2 Crime Analysis and Classification

2.1 Crime Data Mining and Knowledge Discovery

An important modification was the identification of certain fields for analysis and learning. This included both fields that can be used as search terms and fields that are returned by the Learning system of crime knowledge. The relationships between search terms are an important issue not focused upon in prior concept space applications. For an investigator, relationships between objects must not only be identified, but for crime analysis must also be explicitly understood. The format of the query is consequently yet another consideration. Officers may use the crime data mining in search for particular relationships, so their being able to request a specific output format is a valuable component of this application for law enforcement personnel.

The field of crime data mining is concerned with the efficient storage, access, modeling, and, ultimately, understanding of large crime data sets. A detailed discussion of these various aspects of data mining, both from a theoretical and from an implementation viewpoint, can be found in [1]. From the viewpoint of our approach, crime data mining can be specifically defined as the analysis of data in order to identify patterns or discover relationships among the various elements of a crime data set. The domain knowledge and experience provides help to analysis in guiding the search and can be represented in a variety of knowledge formats. From such a viewpoint, crime data mining: including classification, clustering, pattern analysis, discrimination, feature selection, and a broad array of related domains-looks as a critically important practice in business, science, and government.

Study show that productivity in data mining based on mathematical modeling is low: reality data is complex, and it usually takes a lot of trials to find a satisfactory mathematical description of the phenomenon under consideration. Due to this

complexity, modeling has to be done by specialists who are required to speak three “languages”: the language of mathematics in which the model is originally described, a programming language or an input-language to a standard package which is needed to solve the particular case, and the language of the user who is ignorant of these “internal representations”, presents his problem in “user-terms” and also needed the relevant features of the model depicted via e.g. graphic means. After all that, the model obtained can only be used for the particular situation and has to be adapted for a new application should relevant factors change. In most cases this means redoing the whole identification and estimation process.

From a decision maker’s point of view these costs associated with making a precise mathematical model require a high utility in terms of good decisions. However, the data sets are often not so good in practice. Thus cheaper methods like flexible queries in data-bases or simple, deterministic models implemented with spreadsheet programs are often preferred to more sophisticated ways of arriving at decisions: patricians judge a method not only according to its "precision" or "optimality", but also consider implementation and maintenance costs, reliability and transparency.

The methods used to solve these difficult data mining have slowly evolved from constructive methods, like uniformed search, human-computer interaction, to local search techniques and to population-based algorithms. Our research goal was to use blend intuition learning-based cooperative algorithms with methods dealing with uncertainty in order to induce extension rules from large data sets. The following is extension classification problems that are relevant in our proposed approach.

2.2 Extension Classification

We introduced the concept of data interval, through which we can describe any features data in data sets and tell whether it belongs to features interval, non belongs or border intervals. Meanwhile, the features data belonging to one interval can be put into different layers according to the dependent degree [4]. According to quantitative expressions, we can give the following definition:

Definition 2.1[5]. Let the features interval of data sets is $F_{ab} = \langle F_a, F_b \rangle$, non-features interval is $\bar{F}_{ab} = \langle \bar{F}_a, \bar{F}_b \rangle$, and dependent interval is C_{ab} , which fulfills :

- ① $C_{ab} = \text{Max } F_{ab} \cap \text{Min } \bar{F}_{ab} \neq \phi$;
- ② $\text{Max } F_{ab} = \bigcup_{i=1}^{\infty} \langle F_{ia}, F_{ib} \rangle$,
- ③ $\text{Min } \bar{F}_{ab} = \bigcap_{i=1}^{\infty} \langle \bar{F}_{ia}, \bar{F}_{ib} \rangle$.

Definition 2.2[5]. Let x_0 any point on an axis, $C_0 = \langle C_a, C_b \rangle$ is any dependent interval of data sets on an actual field, let

$$\rho(x_0, C_0) = \left| x_0 - \frac{C_a + C_b}{2} \right| - \frac{C_b - C_a}{2} \tag{1}$$

be the distance between x_0 and region C_0 , where $\langle C_a, C_b \rangle$ can be an open interval, a closed interval or a half closed interval. The distance between point and interval $\rho(x_0, C_0)$ has a relationship with $d(x_0, C_0)$ “the distance between point and interval” of traditional mathematics :

- ① When $x_0 \notin C_0$ or $x_0 = C_a, C_b$,
 $\rho(x_0, C_0) = \rho(x_0, C_0) \geq 0$;
- ② When $x_0 \in C_0$ and $x_0 \neq C_a, C_b$,
 $\rho(x_0, C_0) < 0, d(x_0, C_0) = 0$

The usage of the concept of distance can accurately depict the location relationship of points and intervals with a quantitative format. When the point is within the interval, traditional mathematics holds that all the distances between the point and the interval are zero; while in the non-optimum analysis, with the concept of distance, we can describe the different locations of a point in the interval according to the values of the distances. This allows the concept of quantitative description to develop from “points are all the same when they are within one interval” to “points can be differentiated even within the same interval”.

In the data mining analysis, except for the relationship of the locations of the features points and intervals representing aims and results, we often also has to consider the location relation between these intervals, points and intervals. Therefore, we have

Definition 2.3[5]. Let $C_0 = \langle C_a, C_b \rangle, C = \langle C_c, C_d \rangle$, and $C_0 \subset C$, then the distance between point x and interval C_0 and interval C is regulated as

$$D(x, C_0, C) = \begin{cases} \rho(x, C) - \rho(x, C_0) & x \notin C_0 \\ -1 & x \in C_0 \end{cases} \tag{2}$$

$D(x, C_0, C)$ describes the location relationship between point x and C_0, C . Based on the distance analysis, according to the extension theory we set up the expression of dependent degree:

$$C(x) = \frac{\rho(x, C_0)}{D(x, C_0, C)} \tag{3}$$

where $C_0 \subset C$, and there is no common point It can be used to calculate the dependent degree of points and intervals, whose value region is $(-\infty, +\infty)$. We use the above equation to express the dependent degree of data features, then we develop the qualitative description to quantitative description.



In the features selection and analysis, if a features classification subset $C_i (i = 1, \dots, L)$ was given in the objects set P , an attribute in P which does not belong to C_{ia} , must belong to C_{ib} in the classical mathematics. But attribute in C_{ia} consists of two kinds of attribute, between which there is the difference in innate character. In order to describe this relationship, we are establishing the concept of extension classification set:

Definition 2.4[5]. The so called extension classification set \aleph in the objects set P under a restraint is indicates to provide a resale number $C_{\aleph}(p) \in (-\infty, +\infty)$, for any $p_i \in P (i = 1, \dots, L)$, by which the relationship of p_i and \aleph is described. The mapping

$$C_{\aleph}(p_i) \rightarrow (-\infty, +\infty), p_i \rightarrow C_{\aleph}(p)$$

Is called the dependent function of \aleph .

(1) $C_{\aleph}(p_i) \geq 0$ shows $p_i \in \aleph$ (common trust optimum subset); and $\aleph = \{p_i : C_{\aleph}(p_i) \geq 0, p_i \in P\}$, is called the complete features field of \aleph (classical classification field).

(2) $-1 \leq C_{\aleph}(p_i) \leq 0$ shows $p_i \notin \aleph$, but under that restraint p_i can be turned into $\aleph^* \in \aleph$:

$$\aleph^* = \{p_i : -1 \leq C_{\aleph}(p_i) < 0, p_i \in P\}$$

Is called the extension features field of \aleph .

(3) $C_{\aleph}(p_i) \leq -1$ shows $p_i \notin \aleph$ and it also cant be turned $\aleph^{**} \in \aleph$ under that restraint; and

$$\aleph^{**} = \{p_i : -\infty \leq C_{\aleph}(p_i) < -1, p_i \in P\}$$

is called the non-features field of \aleph .

3 Crime Data Mining Pattern

3.1 Crime Data-Feasure Analysis

The general area of pattern discoveries based on data mining has attracted an enormous amount of attention in recent years. Although it has been variously defined, it is evident that data-feature exists at the enterprise level and is quite distinct from mere information. Also apparent in this area are the challenges that data analyzing poses to an organization. In addition to being difficult to data analysis, the Basic Data-feature traditionally has been stored on paper or in the minds of people. The Data-feature problems facing many firms stem from barriers to access and utilization resulting from the content and format of information. These problems make Data-feature acquisition and interpretation a complex and daunting process. Nevertheless,

data-feature technologies have been developed for a number of different applications, such as virtual enterprising, joint ventures, and aerospace engineering. The same problems of data-feature exist at the specialized organizations of policeman department. Many record management systems for crime control agencies contain a large amount of data for each case or incident, but although data may be available, they are not available in a form that makes them useful for higher level processing. For example, the ideal data-feature system should be able to provide information about problems that have not been identified previously, and thus be able to give innovative and creative support for new investigations. The conversion of information to crime pattern is an important concern for policeman department agencies. Data-feature is a product that is designed with a purpose in mind, while data serve as the ingredients in this product. Furthermore, addressing the conversion of crime pattern to useful and easily understandable crime pattern is a powerful aspect of knowledge data-feature that has thus far been missing from most law enforcement.

An important modification was the identification of certain fields for analysis and classification. This included both fields that can be used as search terms and fields that are returned by the extension classification of crime knowledge. The relationships between data-feature and non-feature are an important issue not focused upon in prior dependent function applications. For an investigator, relationships between crime features must not only be identified, but for crime analysis must also be explicitly understood. The format of the query is consequently yet another consideration. Officers may use the extension classification system in search for particular relationships, so their being able to request a specific output format is a valuable component of this application for law enforcement personnel.

3.2 Research Design

These objects contain information that is currently being stored in fields of the records management system at Dalian Policeman Department (DPD). A user can choose to enter up to four objects in any combination to begin a search. The different objects contain specific data-features. For example, the name object contains last name, first name, and middle initial. The majority of addresses use a street number, street name, street type, and apartment number. The crime type identifies the type of crime committed in accordance to the standard features classification system that is used in the Dalian. Vehicle information contains a number of possible fields, including the make, model, type, year, color, and license plate number. Users are able to search on any combination of these data-features and across different fields. By employing different search objects, officers are able to easily search by the specific type of information that he/she has available.

The purpose of this search tool is to discover relationships between the different search terms or objects. It is not only important to know that there is a relationship, but it is also important to know what the relationship between objects is. Figures 1 illustrate sample data-features using the extension classification, which provides a detailed description of how relationships can be identified, given that the officer has a limited amount of information.

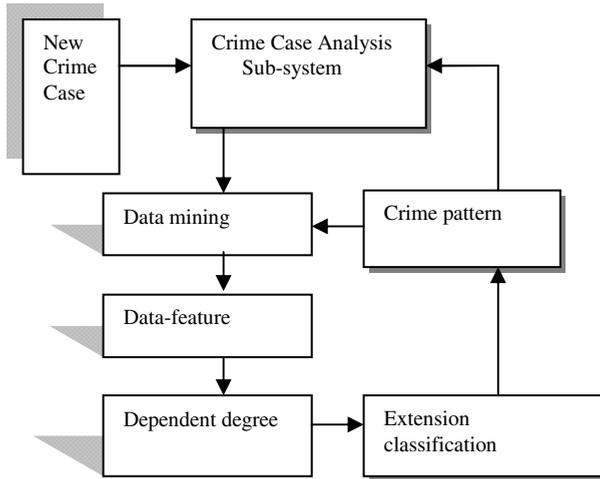


Fig. 1. Crime data-feature analysis system

A basic task for crime data mining at Dalian Policeman Department (DPD) is to create crime information from past data. In this case, information is made up of approximately 60 ten thousand criminal case reports, containing details from criminal events dating back to 2008 [10]. Tacit data-feature has also been described as the means through which new data-feature is generated as well as the practical pattern used to perform a task. It is tacit crime pattern that is used as investigators try to tie together information to solve cases and crimes. This ability to combine data-feature to create crime pattern information is often hampered by the amount of data-feature that exists [7-9].

4 Conclusion and Future Directions

From this new study, we conclude that the use of extension classification as a crime data mining and intelligence analysis tool in a law enforcement environment is quite promising. An important aspect of the study is that it dealt with real criminal data, real cases and search tasks and real crime analysts. In addition to providing an extension feature analysis system (EFAS), the system's combination of different data-features types in its associations provides much data-features of value in the analysis of crimes. Data-feature in a records information system is not useful if the system does not have the ability to pull together the different types of data-features and to present them in an understandable way. The EFAS uses data-feature and transforms them into intelligence that the officer can utilize. Criminals are creatures of habit and being able to understand their habits is an important issue. The EFAS takes advantage of this characteristic by capturing connections between people, places, events, and vehicles, based on past crimes. As a data mining tool, EFAS serves to create new knowledge in the form of links between people, places and objects, which in turn results in possible leads for investigation. From these findings we have been able to determine that EFAS was useful to our participants and, furthermore, provided them with a valuable asset in performance of investigative tasks. These preliminary results also indicate that EFAS

can potentially lead to increased productivity by reducing the amount of time spent for data-feature search. Finally, we have determined that additional development effort is required for redesign of the interface to enable users to be more readily able to interact with and understand the application. It is evident from this study that the use of data mining applications, such as EFAS, can have a significant impact on law enforcement.

Acknowledgment. This work was completed with the supported of the research foundation from Ministry of Public Security (Grant No. 2009YYCXLNST023). Author would like to thank professor He Ping for producing silhouettes with the developed method.

References

1. Adriaans, P., Zantinge, D.: Data Mining, pp. 5–6. Addison Wesley, New York (1996)
2. Han, J., Kamber, M.: Data Mining: Concepts and Techniques, p. 7. Morgan Kaufmann Publishers, New York (2001)
3. Wen, C.: Introduction of Extension. Systems Engineering Theory And Practice 18(1), 76–84 (1998)
4. Li, J., He, P.: Extended Automatic Reasoning of Criminal Investigation. In: International Conference on Industrial Mechatronics and Automation, pp. 345–349. IEEE Computer Society (2009)
5. He, P.: Method of system non-optimum analysis based on Intervenient optimum. In: Proc. of ICSSMSSD, pp. 661–670 (2007)
6. He, J., He, P.: A New Intelligence Analysis Method Based on Sub-optimum Learning Model. In: ETP International Conference on Future Computer and Communication, pp. 116–119. IEEE Computer Society (2009)
7. He, P.: Extension relationship mapping inversion and the automatic reasoning of crime detective. Journal of Pattern Recognition and Artificial Intelligence, 58–62 (2003) (Chinese)
8. He, P.: Fuzzy relationship mapping inversion and automatic reasoning of crime detective. In: Artificial Intelligence Applications and Innovations, pp. 681–690. Springer, Heidelberg (2005)
9. He, P.: Crime Knowledge Management Approach Based on Intuition Concept Space. In: Zhou, Q. (ed.) Proc. of the Int'l. Conf. on Intelligent Information Technology Application, pp. 276–279. IEEE Computer Society (2008)
10. He, P.: Crime Knowledge Management Based on Intuition Learning System. In: Ma, J. (ed.) Proc. of the Int'l Conf. on Fuzzy System and Management Discovery, pp. 555–559. IEEE Computer Society (2008)

A New Data Mining Approach with Data-Intuition Cooperative Analysis

Weidong Tao

Department of Information
Liaoning Police Academy
Dalian, 116036 China
lnpolice@126.com

Abstract. This paper describes how we have applied the design criteria of platform independence, stability, scalability, and an intuition graphical user interface to develop the Data-Intuition Cooperative Analysis (DICA). The Data-Intuition Cooperative is left behind without consideration. In this paper, a literature in data mining is to be reviewed and the technology aspect of data mining is to be analyzed. Also, general design principles are to be reviewed. According to all these issues, recommendations to designing an intuition-computer environment for trusted intuition-computer cooperative system are going to be analyzed and proposed.

Keywords: Data mining, Data-intuition, DIL model, cooperation analysis system.

1 Introduction

The field of data mining is concerned with the efficient storage, access, modeling, and, ultimately, understanding of large data sets. A detailed discussion of these various aspects of data mining, both from a theoretical and from an implementation viewpoint, can be found in [1-4]. From the viewpoint of our approach, data mining can be specifically defined as the data analysis in order to identify patterns or discover relationships among the various elements of a data set. From such a viewpoint, data mining: including classification, clustering, pattern analysis, discrimination, feature selection, and a broad array of related domains. But, conventional environment of data mining is “man round computer”. It is difficult to deal with the information of being of brain thinking procedure. It make person have definite dependence that computer bring into play great power while dealing with information. Whether, when the result from computer is different from reality significantly, person will doubt it. It is the reason that occur incompatible question when we applicant data mining ways on solving the conducting of complexity problem.

Classical methods of data mining are mathematical analysis, the catalog of mathematical methods used for data mining applications encompasses neural networks, the theory of fuzzy sets (Zadeh, 1965), the theory of rough sets (Pawlak, 1991), losupport vector machines, Bayesian analysis, and Markov blankets, among others (see [1-8] for useful descriptions and tutorials).

Study show that productivity in data mining based on mathematical modeling is low: reality data is complex, and it usually takes a lot of trials to find a satisfactory mathematical description of the phenomenon under consideration. Due to this complexity, modeling has to be done by specialists who are required to speak three “languages”: the language of mathematics in which the model is originally described, a programming language or an input-language to a standard package which is needed to solve the particular case, and the language of the user who is ignorant of these “internal representations”, presents his problem in “user-terms” and also needed the relevant features of the model depicted via e.g. graphic means. After all that, the model obtained can only be used for the particular situation and has to be adapted for a new application should relevant factors change. In most cases this means redoing the whole identification and estimation process.

The objective of this paper is to analyze intuition-computer cooperation system and to discuss a method for designing data-intuition analysis based on the theory of intuition learning.

The paper is structured as follows: He Ping’s Data-intuition analysis theory is summarized in sections 2. Section 3 contains an adaptation of what is known about the process of learning from autonomous data to intuitive model. Sections 4 contain the author’s main ideas on how to characterize and design data-intuition cooperative system and more specifically the relationship among data and intuition. Finally, section 5 contains the main conclusions.

2 Data-Intuition Analysis

2.1 Intuition Recognition

Intuition is part of that “knowing” referred to by Polanyi (1966) when he declared that “we know more than we can tell” [7]. An intuition is a recognition or judgment that is: (a) arrived at rapidly, without deliberative rational thought; (b) difficult to articulate verbally; (c) based on a broad constellation of prior learning and past experiences; (d) accompanied by a feeling of confidence or certitude; (e) affectively-charged [Sinclair, M. and Ashkanas, 2005]; (f) experience mapping and inversion [He Ping, 2008].

In fact, we acquire our capability to intuit through experiences in particular domains (intuition is not a generic ability; it is domain specific) via explicit and implicit learning processes [Reber, A.S. 1993] , which result in the acquisition of highly complex and subtle patterns of tacit knowledge that cannot be described or explained easily. These mental models are stored in long term memory under a variety of sophisticated rules (often too complex for verbal exposition) for how to achieve specific goals in particular situations and which are activated by the cues that we perceive from the environment [Klein. G.1998]. Intuitions enable us to solve problems, take decisions, achieve insights and generate scientific discoveries and artistic creations.

Intuition is synonymous in many languages and cultures with visceral signals often referred to as “gut feelings” Neurological research has identified an awareness that operates below our level of consciousness and which may serve like a physical “alarm

bell". Other research by Jung-Beeman and his colleagues (2004) using neuro-imaging techniques has identified brain regions that are implicated in those insights (the "aha" or eureka moments) where we experience the pieces of a problem that has been perplexing us falling into place often after a period of unconscious "incubation". Alongside this, the methods of neuroscience are also beginning to shed light upon the brain regions and processes that are involved in intuitive judgments. For example, Le Doux (1996) has discussed the role of working memory in feelings [8], emotion, unconscious processing and human consciousness itself. Lieberman and his colleagues (1999) found that high-experience domain judgments (i.e. high familiarity) produced activation in a network of neural structures involved in automatic social cognition (referred to as the X-system, i.e. reflexive) [9]. Low-experience domain judgments (i.e. low familiarity) produced activations in a network involved in effortful social cognition and propositional thought (the C-system, i.e. reflective). It seems as though the physical locations of insight, "gut feel" and intuition in the body and brain are being uncovered. For researchers, there are important questions about intuition learning in data mining, psychology, neuroscience, philosophy and linguistics, and computer science.

2.2 Data-Intuition Analysis

From a decision maker's point of view these costs associated with making a precise mathematical model require a high utility in terms of good decisions. However, the data sets are often not so good in practice. Thus cheaper methods like flexible queries in data-bases or simple, deterministic models implemented with spreadsheet programs are often preferred to more sophisticated ways of arriving at decisions: practitioners judge a method not only according to its "precision" or "optimality", but also consider implementation and maintenance costs, reliability and transparency.

The methods used to solve these difficult data mining have slowly evolved from constructive methods, like uniformed search, human-computer interactive, to local search techniques and to population-based algorithms. Our research goal was to use blend intuition learning-based cooperative algorithms with methods dealing with uncertainty in order to induce extension rules from large data sets.

The general area of pattern discoveries based on data mining has attracted an enormous amount of attention in recent years. Although it has been variously defined, it is evident that data-pattern exists at the enterprise level and is quite distinct from mere information. Also apparent in this area are the challenges that data analyzing poses to an organization. In addition to being difficult to data analysis, the Basic Data-intuition (BDI) traditionally has been stored on paper or in the minds of people (He Ping, 2009). The BDI problems facing many firms stem from barriers to access and utilization resulting from the content and format of information. These problems make BDI acquisition and interpretation a complex and daunting process. Nevertheless, BDI technologies have been developed for a number of different applications, such as virtual enterprising, joint ventures, and aerospace engineering. For example, the same problems of BDI exist at the specialized organizations of policeman department [6]. Many record management systems for crime control

agencies contain a large amount of data for each case or incident, but although data may be available, they are not available in a form that makes them useful for higher level processing. For example, the ideal BDI system should be able to provide information about problems that have not been identified previously, and thus be able to give innovative and creative support for new investigations. The conversion of information to crime pattern is an important concern for policeman department agencies. BDI is a product that is designed with a purpose in mind, while data serve as the ingredients in this product. Furthermore, addressing the conversion of crime pattern to useful and easily understandable crime pattern is a powerful aspect of knowledge BDP that has thus far been missing from most law enforcement.

3 Data-Intuition Learning

3.1 Basic Concept

The word “Data-intuition learning” (DIL) means to discovery a trust intuition from data mining. Specifically, DIL is intelligent techniques which seek good (near-optimal) solutions at uncertainty decision making.

According to [Ref. 3-8] intuition learning theory, DIL is the process whereby trust data mining is created through the cooperation of the data analysis and the intuition. In this subsection we review the method of DIL. Many literatures have studied the process of data mining, and much of what we know today comes from their theories. Perhaps the most basic conclusion from these early studies is that people do learn from their experiences particularly from their mistakes of data analysis. One of the most fundamental requirements that facilitate learning is an appropriate pattern where learners can have experiences of intuition decision. Intuition learning emphasizes the role that appropriate pattern of intuition and experiences play in the data analysis. In DIL the learner is directly in touch with the realities being studied. It involves direct encounter with the phenomena being studied rather than merely thinking about the encounter or studying the experience of others with such phenomena. The learning process of data-intuition is not identical for all human beings, and people enter learning situations with a preferred learning style. Associated with this learning style there is a theory about how people learn, or more specifically, about how they themselves learn best. Learning pattern that operate according to a learning method that is dissimilar to a person’s preferred style of learning are likely to be rejected or resisted by that person [8]. Thus an understanding of learning method is important for learning environments based on the experience and the knowledge.

At present, most DIL [6] is designed manually based on past experience of their intuition and data analysis. Since the number of possible DIL is very large for realistic applications of reasonable complexity, learning system designed manually may not work well when applied in new problem instances. Further, there is no systematic method to evaluate the effectiveness of DIL designed manually. For these reasons, an automated method for discovering the proper DIL for a particular application is very desirable. This leads to the development of our system for automated learning of intuition.

3.2 DIL Model

There are stages and spaces to the understand of data-intuition pattern, there be a set of data attributes $D(x)$ of data-intuition pattern, and a set of intuition features $I(x)$ in any data mining process,. We have

Definition 1: Let $DI(x) = DI\{D(x), I(x)\}$ to be input vector $x = \{x_1, x_2, \dots, x_n\}^T$, where $D(x) = \{D_1(x), \dots, D_p(x)\}$

be a set of data attributes , and let $I(x) = \{I_1(x), \dots, I_p(x)\}$ be a set of intuition features, then must have $DI(x)$ from an experience pattern of $D(x)$ and $I(x)$

$$DI(x) = \{DI_1(x), DI_2(x), \dots, DI_p(x)\} .$$

Definition 2: Let $\gamma_i(x)$ be a trust degree on data-intuition feature $DI_i(x)$ and $\gamma_i(x) = \alpha_i(x) \cap \beta_i(x)$, where $\alpha_i(x)$ be a trust degree of $D_i(x)$, and $\beta_i(x)$ be a trust degree of $I_i(x)$, $i = 1, \dots, p$.

Definition 3: Let $\gamma(DI)$ be a trust degree of data-intuition analysis, we then have a trust data analysis, that is ,

$$\gamma(DI) = \frac{1}{p} \sum_{i=1}^p (\alpha_i(x) - \beta_i(x))$$

Without loss of generality we can restrict $\gamma(DI)$ within the interval $[0, 1]$,

4 Design of DIL Cooperative System

DIL system is an intelligent data analysis system for helping viewers through the experience of the data analysis and intuitive decision, and for allowing them to “self-dose” on help, while maintaining their suspension of disbelief and dramatic immersion. There are several different levels of DIL. The first is an across the boards reconfiguring of the experience. Playing with Cooperative Data-Intuition Learning Systems (CDILSs) turned off results in a more game-like experience, while playing with it on causes the title to change subtly, making things easier and more obvious. The next level of DIL is the state of the Artificial Data-Intuition networks. When it is dimmed, it indicates that a location or character holds no further relevance for the viewer. If brighter, it is signifying that locations and characters still need further exploration or interaction.

The designers of CDILSs, then, are training data-intuition levels of data mining, up to this point without a set task, without a programmed. It is purely and simply training the unrestricted, taken-for-granted yet highly personal nature of sketches. Using the three-dimensional programming of the computer, lines and spaces can be brought into

view and prepared for experiencing by the senses, in a way unavailable to any other instrument. Computers, then, are essential to this task.

We have shown most clearly the desire to be able to sketch even more directly in Intuitive Experience Space (IES). It should be possible to translate manual three-dimensional movements into IES at a single step. There are programs for doing this, but they are then out of the question due to the costs involved and the availability. The capacity for visual representation offered by the computer shows great promise. So much so that it makes you long for more. The experimental nature of intuitive concepts enables their executants to then put them to the test. A scientific analysis of the intuitively designed concept, in the shape of an AIN, gives those with decision-making powers a tool with which to make a decision. Intuition is an inductive process, analysis a deductive process. So the participants begin by communicating with themselves by way of the computer.

The subsequent phases of the CDILSs are intended as training for communication with others. The computer is the pre-eminent tool for enabling communication between the most wide-ranging disciplines, in that it couples image and language. Architectural AIN are dictated by a great many parameters. One of these is the intuition of those involved, the strictly personal premises of, say, the architect, or the IES of the expert in his own field.

The CDILSs starting-points are brought together in a Data-Intuition Concept Sets (DICS). This DICS is an experimental hypothesis which can then be tested as to its validity and suitability. In the DICS the will of the designers is illustrated and defined. We ask each participant of the CDILSs to look at learner personal intuitive sketch, and be inspired by it to arrive at a DICS. To give a depiction of what the sketch might mean. What was originally a value-free sketch is now attributed with a meaning.

Still proceeding from the initial sketch the AIN participants now take a number of elements in the sketch and rearrange them, sometimes drawing them again. Learner adds new elements so as to compile the structure they wanted. As yet gravity is conspicuous by its absence. Working like astronauts in CDILSs, they examine the construction from all sides. It is as easy to add something to the bottom as to the top. This exercise is just as valuable for real-life construction. After all, the construction material proceeds from the same obliviousness to gravity as the image envisioned by the designer. The effect of gravity on a steel beam is only calculated when that beam is in place. Before then the beam can only be described in terms of moments of inertia, moments of resistance. The properties of the material are not connected to the situation on site. They are, however, designed to deliver a particular performance in a particular situation. So both material and image are developed without giving consideration to gravity. It is of great importance, I feel, to recognize this fact, so that the potentials offered by material and construction can be better exploited in real life. The longer you spend working in weightless digital CDILSs, the greater the tendency will be to challenge gravity. You will then try to see how far you can go in making a gravity-free image.

In founding the CDILSs, the focus is finding the proper methodology in realizing the learn aims. Then it may embody the intellectual aspect in the learning system. By analysis, we find that Interactive Mapping Intuition Pattern inversion (IMIPI) (He

Ping, 2008) is a useful tool in constructing this brain sense. Here firstly we give a supposition. In brain thought practice, by analyzing the attribute of the brain thought, and can construct a sensation model. Thus, we may found a similar image model of the initial sensation. Normally, the detector could not see the crime procedure on spot. After crime happened, people will never experience the scene again. Consequently, only by mocking and recognizing could people realize and grasp its changing regular patterns.

5 Conclusions

We found that Data-intuition Cooperative Analysis System (DICAS) of data mining and knowledge discovery is effective tools that they build intelligence activities [5]. This paper is building on data mining intelligence system, through concrete case's emulate resolution, machine learning action of human being's thinking procedure.

We propose an approach that incorporates an extension data mining module to identify relevant system inputs and to analyze the way these inputs interact within the system. This extension data mining model uses traditional methods to analyzed relation between data and intuition. Its principal contribution makes use of a cooperative data mining procedure that is activated at certain intervals during the sub-optimum process in order to make use of information obtained during that process, with the goal of speeding the search for trust data analysis. However, the approach has not yet been applied by others to the areas of data analysis research, finance, and other business settings, which we plan to explore in this paper.

Acknowledgment. This work was completed with the supported of the research foundation from Ministry of Public Security (Grant No. 2009YYCXLNST023). Author would like to thank professor He Ping for producing silhouettes with the developed method.

References

1. Prakken, H., Sartor, G.: A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law* 4, 331–368 (1996)
2. Prakken, H.: Modelling Reasoning about evidence in legal procedure. In: *Proceedings of the 8th International Conference on AI and Law*, pp. 119–128 (2001)
3. He, P.: Design of Interactive Learning System Based on Intuition Concept Space. *Journal of Computer* 5(3), 478–487 (2010)
4. He, P.: Crime Knowledge Management Based on Intuitionistic Learning System. In: *FSKD*, pp. 467–471. IEEE (2008)
5. He, P.: Crime Knowledge Management Based on Intuition Concept Space. In: *Intelligent Information Technology Application*, pp. 276–279. IEEE Computer Society (2008)
6. He, P.: Crime Pattern Discovery and Fuzzy Information Analysis Based on Optimal Intuition Decision-Making. *AISC*, vol. 54, pp. 426–439. Springer, Heidelberg (2008)

7. He, J., He, P.: Fuzzy Relationship Mapping and Intuition Inversion: A Computer Intuition Inference Model. In: IEEE 2008 International Conference on MultiMedia and Information Technology, pp. 298–301 (2008)
8. He, P.: Crime Pattern Discovery and Fuzzy Information Analysis Based on Optimal Intuition Decision Making. AISC, vol. 54(1), pp. 426–439. Springer, Heidelberg (2008)
9. Tao, W., He, P.: Intuitive Learning and Artificial Intuition Networks. In: 2009 ETT, pp. 347–360. IEEE Computer Society (2009)
10. Tao, W., He, P.: Measurement of Network Security Based on Sub-optimum Degree. In: ICTM 2009, pp. 283–286 (2009)

A Discovery for the Association Rules of Operation Sequence Based on Improved Apriori Algorithm

Tan Yanhua

College of Physics & Electronic Information Engineering of Wenzhou University, Wenzhou,
Zhejiang, China
jsj_tyh@wzu.edu.cn

Abstract. With the application of the computer aided process design technology, a great deal of process data has been accumulated by enterprises. Inherent linking with these data is found by exploiting Apriori algorithm, which will help enterprises with the process design. In this paper, the association rule model of operation sequence is introduced briefly at first. Then the bottleneck problem of apriori algorithm is analysed, on basis of this analysis, an improved apriori algorithm is provided. At last, combining with actual data, a detailed description of the improved Apriori algorithm's application in the discovery for association rules of operation sequence is put forward.

Keywords: Apriori algorithm, operation sequence, association rule.

1 Introduction

Data mining is a semi-automatic process to extract some potential and useful information and knowledge from much incomplete, noisy, fuzzy, random data. At present, with the growing popularity of computers, computer Aided Process designing(CAPP) has been widely used in the enterprises. In the enterprises, a large number of process data are accumulated, a variety of knowledge implied in these data is useful, but it is not easy to obtain these processes knowledge. At first, the process designing itself is a complex process, considering to analyse and process a large number of process information. It is not only to consider the structure, shape, material and heat treatment and other factors of accessory diagram, but also to understand the processing methods and manufacturing conditions and other complex information and knowledge factors. What's more, because of huge amount of data, it is not only time-consuming, laborious and reliability to take artificial way, but also has no good maneuverability. Therefore, the introduction of association rule mining technique with respect to process sequence for mining association rules has become a meaningful work.

2 Association Rule Model of Operation Sequence

When enterprises do new process designing, the introduction of new data mining technology can mine the association rules between process sequence from these data.

Definition 1: Let $I = \{i_1, i_2, \dots, i_m\}$ be the set consisting of m different process sequence items, given a process sequence transaction database D , of which each process sequence transaction T is a set of process sequence items I , that is $T \subseteq I$. Suppose T has a unique identifier TID. Process item set $I = \{i_1, i_2, \dots, i_m\}$ is the set of all the data in T , that I is the set of appeared process sequence in all the process routing, all the elements i_j ($1 \leq j \leq m$) in I denote a process.

Definition 2: Let A be a set of process items, process transaction T contains A if and only if $A \subseteq T$. Association rule such as the implication style $A \Rightarrow B$, where $A \subseteq I$, $B \subseteq I$ and $A \cap B = \Phi$. The rule $A \Rightarrow B$ is established in the transaction set D with the support degree s , where s is the percentage of transaction D contains $A \cup B$, which is the conditional probability $P(A \cup B)$. The rule $A \Rightarrow B$ has a confidence degree c in the process transaction database D , if the percentage of transaction containing A in D is also contains B is c , that is the conditional probability $P(A | B)$. The support degree ($A \Rightarrow B$), that is the percentage of containing $A \cup B$ in T , the support degree is a measure of the importance of association rules, support degree = support count / total number of transactions, so you can use support count to measure support degree, Confidence ($A \Rightarrow B$) = support($A \Rightarrow B$) / support(A). That is the percentage of the transaction containing A also containing B in T , the confidence degree is a measure for the accuracy of association rules.

Definition 3: The process sequence of operation sequence association rules meets the minimum support threshold (min_sup) and minimum confidence threshold (min_conf) often called a strong association rule.

According to the property of Apriori algorithm, obviously, the association rules model of processes sequence has the following properties:

Property 1: All the non-empty sub-sequences of frequent operation sequences are frequent.

Property 2: The superset of non-frequent process sequence is non-frequent for sure.

3 The Improved Apriori Algorithm

The Apriori algorithm exploits an iterative method of searching k -itemsets layer by layer, searching $(k + 1)$ - itemsets based on k -itemsets. Because facing large-scale databases in the process of association rule mining, the reasons results in the low efficiency of Apriori algorithm are main two aspects: First, the number of generated candidate items is huge; Second, the number of records in the database is many, causing excessive expense of I / O. In this paper, a method of optimizing Apriori algorithm in two aspects is provided.

3.1 Optimize the Pruning Operation of Candidate Itemsets

Deduction 1: T_k is a k -dimensional data items-set, if all the $k-1$ -dimensional frequent item-set set L_{k-1} including the number of $k-1$ dimensional subset of T_k is less than k , then T_k is not a k -dimensional frequent item set.

Proof: the number of $k-1$ dimensional subset of data items set T_k is k , if the frequent itemset L_{k-1} including the number of $k-1$ dimensional subset of T_k is less than k , then there exists a $k-1$ dimensional subse of T_k is not frequent itemset, according to Property 1 we can know that T_k is not K -dimensional frequent item set.

3.2 Reduce the Expense of I/O

Deduction 2: T is a transaction record in database D , if the number of valid data m in T is less than the dimension k of frequent itemsets L_k , then it will not search any element X of frequent itemsets L_k in T .

Proof: In common sense, when the number of valid data $m = 2$ in the transaction record T , while the dimension of complex itemset L_3 is 3. Therefore, an arbitrary element X in L_3 has at least three items, then in any case it can not be found the set element X including three items in the record that has only two valid data.

Therefore, according to Deduction 2 we know that any subset of $k + 1$ dimensional frequent itemsets L_{k+1} can not be found in this transaction, therefore, this transaction record can be deleted.

3.3 The Description of Improved Algorithm

Input: D : Database of transactions; min_sup : minimum support threshold

Output: L : frequent itemsets in D

```
(1)  $L_1 = \text{find\_frequent\_1-itemsets}(D)$ ;
(2) For( $k=2$  ;  $L_{k-1} \neq \Phi$  ;  $k++$ ) {
(3)  $C_k = \text{apriori\_gen}(L_{k-1}, \text{min\_sup})$  ;
(4) for each transaction  $t \in D$  {
(5)    $C_t = \text{subset}(C_k, t)$  ;
(6)   for each candidate  $c \in C_t$ 
(7)      $c.\text{count}++$  ;
(8) }
(9)  $L_k = \{ c \in C_k \mid c.\text{count} \geq \text{min\_sup} \}$  ;
(10) if( $k \geq 2$ ) {
(11) delete_datavalue( $D, L_k, L_{k-1}$ ) ;
(12) delete_datarecord ( $D, L_k$ ) ; }
(13) }
(14) return  $L = \bigcup_k L_k$  ;
procedure apriori_gen( $L_{k-1}$ : frequent( $k-1$ )-itemsets)
(1) for each itemset  $l_1 \in L_{k-1}$  {
(2) for each itemset  $l_2 \in L_{k-1}$  {
(3) if( $(l_1[1] = l_2[1]) \wedge (l_1[2] = l_2[2]) \wedge \dots \wedge (l_1[k-2] = l_2[k-2]) \wedge (l_1[k-1] < l_2[k-1])$ )
then {
(4)    $c = l_1 \bowtie l_2$  ;
(5) for each itemset  $l_1 \in L_{k-1}$  {
```

- (6) for each candidate $c \in C_k$ {
- (7) if I_1 is the subset of c then
- (8) $c.num++$; } } } }
- (9) $C'_k = \{ c \in C_k \mid c.num = k \}$;
- (10) return C'_k ;

Procedure delete_datavalue (D:Database ; L_k :frequent(k)-itemsets ; L_{k-1} :frequent(k-1)-itemsets)

- (1) for each itemset $i \in L_{k-1}$ and $i \notin L_k$ {
- (2) for each transaction $t \in D$ {
- (3) for each datavalue $\in t$ {
- (4) if (datavalue= i)
- (5) update datavalue=0 ;
- (6) } }

Procedure delete_datarecord (D: Database ; L_k :frequent(k)-itemsets)

- (1) for each transaction $t \in D$ {
- (2) for each datavalue $\in t$ {
- (3) if (datavalue!=null and datavalue!=0) {
- (4) $datarecord.count++$; }
- (5) if (datarecord.count<K) {
- (6) delete datarecord ; }
- (7) }

4 Example for Finding Association Rule of Operation Sequence

In this paper, in order to illustrate the validity of the algorithm, we refer to the data of paper [5], and revise the tenth record. Each transaction record represents a part of processing lines, 0 denotes no operation sequence.

Table 1. Handicraft routing data

| TID | The name of accessory | Process 1 | Process 2 | Process 3 | Process 4 | Process 5 | Process 6 | Process 7 | Process 8 |
|------|--------------------------|-----------|-----------|----------------|-------------|------------|--------------|--------------|------------|
| T001 | Isolation circle A | material | vehicle | mill | plier | vehicle | plier | inspection | 0 |
| T002 | rotor | material | vehicle | plier | vehicle | mill | plier | inspection | sand |
| T003 | Weight ring | material | vehicle | mill | plier | inspection | polarization | 0 | 0 |
| T004 | Lower coupling sleeve | material | vehicle | prescription | vehicle | plier | vehicle | plier | inspection |
| T005 | Parting block | material | vehicle | mill | plier | vehicle | inspection | polarization | 0 |
| T006 | Installing screw | material | vehicle | mill | plier | inspection | passivation | 0 | 0 |
| T007 | Gear bowl | material | vehicle | vehicle | plier | vehicle | inspection | polarization | 0 |
| T008 | Connection bolt | material | solid | vehicle | grind | plier | inspection | prescription | 0 |
| T009 | circular | material | vehicle | heat treatment | vehicle | mill | plier | inspection | 0 |
| T010 | Half-finished goods bolt | grind | plier | inspection | passivation | 0 | 0 | 0 | 0 |



Table 2. Process mark

| material | vehicle | mill | plier | grind | polarization | sand | prescription | inspection | Heat treatment | solid | passivation |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|-----------------|-----------------|
| I ₁ | I ₂ | I ₃ | I ₄ | I ₅ | I ₆ | I ₇ | I ₈ | I ₉ | I ₁₀ | I ₁₁ | I ₁₂ |

Let the minimum transaction support count $\min_sup=5$ and minimum confidence degree be 60%, the main steps are shown as follows:

Step 1: In the first iteration, each process is the candidate member of 1-itemset collection C_1 , thus by scanning all the processes of transaction database D to calculate the support count of each item.

Step 2: According to $\min_sup=5$, then determine the set of frequent 1 - item collections L_1 .

Step 3: The minimum transaction support count of $I_5, I_6, I_7, I_8, I_{10}, I_{11}, I_{12}$ process sequence does not meet the $\min_sup = 5$,

which not appear in frequent 1 - itemsets in set L_1 , and by Property 2 the supersets include non-frequent process sequence are not be frequent, you can remove these processes from the D , then obtain D_1 .

Step 4: To find L_2 , the algorithm generates a candidate 2 - itemsets C_2 by using $L_1 \bowtie L_1$.

Step 5: Continue to scan the affairs in D_1 , and calculate the supporting count of each candidate set in C_2 .

Step 6: Determine the set L_2 of frequent 2 - item, which consists of candidate 2-itemset satisfy the minimum support count .

Step 7: After generating L_2 , you can find some affairs records such as T_{10} in D_1 , the number of effective data is 2,

then by Deduction 2 we can see the searching process of collecting frequent 3-itemset L_3 , there will be no

elements in C_3 appearing in this record, which can be deleted directly, then obtain D_2 .

Step 8: To find L_3 , the algorithm generates a candidate 3 - itemsets C_3 by using $L_2 \bowtie L_2$.

There are more than one elements C_3 produced by connecting, then according to Property 1 to prune: all non-empty process sequences of frequent operation sequences are to be frequent. All 2 -item subset of $\{I_2, I_3\}$, $\{I_2, I_4\}$ and $\{I_3, I_4\}$ except for $\{I_2, I_3, I_4\}$ are the elements of L_2 , which can be verified. The other elements C_3 that have 2 - item subset does not belong to L_2 , so they should be removed from C_3 .

Step 9: Continue to scan the affairs in D_2 , calculate the supporting count of each candidate set in C_3 . L_3 is generated by C_3 .

Step 10: Because there is only 3 - itemsets in L_3 , thus $C_4 = \Phi$, the algorithm terminates, that finding all the frequent itemsets.

Step 11: Calculate the confidence degree.

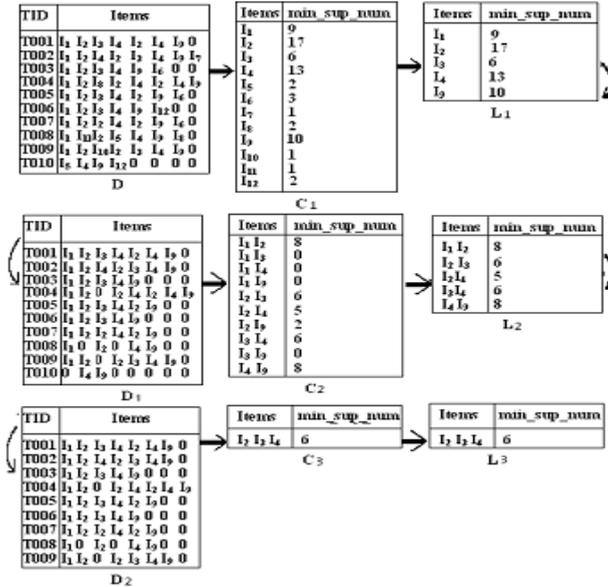


Fig. 1. The example for extracting operation sequence association rule based on improved apriori algorithm

For frequent_2-itemsets in L₂, calculating confidence degree is shown as follows:

- $I_1 \Rightarrow I_2$, confidence=8/9=88.8% ;
- $I_2 \Rightarrow I_3$, confidence=6/17=35.3% ;
- $I_2 \Rightarrow I_4$, confidence=5/17=29.4% ;
- $I_3 \Rightarrow I_4$, confidence=6/6=100% ;
- $I_4 \Rightarrow I_9$, confidence=8/13=61.5% ;

For frequent_3-itemsets in L₃{ I₂, I₃, I₄}, calculating confidence degree is shown as follows:

- $I_2 \Rightarrow I_3 \wedge I_4$, confidence=6/17=35.3% ;
- $I_2 \wedge I_3 \Rightarrow I_4$, confidence=6/6=100% ;

Therefore, according to the minimum confidence degree is 60% , $I_1 \Rightarrow I_2$ (material=>vehicle), $I_3 \Rightarrow I_4$ (mill=>plier),

$I_4 \Rightarrow I_9$ (plier=>inspection) , $I_2 \wedge I_3 \Rightarrow I_4$ (vehicle∧mill=>plier) are stronger type rule.

5 Conclusion

By making use of Apriori algorithm in data mining to find some interesting knowledge in the daily process data in these enterprises, which will help process designing in business. What's more, because there exists the performance bottleneck problem in the massive data processing of the typical Apriori algorithm, it needs to try to use a variety

of ways to optimize. In this paper, the improved method of the proposed algorithm not only reduce the pruning operation of candidate itemsets C_k , but also reduce the expense of I/O by deleting the data in a database. Finally the performance of Apriori algorithm is optimized.

References

1. Han, J.-W., Kamber, M.: Data Mining Concepts and Techniques. Beijing Higher Education Press (2001)
2. Li, X.-C., Wang, B.-B.: An Improved Apriori Algorithm for Mining Association Rules. *Computer Engineering* 28(7), 104–105 (2002)
3. Zhang, C.-S., Li, Z.-Y.: An Improved Algorithm for Apriori. In: International Symposium on Education and Computer Science (ECS 2009), pp. 995–998 (March 2009)
4. Zhang, C.-S.: A Modified Apriori Algorithm with Its Application in Instituting Cross-Selling Strategies of the Retail Industry. In: 2009 International Conference on Electronic Commerce and Business Intelligence (ECBI 2009), pp. 515–518 (June 2009)
5. Liu, S.-N., Tian, X.-T., Zhang, Z.-M., Xu, J.-X., Zhu, M.-Q.: Typical operation sequence capture based on Apriori algorithm. *Computer Integrated Manufacturing Systems* 12(8), 1279–1282 (2006)
6. Tang, Y., Wang, Y., Yang, H.: Optimized Method for Mining Maximum Frequent Itemsets. *Computer Engineering And Application* 42(31), 171–173 (2006)

A Parameter-Driving Method of Redeveloping Mould Standard-Part Library Based on SolidWorks

Tan Yanhua

College of Physics & Electronic Information Engineering of Wenzhou University,
Wenzhou, Zhejiang, China
jsj_tyh@wzu.edu.cn

Abstract. In this paper, on the basis of the 3D feature modeling system SolidWorks, the principle of redevelopment by making use of parameter-driving approach is introduced. And a practical example is designed to illustrate the approach of using dimension parameter-driving to develop mould standard-part library.

Keywords: Standard-part library, parameter-driving, solidworks, visual c++.

1 Introduction

Making use of parametric feature-based modeling technique, the applicability of CAD systems is improved greatly, the function of CAD systems is enhanced, and the design level of the engineering designer staff is enriched. At the same time, for secondary CAD developers, the parametric feature-based modeling technique also has brought great convenience, which improves the developer's productivity greatly. Just as the same as many and many commercial CAD / CAM systems such as PRO / E, UG, AUTOCAD, etc. the SolidWorks are designed as a general mechanical design and manufacture software, not especial for mould. Therefore, in order to improve the efficiency and accuracy of designing mould, it needs redevelopment. While the developing of three-dimensional mould standard-part library is the groundwork of redeveloping SolidWorks.

2 Theoretical Basis of Parameter-Driving Method

2.1 Parametric Principle

Parametric designing(also called Dimension-Driven) is a subject raised by the practical application of CAD technique. It not only makes CAD system have an interactive graphics function, but also have an automatic mapping function. At present, parametric designing is an important but further research subject in the field of CAD technique. The developed special product design system by making use of parametric designing that can release designers from drawing a large number of heavy and trivial work, which improve the speed of designing greatly and reduce amount of information storage.

While Solidworks is a software based on parametric feature-based modeling technique. For its redevelopment, it should follow the principle of parametric CAD. For the products of generalization and high standardization in the enterprises, the mathematical model and the product structure adopted by product designing are fixed, only the differences from structure size. Those differences are caused by the same number and known conditions of different values in the product designing with different specifications. For these products, you can replace the known conditions and other basic parameters by the corresponding variables. And then based on these known conditions and basic parameters, the program can automatically query the database, or calculate all the necessary drawing data by the program and at last design the graphics automatically on the screen.

2.2 Parametric Designing Idea

Parametric designing is the parameter-driven based on the drive mechanism, the parameter-driven mechanism is the operation based on graphics data. Through parameter-driven mechanism, the geometric data in graphics can be parametric modify. However, we must satisfy the constraints of graphics in the modifying process, and also need the driven means to restrain the linkage, but the linkage is through relationship between binding to implement driven. For a graph, the constraints may be very complex and significant. In fact, it can be controlled by user, that is the number of independent parameters is only a few, that is called the main argument or the main constraints; other constraints determined by the graph structure or have relationship with the main constraints are referred as secondary constraints. The main constraints are not simplified, the simplification of secondary constraint has two ways: graphical features linkage and related parameters linkage.

Establishing the relationship between secondary constraints and main constraint on the numerical and logical is called related parameter linkage. In the parameter-driven process, it always keeps this relationship unchanged. The linkage approach of related parameter establishes the relationship for judging from some driving points to driven points not by topology. Using this method, a driven tree is introduced, the tree is shown the constraint relations between active points and driven points, so as to judge the driven and constrainting situations of the graphics visually.

The so-called graphic characteristics linkage is to ensure the graph topology relationship unchanged. Reflected in the parameter-driven process, according to various related criterias to judge the above entities and geometry data, the new geometry data is obtained. We call these geometric data as driven points. In this way, the constraints on the driven point has been linked with the drive parameters. Relying on this link, the driving of the driven point has been obtained. At the same time, the driven mechanism expands its scope.

2.3 Parameter-Driving Method

The parameter-driven method is to keep accessory structure unchanged, see the size of the accessory as a variable, giving different values of size, so as to obtain a series of similar accessories with the same structure but different sizes. It allows users to define a typical accessory, but not to consider the exact position of geometric elements in the

accessories, only to ensure its topology correct. Through the changes in the values of parameters to generate accessory family that of the same structure but different parameters, which shows great convenience with the aspects of re-generating and modifying.

The basic principle of parameter-driving method is to take the combination way of 3D model and programme control. The original feature model is not created by the program, but use interactive way to generate. On the basis of original created feature model, according to the design requirements to establish a set of full control on the original features models and designing parameters. The parametric programme programs designing parameter of this feature model, to design the retrieval of parameters, modification and generate a new feature model according to the new value of parameters.

Because this method is on the basis of the original feature model, a new feature model is derived by modifying and designing parameters. In order to distinguish with a general three-dimensional model, these original characteristics models are called as template accessory or template mold.

3 Key Techniques

In the standard parts library, the parametric process must have the extraction of main parameter and designing problem of a parameter-driven algorithm, which is the key un-solved technical problems in characteristic parametric designing of the standard parts library. Here the mentioned parametric designing is the parameters of the driven sizes, the extracted parameters are generally characteristic of the size parameters, mainly for structural parameterization.

3.1 Extracting Main Parameters

The main parameter is an independently changing parameter. The other parameters can be determined after main parameter determining by certain constraints conditions or determine the relationship with main parameters, which called as sub-parameters.

Extraction of main feature parameter: the main parameter must satisfy independency, the correlation often leads to contradictions. In addition, designing specifications and constraints determined by user requirements are closely related and can be controlled by users. The intention of standard characteristic parameters is to take advantage of users' requirements, then directly obtain standard digital models, depending on different inputting parameters .

The main geometric characteristics parameters: Once the main geometric characteristics parameters are established, other geometric sizes of accessories are determined by parameter-driven. Therefore, some performance indicators are used as the main geometric characteristics parameters, these parameters have their autonomy, independence and can not change.

3.2 Designing Parameter-Driving Algorithm

The parameter-driving algorithm operates graph data by using parameter-driving mechanism, to control the running of driving mechanism. Compared with the existing

parametric methods, this method can not change graphs to other expressing forms, such as equation, sign, etc; and it not inquires drawing process, but stress to understand the drawing process, just see graph as a model, a parametric basis. As a medium for drawer to communicate information. Drawers give their intention to parametric programme by graph, then parametric programme returns the needed graph for drawers. Therefore, the running is very simple and convenient.

Thus, the parameter-driven algorithm is the knowledge core of designing characteristic product category, the algorithm should demonstrate the transitive of the main parameters ascertaining sub-parameters, should have the relationships between accessory and standard-part library. The algorithm should require strict accordance for the main parameter-driven and the second binding constraints with the process order, it can not be free to drive the main parameters randomly.

4 Designing Standard-Part Library Based on Parameter-Driving

4.1 Designing Principle

To achieve parametric designing, we must establish the parametric model--accessory template or assembly template. With SolidWorks software, it only to do parametric designing, just design the parts of this series to be a template file (accessory template or assembly template) and mark size on the template file. The size line can be seen as a directed segment, whose above figures are parameter names, whose direction reflects the trend of geometry data, whose length reflects the value of parameter, thus the relationship between geometric entities and parameters is established. Through inputting parameter name, the corresponding entity is found, so as to edit and modify the entity according to the value of parameter. Thus the new entity is obtained, the parametric designing is achieved.

The main designing idea and developing steps of parameter-driving method are to make use of database to administer standard-part library parameters. In SolidWorks, the template of standard-part accessory is drawn, the file '*.sldprt' is saved. What's more, the size of size-driven is named, such as 'L@ draft 1'. Then open the '*.sldprt' file in the program running, replace 'L@ draft 1' size for the extracted parameter from the database, new accessory is produced finally.

4.2 Design Example

(1) Main steps of parameter-driving method

1) At first, construct a accessory on SolidWorks, and denote the size of needed size-driven (such as size name 'h@matrix - tension '), so as to preparing extract, then save the file with sign '.sldprt'.

2) In the program, use the OpenDoc4 function to open the front saved '.sldprt' file. Note that OpenDoc4 open the file only in memory, but does not activate and display this file, and OpenDoc4 function should return a valid ModelDoc pointer.

3) Then need to use ActivateDoc2 (with SolidWorks version increasing, the numbers behind the function will change accordingly, the parameter list may be changed) or get_IActiveDoc Function activates and display this file. The function of ActivateDoc2 Function is to activate a loaded file, making this file into the Active File.

4) Use `IParameter (stringIn, &retval)` to return the replaced parameter pointer. Pay attention to parameter `stringIn` must be full name such as 'h@ matrix - tension ', can not only be 'h'.

5) Use the function of `SetValue2(newValue, whichConfigs, &retval)` to set up new values.

6) Use `EditRebuild()` to establish model newly at last.

(2) Case

The completed function of this case is to extract punch head size ' h@ matrix - tension ' from the file ' B type circular punch.sldprt', then use the value of variable S to replace, from this case we can see a simple and complete parameter-driven process, which includes several steps as follows:

1) drawing three-dimensional template file

Using SolidWorks to draw 'B type circular punch.sldprt' accessory, denoting punch head size as "h@matrix - tension" and save.

2) building new project

Using SolidWorks Add-In AppWizard guide to build a new project, called as "parameter-driving method", accept all default options.

3) Writing code

Add void type `OPEN()` function for class `CMyApp`, and add codes shown as follows:

```
void CMyApp::OPEN( )
{
    LPMODELDOC m_ModelDoc=NULL;
    HRESULT hres;
    LPPARTDOC m_PartDoc = NULL;
    long errors;
    hres=TheApplication->GetSWApp()->
    IOpenDoc4(_T("E:\\B type circular punch.sldprt"),
        swDocPART, swOpenDocOptions_Silent,_T("default ")
        ,&errors,&m_ModelDoc);
    hres = TheApplication->GetSWApp( )->
        get_IActiveDoc( &m_ModelDoc );
    if( m_ModelDoc == NULL )
        return;
    LPDIMENSION m_Dimension=NULL;
    long retval;
    hres=m_ModelDoc->IParameter(
        auT("h@matrix - tension@B type circular punch.sldprt"),
        &m_Dimension);
    AfxMessageBox(_T(" replace parameter 2 "),MB_OK,0);
    int s=30; // set up new value of h@matrix - tension
    //replace the value of s for the value of "h@matrix - tension" size parameter
    m_Dimension->SetValue2(s,0, &retval);
    hres=m_ModelDoc->QueryInterface
        (IID_IPartDoc,(LPVOID *)&m_PartDoc);
    m_PartDoc->EditRebuild( );
}
```

```

m_ModelDoc->Release();
m_Dimension->Release();
m_PartDoc->Release();
return;
}
4) The correlation between Menu and OPEN() function
Call OPEN() function in the menu response function.
void MenuItemCB(void)
{
    TheApplication-> OPEN();
return;
}
5) running program

```

After clicking menu item Example App/Example Menu Item, then the program implements OPEN function, the extracted punch head size 'h@ matrix - tension' from the file 'B type circular punch.sldprt' is changed, the value of 'h@matrix - tension' is replaced as s=30. The contrasts before and after loaded program running are shown as Fig 1.

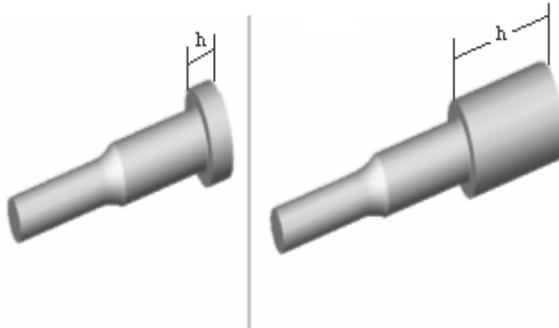


Fig. 1. Accessory contrasts before and after program running

5 Conclusion

All in all, parameter-driven is a new method of parametric designing, whose basic features are of direct operation to database. Thus it has good interactive. Users can modify graphics and its attributes by making use of all the interactive features of CAD planar or three- dimensional drawing system, so as to control the parametric process; compared to other parametric methods, parameter-driven method is convenient to develop and use, and can facilitate to redevelop on the basis of the existing drawing systems.

In this paper, the mould standard-part library is established by using parameter-driven technique based on SolidWork, which fully embodies the facility of the redevelopment of parametric modeling software, should not compile a large number of programs to complete drawing work not as the parametric modeling software of

non-parametric function. The application of this technique can improve the efficiency of the redevelopment, shorten development cycles, and bring some advantages of future maintenance and upgrading of this software.

References

1. Liu, J., Zhu, C., Sun, J.: Research and Realization of Equipment Library Based on SolidWorks. *Application Research of Computers* 23(01), 132–134 (2006)
2. Qi, C.: Development of Three-dimensional Standard Parts Library Based on SolidWorks. *Mechanical Management and Development* (06), 168–170 (2009)
3. Zhang, C.: Development of Die & Mould Standard Parts Library with SolidWorks-based Visual C++. *Die & Mould Industry* (03), 13–16 (2006)
4. Huang, X.: Development of Three-dimensional Standard Parts Library of Rolling Bearing Based on SolidWorks. *Journal of ChangZhou Institute of Technology* 22(03), 7–9 (2009)
5. Du, M.: Studies on the Method of Secondary Development of Solidworks Using Visual C++. *Science Technology and Engineering* 06(10), 1425–1427 (2006)

Study on Magnesium Alloys Ignition Temperature Test System

Li Yongyan, Zhao Weimin, Xue Haitao, and Ding Jian

School of Material Science & Engineering
Hebei University of Technology
300132, Tianjin, China

liyongyan@126.com, wmzhao@hebut.edu.cn, xuehaitao@126.com,
djian0122@yahoo.com.cn

Abstract. The hardware system and software system of magnesium alloy ignition temperature test system were designed. The hardware system was composed of well type electric resistance furnace, temperature controller, thermocouple, data acquisition card and computer. The software developed by LabVIEW 8.5 could deal with data acquisition, dynamic display, experimental data storage and data analysis. The test result shows that the test system can acquire the temperature –time curve well. The method of determining ignition temperature was presented that could accurately determine the magnesium alloy ignition temperature.

Keywords: Data acquisition, Magnesium alloy, Ignition temperature, Test and measurement.

1 Introduction

Magnesium and magnesium alloys are widely used in auto, electronics, aerospace field etc due to a number of good properties such as high specific strength, specific stiffness, good isolation performance, easy recovery, easy cutting and so on. Despite these properties, the use of magnesium alloys is still relatively limited, because magnesium oxidates and ignites easily at high temperature, and it can't automatically form compact oxide film on the surface of melt after melting which brings many difficulties during the melting and processing of magnesium alloys. So it is necessary to study and improve the ignition-proof performance. In 1950s, modern scientists put forward the opinion on alloying ignition-proof, which means, during the process of melting, magnesium alloys can automatically form protective oxide film through adding alloy elements, which will prevent further oxidation or even burning of magnesium alloys. However, how to determine correctly the ignition temperature point is a critical problem during the study of ignition-proof magnesium alloys. In this paper, a system used for determining the ignition temperature of magnesium alloys was developed by using data acquisition technique.

2 Developing of Magnesium Alloys Ignition Temperature Test System

The ignition temperature test system of magnesium alloys was composed of hardware system and software system. Functions of hardware system mainly include temperature sensing, signal conditioning and data acquisition. Functions of software system mainly include show temperature data in real-time, storage experimental data and analysis experimental data.

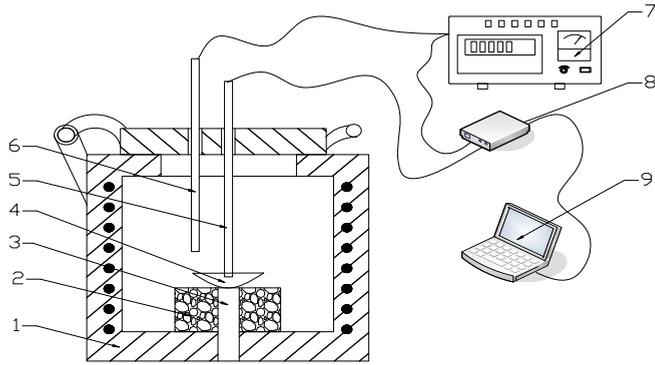
2.1 Design of Hardware

The hardware system was composed of well type electric resistance furnace, temperature controller (TCW-32B), thermocouple, data acquisition card and computer, Fig.1 represents its structure. This electric resistance furnace and temperature controller has the advantages of energy-saving, programmable and heating rate controllable and so on. Temperature sensor is two K-type thermocouple. As the temperature acquisition doesn't require a high sampling rate, Advantech USB-4718 data acquisition card was in used. This data acquisition card is 8-channel thermocouple input USB module. Because the data acquisition card supports USB 2.0 and Bus-powered, the data acquisition system based the USB-4718 data acquisition card and notebook PC is portable. Besides, this data acquisition card has advantages such as 2,500 VDC isolation, supporting 4-20 mA current input, embedded signal conditioning module and so on, the developing costs and time is reduced. When the temperature is tested, two thermocouples are used. One thermocouple is used to control furnace temperature, so it is connected firstly to temperature controller and then connected to data acquisition card. Another one thermocouple is used to test samples, it is connected directly to the data acquisition card. Both two channel temperature signals are sent to computer by USB interface from data acquisition card, therefore, the temperature data can be saved and displayed by computer.

2.2 Software Functions and Key Technologies

The software system of magnesium alloys ignition temperature test system was developed by LabVIEW8.5, the interface is shown in fig. 2. The software interface is functionally divided into three areas: setting area, dynamic display area and control area. The parameters such as the type of thermocouple, the channel of data acquisition card, the experiment name, the storage path of experimental data and the sampling rate etc can be set in experimental parameter setting area.

Dynamic display area of experimental parameter can dynamic display temperature data of two channel and temperature-time curve. Software function control area can control experimental progress, such as the beginning and ending of data acquisition, the refreshing of interface, the playback and display of temperature-time curve and exit the system etc.



1. well type electric resistance furnace; 2. firebrick pad; 3. vent; 4. stainless crucible; 5、6. electric thermo-couple; 7. temperature controller; 8. data acquisition card; 9. computer

Fig. 1. The sketch of magnesium alloy ignition temperature test system

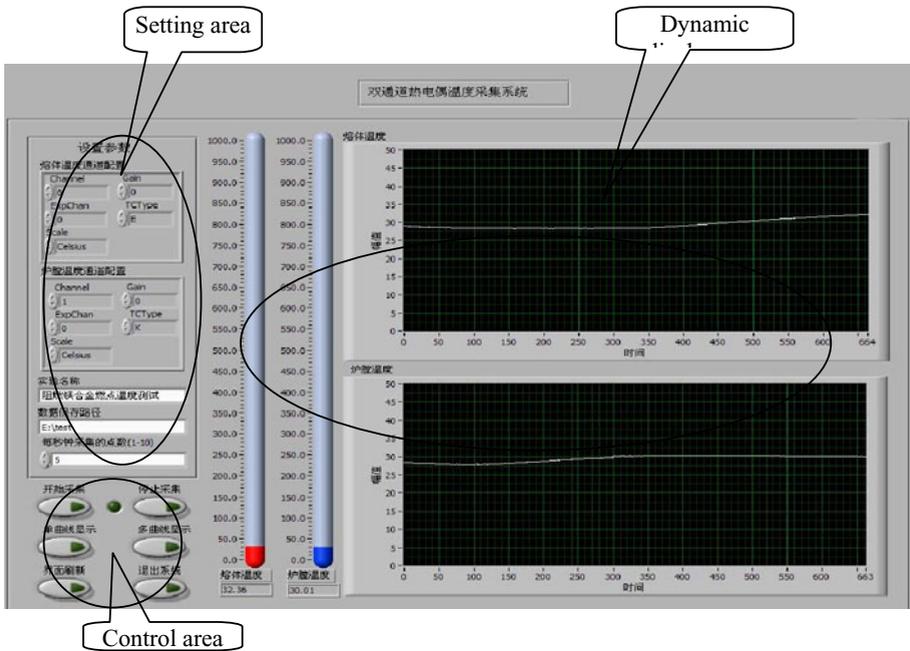


Fig. 2. Software interface of magnesium alloy ignition temperature test system

The communication between hardware and software was achieved through function of the underlying drivers provided by the Advantech. Fig. 3 represents a program diagram of data transfer. This data acquisition card doesn't support the parallel communication of 8-channel signals, so the transmission of multiple signals must be finished by software. In this paper, the method of achieving dual-channel data transfer is time-sharing use sequence structure in LabVIEW.

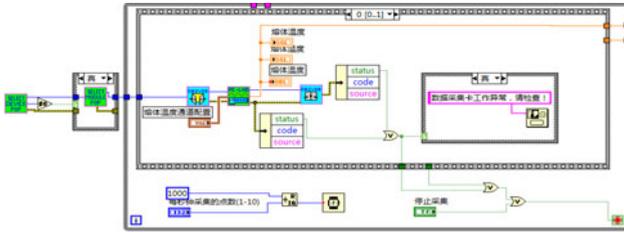


Fig. 3. The program diagram of data transfer

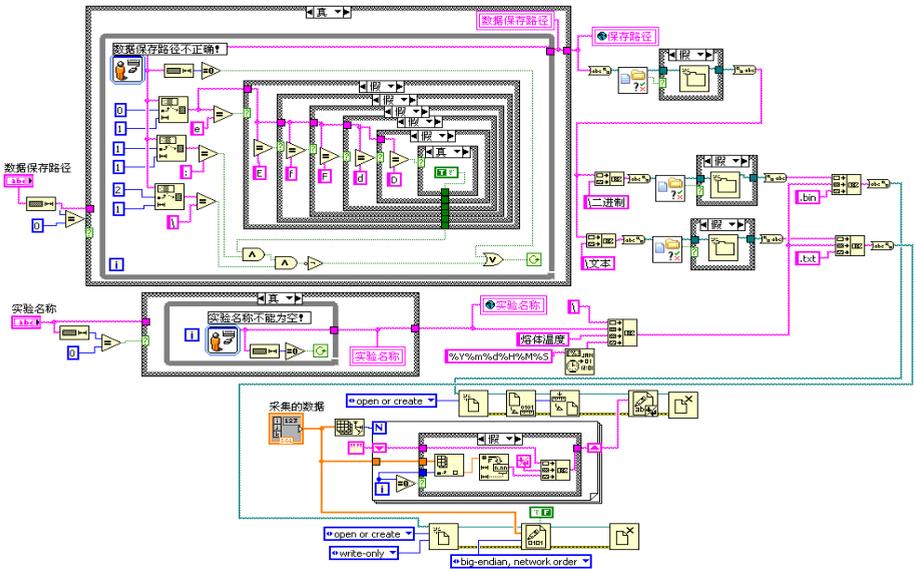


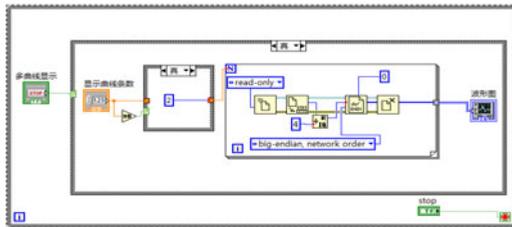
Fig. 4. The program diagram of data storage

In order to analyze the experimental data after testing, software must provide the function of data storage. The saved data must be corresponding to match experiment. The every experimental data is named by method of man named part adding computer named part. Man named part is experiment name, computer named part is time character string created by computer. For example, an experiment name is “mg0.5y1ce” and time character string is “20100531204001”, so the experimental data name is

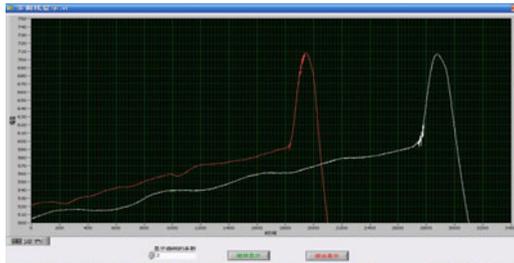
“mg0.5y1ce20100531204001”. The saved formats of experimental data are binary system and text formatting. The program diagram of creating experimental data name and data storage sees Fig.4.

The display process of experimental data is dynamic process when temperature was tested. In this process, the panorama of experimental data display can't be showed. Therefore, the software has a function of data redisplay. In this function, one or more temperature curve could be showed. This function was realized by calling a subprogram (Sub Virtual Instruments). Selecting “single curve display” button or “multi curve display” button in main interface will call a new display widows to realize data redisplay. The number of curve that showed in “multi curve display” must be input by user. The interface and program diagram of “multi curve display” see Fig. 5.

For convenience, the software has a function of interface refurbish. The software interface will resume default interface when selecting the button of “interface refurbish”. In order to software runs smoothly, some fault tolerance technique were used in this software. The application of fault tolerance technique also avoid computer crash which brought by misuse software or hardware trouble.



(a) Program diagram



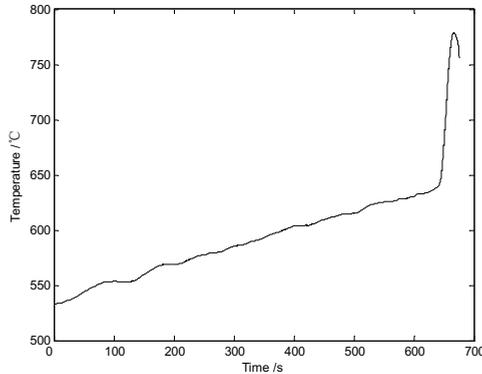
(b) Program interface

Fig. 5. Multi-curve display

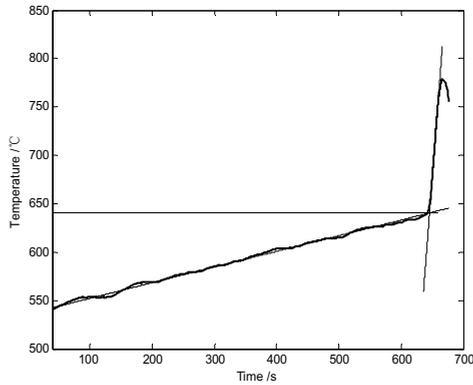
3 The System Test and Ignition Temperature Definition

Heat the electric resistance furnace to about 500°C, then put into a stainless steel crucible loaded with a magnesium alloy block, and, keep closed contact between thermocouple which test magnesium alloy temperature and the magnesium alloy block. Finally, heat the furnace at constant heating rate and began to collect data. The temperature-time curve collected as shown in Fig. 6(a). From the Fig. 6(a), it can be

seen that a inflexion point at about 640°C, this is because when magnesium alloy ignites, it will release large amount of heat, so that the temperature rise suddenly, therefore, the inflexion point is defined as ignition temperature of magnesium alloy [1-4]. Fitting data on the left and right of the inflexion point separately, two beelines will be achieved, and the intersection of two beelines is defined as ignition temperature point, the Fitting result shown in Fig. 6(b).



(a) Temperature-time curve



(b) Calculate result of ignition temperature

Fig. 6. Temperature-time curve and determine of ignition temperature

4 Conclusion

1) Designed the hardware system used for testing the magnesium alloys ignition temperature. The system was composed of well type electric resistance furnace, temperature controller, electric thermocouple, data acquisition card (Advantech USB-4718) and computer. The system has many advantages, such as low-cost, simple structure and portable system.

2) The software system used for testing the magnesium alloys ignition temperature was developed by LabVIEW8.5. The functions of software have data acquisition, data storage, data dynamic display and redisplay, experimental data analysis and so on. It has been proved that the software has the friendly interface, the simple operation and the strong fault tolerance.

3) The tested results show that the system could accurately test the ignition temperature of magnesium alloys. Finally, the method to determine ignition temperature was presented in this paper.

Acknowledgment. This work was supported by the Natural Science Foundation of Hebei Province, China (No. E2008000045, E2010000057 and E2010000121).

References

1. Ravi Kumar, N.V., Blandin, J.J., Suery, M., et al.: Effect of alloying elements on the ignition resistance of magnesium alloys. *Scripta Materialia* 49(3), 225–230 (2003)
2. Huang, X.F., Zhou, H., He, Z.M.: Influence of Ce rare-earth on ignition temperature of magnesium alloy. *The Chinese Journal of Nonferrous Metals* 11(4), 638–640 (2001)
3. Liu, Z.J., Li, F.Z., Zhang, L., et al.: Study on burning point and the corrosion resistance of magnesium and magnesium alloy. *Journal of Harbin University of Science and Technology* 5(6), 56–59 (2000)
4. Chen, P., Zhang, M.X.: Measurement for the flammability point of magnesium and magnesium alloy. *Special Casting & Nonferrous Alloys* (2), 75–76 (2001)

Defect Recognition of Resistance Spot Welding Based on Artificial Neural Network

Li Yongyan, Zhao Weimin, Xue Haitao, and Ding Jian

School of Material Science & Engineering,
Hebei University of Technology, 300132, Tianjin, China
liyongyan@126.com, wmzhao@hebut.edu.cn, xuehaitao@126.com,
djian0122@yahoo.com.cn

Abstract. The nugget size forecast model was built based on BP algorithm of Artificial Neural Network. The input parameters of the model are two characteristic number extract from electrode displacement curve. The output parameter of the model is nugget size. The model has three layers and the hidden layer have five nodes. The transfer function of hidden layer is Sigmoid function and the transfer function of output layer is linear function. Measured value and forecast value was analyzed by comparative method and the difference value between them was calculated. The results showed that 83% of the difference value is less than 1 millimeter. Based on the nugget size forecasted by the model, a method to identify incomplete fusion defect of resistance spot welding was suggested. In this method, 7 millimeter nugget size was regarded as criterion of recognition incomplete fusion defect of resistance spot welding. The result showed that recognition accuracy rate up to 94.3%.

Keywords: Artificial neural network, Nugget size forecast, Defect recognition, Incomplete fusion defect.

1 Introduction

Incomplete fusion defect is a general term of welding area which includes the faying surface of resistance spot welding no melt, plastic connection, ring nugget or undersize nugget [1]. All of these states mentioned above lead to decreased strength of spot weld. Incomplete fusion defect is a serious defect in resistance spot welding. Because incomplete fusion defect can't be distinguished by general NDT method, when spot welding structure which contain incomplete fusion defect works, loose weld may take place or parts may break off under affected by complex stress. Especially in aerospace, the parts breaking off may bring on short circuit or Drawn into engine High-speed rotated, these situations will lead to Catastrophic incidents. Nugget formation process is closed, so welding area can't be saw directly, therefore, it is difficult to distinguish if there is incomplete fusion defect in welding area from Weld appearance.

Recently, the study on spot welding process parameter shows that there is much information of spot welding quality in spot welding process parameters [2-5]. If this

information can be extracted, the spot welding quality would be assessed by spot welding process parameters.

Artificial Neural Network (ANN) is a new interdisciplinary technology which widely used in pattern recognition, system identification, forecast, data mining, optimization, signal processing and so on. It is a research focus to forecast quality of resistance spot welding by using ANN recently. Because the welding process is complex, the factors affected welding quality are various and the welding equipment and welding materials are different, every ANN established by every expert is various and the input parameters of ANN is distinguishing feature^[6-10]. In this study, a method which can identify incomplete fusion defect was suggested based on BP algorithm of ANN. The feature Information used in the method was extracted from resistance spot welding process parameter curve.

In this study, the resistance spot welding process parameter collected from direct current resistance spot welding equipment. The pressure curve of welding process parameter is staircase curve and the welding test plates are aluminum alloy which thickness is 4 millimeter and 2 millimeter.

2 The Input Parameters of ANN

The main feature of incomplete fusion defect is no nugget or undersize nugget in faying surface after welding. Therefore, if the nugget size of resistance spot welding can be forecasted accurately, the recognition of incomplete fusion defect become easily.

The studied shows that the electrode displacement signal contained plenty of nugget size information. Figure.1 respectively shows electrode displacement signal of acceptable spot weld and incomplete fusion spot weld. The moment marked ① is the beginning power moment, at this moment, workpiece begin to expand because of resistance heating and electrode begin to move upward slightly. More quantity of heat produced in welding area, more expansion produced in workpiece and more displacement produced in electrode. However, if the quantity of heat produced by resistance heating is less, the displacement produced by electrode is smaller. The quantity of heat produced by resistance heating in welding area determined the amount of melted-metal, that is to say the quantity of heat produced by resistance heating in welding area determined the nugget size. Therefore, the electrode displacement produced in power time which marked by ① and ② at electrode displacement curve is corresponding to the nugget size. On the other hand, in forging time which marked by ② and ③ at electrode displacement curve, if the amount of melted-metal in welding area is more, the depth of indentation is bigger. Otherwise, the depth of indentation is smaller. The indentation represents the change of electrode displacement in forging time. So the change of electrode displacement in forging time also is corresponding to the nugget size. Therefore, both the volume of increase in electrode displacement curve marked by ① and ② and the volume of decrease in electrode displacement curve marked by ② and ③ are used as input parameters of ANN forecast model.

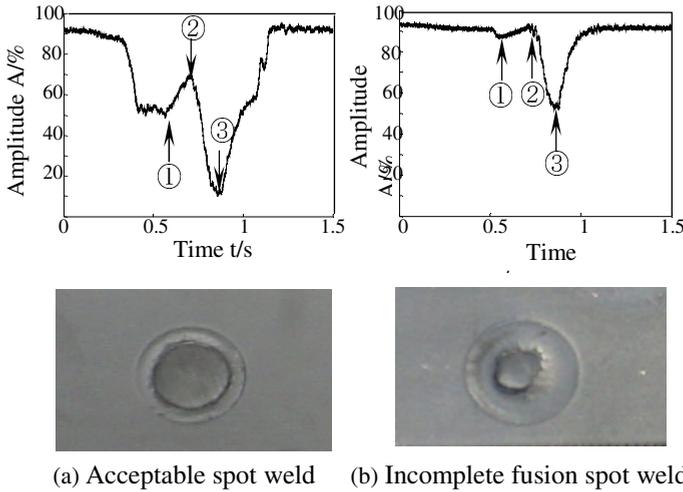


Fig. 1. The electrode displacement curve of acceptable spot weld and incomplete fusion spot weld

3 Ann Model Building and Its Training Based BP Algorithm

Two input parameters chosen above determined the input node number of ANN forecast model. The ANN output is generally the function of the model, so the nugget size is used as the output of ANN forecast model. Thus, the number of ANN input and output layer nodes are determined simultaneously, that means input layer nodes is 2 and output layer nodes is 1.

Theoretical analysis shows that single hidden layer feed-forward ANN can map all the continuous functions, only when learning discontinuous function, it requires two hidden layers, and therefore, feed-forward ANN requires at best two hidden layers. When a multilayer feed-forward ANN was designing, usually design one hidden layer firstly, only when a large number of hidden layer nodes still cannot improve ANN performance, it should be considered to add a hidden layer again. The ANN forecast model built in this paper initially set one hidden layer.

The role of hidden layer nodes is to extract and store its internal rules from the samples, each hidden layer node has a number of weights, and each weight is a parameter increasing the mapping capability of the ANN. So if set fewer hidden layer nodes, the ANN capacity to learn information from the samples will be poor, so that it will not sufficient capacity to summarize and reflect the law of training samples; But if set too many hidden layer nodes, the ANN possible remember some non-law content such as noise, resulting in an “over-match” problem and reducing the generalization capability. In addition, excessive number of hidden nodes will increase the training time. The method of setting hidden layer node is to set less nodes firstly, then gradually increase the nodes and train using the same samples, when the output error of ANN is the least, the corresponding number of nodes is acceptable.

Ref[11] given an empirical formula of an initial number of hidden layer node: $m = \sqrt{n + l} + a$, m is the number of hidden layer node, n is input layer nodes, l is output layer nodes and a is a constant between 1 and 10. This study takes $a=4$, so according to the formula above, it can set the number of ANN hidden layer node is 5.

Initial weights of ANN determined the ANN training starting from which point on the error surface, so the initialization method is essential to shorten the ANN training time. Transfer function of neurons are symmetric about zero, if the net input of each node are near zero, then the output are in the middle of the transfer function. This location not only away from the two saturated zone of the transfer function, but also is the most sensitive region, so it will certainly speed up the learning speed of the ANN. General initialization method is to initialize all the initial weights to a random value near zero.

Generally, the transfer function of feed-forward ANN based BP algorithm uses the Sigmoid function. In the model built in this study, the target output is forecasted nugget size, in order to minimize the impact of Sigmoid function saturated zone to the forecast result, the transfer function of output layer is linear function.

According to the analysis above, the spot weld nugget size forecast model based on BP algorithm of ANN as follows:

The forecast model is multilayer feed-forward ANN based on BP algorithm, there are three layers in the model, including a input layer, a hidden layer and a output layer; the nodes number of three layers is 2, 5 and 1 respectively; the transfer function of hidden layer is Sigmoid function, and that of output layer is linear function. The structure of ANN model shows in Figure 2.

The number of training sample is 145. These samples include acceptable spot weld and incomplete fusion spot weld. All the samples will be divided into three parts, in which, one half of the samples will be used for ANN training, a quarter for the test of ANN and the remaining part for the ANN validation. The input and output data of the ANN should be subject to standardization.

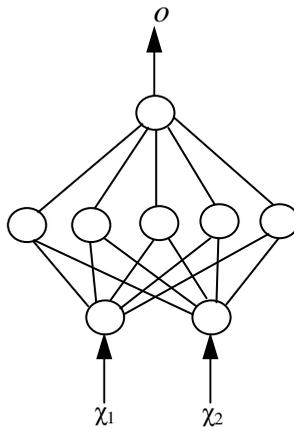


Fig. 2. The structure of ANN forecast model

The ANN can be trained after initialized the weights and thresholds. In the training process, in order to reduce the mean square error of the ANN output and target output, the weights and thresholds of the ANN will be adjusted repeatedly.

MATLAB Neural Network Toolbox provides a variety of training algorithm, which mainly can be divided into general training algorithm and fast training algorithms. Fast training algorithm has a faster convergence rate, and some even faster than normal training function a few times. In this paper, a faster algorithm named Levenberg–Marquardt was in used, the mean square error trend during ANN training process shown in figure 3. It can be seen from the figure 3 that with the training process progresses, the ANN output error is lower, and after 11 times of training, reached a stop condition.

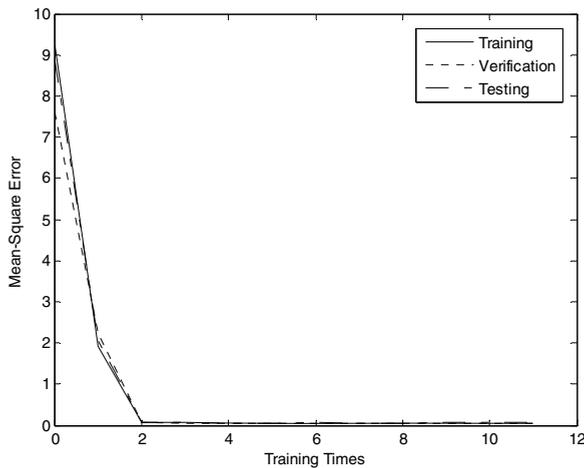


Fig. 3. The trendline of mean square error during training

4 Forecast Performance Evaluations of Ann Model and Defect Recognition

Using the validation data simulates the trained ANN, and compares with measured results. ANN simulation results should go through denormalization processing. ANN simulation results compared with the measured results shown in figure 4.

The Compared results showed that 83% of the differential value between forecast value and measured value is not more than 1mm.

Through further regression analysis to the simulation results, quantitative indicator used for forecast model evaluation can be achieved. Take measured value as the abscissa and the ANN forecast value as the vertical axis, depict these points and regression analyses these points, the results shown in figure 5. The solid line in figure 5 is the regression line for points, the dotted line is the ideal regression line, that is line $A=T$, that means ANN output error is zero, forecast values are equal to the measured values. The higher coincidence degree between the solid line and dotted

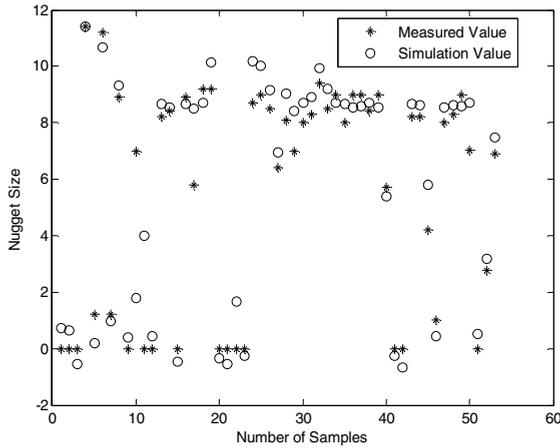


Fig. 4. The compare of simulation value and measured value

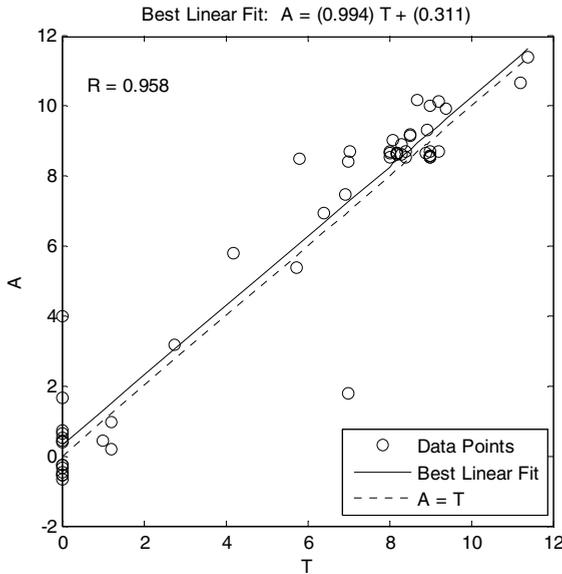


Fig. 5. Regression analysis of ANN forecast model performance

line, the better the ANN forecast model performance. Figure 5 also shows three quantitative indicators to describe the ANN forecast model performance. In the equation $A=0.994T + 0.311$ which on the top of figure 5, the coefficient 0.994 represents the slope of the regression line, 0.311 represents the intercept of regression line. The slope of ideal regression line and the actual regression line. The correlation coefficient is closer to 1, the better the ANN the regression line is closer to 1, the intercept of regression line is closer to 0, and then the ANN forecast values are closer to measured values. In addition, there is another parameter R in figure 5, which

represents the correlation coefficient of the forecast model performance. Of course, when the ANN forecast model performance is evaluated, three parameters should be taken into account simultaneously.

Take 7 millimeter nugget size as criterion of recognition incomplete fusion defect of resistance spot welding and according to the forecast results of nugget size, the incomplete fusion defect can be identified. If the forecasted nugget size is less than 7 millimeter, the spot weld is determined incomplete fusion spot weld. The results show that the defect recognition accuracy rate is up to 94.3%.

5 Conclusions

(1) The nugget size forecast model was built based on BP algorithm of Artificial Neural Network. The input parameters of the model are two characteristic number extract from electrode displacement curve. The output parameter of the model is nugget size.

(2) The tested results shown that 83% of the differential value between forecast value and measured value is less than 1 millimeter.

(3) Take 7 millimeter nugget size as criterion of recognition incomplete fusion defect of resistance spot welding and according to the forecast results of nugget size to recognize incomplete fusion defect. The recognized results show that the defect recognition accuracy rate is up to 94.3%.

Acknowledgment. This work was supported by the Natural Science Foundation of Hebei Province, China (No. E2010000057 and E2010000121) and International S & T Cooperation Program of China (No. 2010DFA51850).

References

1. Institute of Welding, Chinese Mechanical Engineering Society. The Theory and Practice of Resistance Welding. Machinery Industry Press, Beijing (1994)
2. Luo, X., Ji, C., Zhang, C., et al.: Quality Control Visualization of Direct Current Spot Welder with Secondary Rectification. Transactions of The China Welding Institution 22(1), 59–61 (2001)
3. Hao, M., Osman, K.A., Boomer, D.R., et al.: Developments in Characterization of Resistance Spot Welding of Aluminum. Welding Journal 75(1), 1s–8s (1996)
4. Xue, H., Li, Y., Cui, C., et al.: Extraction of diagnostic information of expulsion defect in resistance spot welding process by wavelet analysis method. Transactions of The China Welding Institution 28(5), 38–40 (2007)
5. Niu, Y., Xue, H., Zeng, Z., et al.: Data acquisition and defects analysis in resistance spot welding process based on parallel port. Transactions of The China Welding Institution 28(6), 61–64 (2007)
6. Osman, K.A., Higginson, A.M., Kelly, H.R., et al.: Monitoring of resistance spot-welding using multi-layer perceptrons. Intelligent Engineering Systems Through Artificial Neural Networks 4, 1109–1114 (1994)

7. Cho, Y.J., Rhee, S.: Quality Estimation of Resistance Spot Welding by Using Pattern Recognition With Neural Networks. *IEEE Transactions on Instrumentation and Measurement* 53(2), 330–334 (2004)
8. Lee, H.-T., Wang, M., Maev, R., et al.: A study on using scanning acoustic microscopy and neural network techniques to evaluate the quality of resistance spot welding. *International Journal of Advanced Manufacturing Technology* 22(9-10), 727–732 (2003)
9. Chen, Y., Hu, D., Ma, L., et al.: Neural Network Model for DC Spot Welding of Aluminum-alloy. *Transactions of The China Welding Institution* 21(4), 20–23 (2000)
10. Fang, P., Tan, Y., Wu, L., et al.: Artificial Neural Networks Applied to the Quality Control in Alternating Current Resistance Spot Welding. *Acta Aeronautica Et Astronautica Sinica* 21(1), 94–95, 86 (2000)
11. Han, L.: Artificial neural network theory, design and application. Chemical Industry Press, Beijing (2002)

Theoretical Research on University Outward Bound Curriculum System

Xinhua Zhu*

P.E. College of Jiangnan University Wuhan Hubei, China
tigerandsheep@163.com

Abstract. This paper uses the methods of documentary, interview, inductive analysis and so on, give a profound research on the meaning, property, the construction of curriculum system, the formulation of teaching outline of outward bound curriculum. Aim to establish a outward bound curriculum system which helps to cultivate students' attic faith, strong self-confidence and brave strong will; formulate a outward bound curriculum teaching outline which can effectively promote the implementation of the outline, can carry out reform and innovation of the traditional P.E. class.

Keywords: College sports, outward bound, curriculum system.

1 Introduction

Through the consulting of document and literature about outward bound in the recent 5 years, I find that outward bound into college P.E. curriculum not only enrich the content, strengthen physical education curriculum's practicability, enjoyment, and challenge, embodies the varieties and value of P.E. curriculum, but also good for universities to better realize the curriculum target and character of P.E. curriculum, to construct the new teacher-and-student view and the positioning of the role of teacher. However, how to blend the projects which are suitable for universities together, construct and perfect college "outward bound curriculum", promote the full expansion of outward bound in universities, all the above problems are what this paper going to solve.

2 The Construction of Outward Bound Curriculum

2.1 The Definition of Outward Bound Curriculum

The original meaning of outward bound is a small ship departs the safe bay, and sailing for the dangerous sea to meet challenges.

Outward bound is a modern people and modern organized learning method and trainingmethod, it uses varieties of typical scenes and active ways let team and individual experience a series of tests, cultivate their perseverance of overcoming difficulties, healthy mental quality, and positive attitude towards life, and enhance their sense of team work, those challenging and niche-targeting sports groups are.

* Corresponding author.

In the comprehensive analysis of the basic characters of all kinds of outward bound, we put forward some concepts as follows: outward bound curriculum is to make full use of the natural environment, with the common basic knowledge and skills of those item groups as the main teaching content, to cultivate students' physical and mental quality, and is the process conducted according to the teaching plan.

2.2 The Construction of Curriculum Structure System

For the theoretical basis of outward bound curriculum, the academic circle has relatively comprehensive statement of it: "through the scene setting of the outdoor experience program, make participants fully experience all kinds of emotions, especially the negative emotions, thus get further understanding of the mental reaction and effect of external stimuli, so as to control and realize the overtake of the past ourselves, and cooperate with others."

Good physical quality, healthy, mature psychological quality and basic theoretical knowledge and skills of outward bound are the necessary qualities of completing outward bound. Those qualities can be realized by the theory and skill teaching and training.

3 The Formulation of Curriculum Teaching Outline

Complete the outline of outward bound curriculum, make concrete arrangements of course schedule, content, tasks, requirements, organizations. The teaching outline of outward bound curriculum please see table 1.

Table 1. The teaching outline of outward bound

| Teaching Progress | Items | Purpose | Teaching Content | Teaching Requirements | Organizing Methods | Number of Lessons |
|-------------------|-------------------------------|--|--|---------------------------|---------------------------|-------------------|
| 1 | Introduction of outward bound | 1.preliminary understanding the origin, character, intention and meaning of outward bound 2.stimulating studying interest | 1.origin of outward bound, characters of development course 2.meaning of the curriculum 3.domestic and foreign developing 4.content of curriculum | learn about outward bound | classroom theory-teaching | 2 |

Table 1. (Continued)

| | | | | | | |
|---|------------------------------|--|---|--|--|---|
| 2 | Security awareness education | <p>1.Master common sports injury disposal methods</p> <p>2.strengthen the knowledge, awareness and ability of self-protection</p> | <p>1.the reason of sports injury and the judgment and handling</p> <p>2. common diseases and processing (haemostatic, bandage, fracture and fixed, the wounded handling, etc.)</p> <p>3.diet and health</p> | <p>master prevention and treatment</p> | <p>classroom theory-teaching</p> | 2 |
| 3 | Physical exercises | <p>1.fully understand good physical quality is the prerequisite for expanding movement</p> <p>2. master the methods of training physical quality.</p> <p>3. learn to make exercising schedule.</p> | <p>1.the necessity of physical training</p> <p>2. ways and contents of physical quality training.</p> | <p>1. learn the methods and ways of body exercising .</p> <p>2.learn to make exercising plan.</p> | <p>practice lessons: 5-8 people form a group, and do exercises or competitions assigned by teachers.</p> | 2 |
| 4 | Along-the-rope-down | <p>1. cultivate the ability of understanding and mastering targets, the skill of effective communication, and team work spirit.</p> <p>2. develop self-potential, trigger their self-transcendent conscious.</p> | <p>1.psychological outward bound</p> <p>2.physical training</p> | <p>1.every student complete the task as required</p> <p>2. students should encourage each other and be responsible.</p> <p>3.complete the quality training</p> | <p>practice lessons: 5-8 people form a group, and do exercises or competitions assigned by teachers</p> | 2 |

Table 1. (Continued)

| | | | | | | |
|---|-----------------------|---|--|--|--|---|
| | | 3. cultivate high sense of responsibility and the mutual trust between members. | | | | |
| 5 | DOA wall | 1. cultivate the ability of understanding and mastering targets, the skill of effective communication, and team work spirit. 2. develop self-potential, trigger their self-transcendent conscious. | 1.the creation and repair of tools for outward bound: stoves, hammocks, stretcher, ladders 2.physical training | 1.every student complete the task as required 2. students should encourage each other and be responsible. | practice lessons: 5-8 students form a group, and combine theory and practice together. | 2 |
| 6 | The creation of tools | 1. cultivate students' ability. 2.cultivate teamwork spirit | 1.the creation and repair of tools for outward bound: stoves, hammocks, stretcher, ladders 2.physical training | 1. can make and use tools. 2. complete the quality exercise. | practice lessons: 5-8students form a group, and combine theory and practice together. | 2 |
| 7 | Electric net | 1.know the origin, development, character and meaning of electric net 2.master the essential methods of crossing the electric net 3. learn the importance of cooperation division and obey the arrangement of organization. | 1.introduce the origin, development, character and meaning of electric net 2. the technical essentials and methods of crossing electric net. 3.the team cooperation precautions 4. physical quality training. | 1. learn the origin, character and meaning of electric net. 2.keep team work in mind 3. complete quality training. | practice lessons: 5-8 students form a group, and combine theory and practice together. | 2 |

Table 1. (Continued)

| | | | | | | |
|----|-----------------|---|--|--|---|---|
| 8 | Trust fell back | <ol style="list-style-type: none"> 1. master the skill and protection method of trust fell back. 2. cultivate students' aggressive, positive and cooperating spirit. | <ol style="list-style-type: none"> 1. the action essentials of trust fell back. 2. protection methods. 3. make slogans. | <p>check members' position and gestures.</p> | <p>practice lessons: 5-8 students form a group, and combine theory and practice together.</p> | 2 |
| 9 | Turn leaves | <ol style="list-style-type: none"> 1. improve students' release ability, break among all the barriers. 2. train all participants' cooperation ability. 3. Cultivate solidarity, closely cooperate with the team spirit | <ol style="list-style-type: none"> 1. understand mission requirements. 2. explain safety requirements. 3. physical exercises. | <p>lin flat open field.</p> <p>2. students in activities don't kick to teammates.</p> | <p>practice lesson: 3-4 people form a group. (inner group competition.)</p> | 2 |
| 10 | Water thorns | <ol style="list-style-type: none"> 1. understanding the importance of special talents for completing the task. 2. train all students the ability of cooperating with each other, and try their best. | <ol style="list-style-type: none"> 1. introduce the methods and meaning of water thorns. 2. group exercise. 3. physical quality exercise. | <ol style="list-style-type: none"> 1. check the site of sharp objects and the safety of appliance. 2. group those good physical quality and bad physical quality together. | <p>practice lesson: 5-8 people form a group and practice under the guidance of teacher. (inner group competition can be applied.)</p> | 2 |
| 11 | Bamboo dance | <ol style="list-style-type: none"> 1. understand the importance of bamboo dance. 2. grasp basic technology. 3. cultivate the spirit of unity and cooperation. | <ol style="list-style-type: none"> 1. introduce the origin of bamboo dance. 2. complete the ground of parallel stem dance. 3. group practice. | <ol style="list-style-type: none"> 1. check the field, clear out hard objects. 2. complete quality training. | <p>practice lesson: 5-8 students as a group. Teaching and practice combined.</p> | 2 |

Table 1. (Continued)

| | | | | | |
|--|---------------------------------|---|---|---|---|
| 12 | Theoretical examination | check the teaching effect | physical exercise methods, climbing technology, protection methods, use of maps and compasses, the normal injury treatment and diet health. | answer the test paper. | 2 |
| 13 | Technical and skill examination | check the teaching effect | hammock, stretcher, ladders, rafting, stove building and production, the wild campsite selection, construction, rock climbing technology and drop technology, climbing protection method, etc | draw assessment 1 oral answer 2 operation | 2 |
| Comprehensive training | | | | | |
| Aim and mission | | Teaching content | | Teaching requirements | |
| 1. cultivate the spirit of hard-working and brave heart. 2. cultivate team work spirit 3. challenging oneself and develop self-potential. 4. learn to challenge oneself and the nature. | | Island survival, along-the-rope-down, turn leaves, bamboo dance, trust fell back, etc, or the combination of 3-4 items. | | 1. it must be reasonable on the combination of items, and have abundance learning meaning and good effect on students. 2. teachers' and students' safety must be ensured, and protection measures must be taken when carrying out dangerous items. | |
| 6 | | | | | |

4 The Implementation of Curriculum Teaching

4.1 The Preparation of Curriculum

The arranged time for the curriculum is semesters, before the beginning of the course, introduce the curriculum and management through the university media.

Students enroll this course as the elective course, and proper expenses are required to pay for the course.

Unify organize the utilization of teaching resources reasonably. The number of students of the theory and practice lessons should be controlled within 40 people.

4.2 Way of Teaching and the Relative Content

Teaching adopts the method of combine theory and practice together. The comprehensive training as the improvement of the curriculum, will not be tested.

Theory classes use the multi-media teaching, practice lessons pay attention to students' mental and physical quality and the cultivation of surviving skills and techniques, try to make students, solve their problems through what they've learned, so as to cultivate their ability of innovation and meeting emergencies. Comprehensive training is the general improvement of students' comprehensive quality and studying results, the task of practice lessons is difficult and highly-required, it should be set after the theory lessons.

In practice lessons, according to the teaching requirements, teachers divide students into small groups of 5-8 people, implement leader responsibility system, to practice for unit, and adopt the interactive teaching method. Physical exercise throughout the whole process of practice lesson. It is arranged generally before the ending of the curriculum, and the time is between 20-25 minutes.

The comprehensive training should be indoor first, and then outdoor or the combination, food and equipment should be given out together to all groups, and they decide their food arrangement independently. The implementation should be strictly under the control of the teacher, if the condition is allowed, can let students identify the eatable plants r animals. In the process of teaching, we should cultivate the sense of protecting the environment.

4.3 Attention Items

4.3.1 The Selection and Innovation of Contents

Market-oriented operation of outdoor sports club or outward bound companies' rock climbing, mountaineering, tracing the creek, organic movement project are clearly not completely suitable for college students, and still need to be introduced according to the requirement of students and the conditions of the school. Therefore, from the advanced theory and some relatively mature items of outward bound, we must conform to the goal of personal training, combine students' psychological and physical characters, adapt to the teaching standards, and consider the facility conditions to design a individuation solution, the outward bound curriculum must be a "reform and innovation outward bound".

4.3.2 The Formation and Frequency of Curriculum

Universities organize it all by their own. This way asks universities to hire experienced teachers, and buy the equipments of outward bound, treat it as a compulsory course, cooperate with the professional outward bound center. Through this kind of cooperation, universities will not only solve the problem of funds, but also get better train of students. The most difficult is to cultivate the mutual-help spirit quality in the process. The trigger of this kind of emotion can't be too frequent, or it will reduce the psychological feeling. Thus we take students out to do field training, doesn't need to be like the P.E. class once a week, 1-2 time a semester will be fine.

4.3.3 The Requirement of Implementation

As outward bound is mainly consisted of indoor, outdoor and professional field these 3 kinds, so when we are arranging the curriculum, each team should be with one teacher to handle emergencies, and ensure the safety. Weather shouldn't be the barrier of the arrangement, but it can be arranged according to the weather.

4.4 The Test of Curriculum

The test should include theory, practice and usual performance. Theory takes about 30% of the total score, and the content is all the knowledge of outward bound; practice takes about 50%, and includes the techniques or skills; the usual performance takes 20%. Teacher should do comprehensive analysis towards students on the bases of objectivity and reality.

5 Conclusion

The content of outward bound is rich, the curriculum setting is scientific and reasonable, and students have high enthusiasm of participation, it conforms to the age character and requirement of students, has a deep physical and psychological effect on students, and helps them to cultivate fulfilled personality. The introduction of outward bound is the reform and innovation of college P.E. and it's the construction and developing direction of college sports curriculum.

Acknowledgements. I would like to express my gratitude to all those who helped me during the writing of this thesis.

I owe a special debt of gratitude to all the professors and colleagues of P.E. College, Jiangnan University from whose devoted caring and helping.

References

1. Qian, Y.: *Outward Bound*, 2nd edn. Enterprise Management Press, Beijing (2007)
2. Yang, H., et al.: Research on College Sports Course - Outdoor Sports Teaching System. *Journal of Beijing Sports University* (6), 789-791 (2005)
3. Zhang, W., Zhou, H.: The Effect of Outward Bound Curriculum to Improve College Students' Ability of Cooperation. *Journal of Xi'an Sports College* (3), 113-115 (2006)
4. Zhang, J.: Research on the Setting of Outward Bound Curriculum in College Sports Teaching. *Journal of Wuhan Sports University* (10), 100-103 (2006)
5. Zhang, H., Liu, X.: The Related Factor Analysis on the Setting of College Outward Bound Curriculum. *Sports Adult Education Journal* (10), 75-76 (2006)
6. Zhang, F., Song, J.: Theoretical Analysis and Experimental Research on the Introducing of College Outward Bound. *Sichuan Sports Science* (6) (2005)
7. Han, T.: The Feasibility Analysis on the Opening of College Outward Bound Curriculum. *Journal of Beijing University* (12), 1694 (2008)
8. Wang, W., et al.: Analysis on the Implementation of College Outward Bound. *Education Academic Monthly* (5), 52-53 (2008)

The Effect of Guangzhou's Temperature Change to the Electric Power Consumption

Feng Jing and Pan Anding

School of geographical Sciences, Guangzhou University, Guangzhou 510006, China
panyun698@sina.com

Abstract. The temperature rising already affected people's life. We chose the model of three multinomial to fit tendency of electricity consumption, and we used the model of cooling degree-day to analysis the relation between Guangzhou's temperature changing recent 30 years and the electric power consumption. The result demonstrated that Guangzhou as an important city of electric power consumption, its annual mean temperature near 30 year presented a fluctuated improving trend. After 1997, Guangzhou's temperature rose obviously; it might have strengthened the intensity that people need to decrease living temperature. The electric consumption of meteorological influence in all departments presented a fluctuated improving trend. The data from 2004 to 2009 suggests that April to September was a high-temperature period, people need to decrease temperature, intensity that electric consumption of meteorological influence has a great fluctuation .There was a significant correlation between meteorology electric quantity in the industry and high seasonal temperatures. Meteorology electric quantity of the urban residents living was empirically related to monthly average temperature at different seasons.

Keywords: temperature change, electric power consumption, Guangzhou.

1 Introduction

Climate warming and temperature rising have been the fact that does not dispute. The global average temperatures will rise 1.4 ~ 5.8 Celsius degree in 21 century [1]. Research shows that the temperature changing affects the demand of electricity consumption of residents that is mainly heating and cooling, electricity consumption for cooling degree-day will changes 8 kilowatt-hours per person per year when temperature changes 1 Celsius degree[2]. Temperature changing of Guangzhou greatly influenced the consumption of electricity. Urbanization, expanding energy-intensive industries and improving energy efficiency limited capacity and resources are main factors that lead to increasing of the energy demand rapidly, this makes modeling and predicting power consumption become more and more important [3]. Based on the model of cooling degree-day, we analysed the correlation between interannual variation and monthly variation of the temperature and the influence of electricity consumption.

2 Data and Methods

Data. In this article, we choice the data of annual average temperature, average temperature of per month, the average daily temperature, the high (low) daily temperature of Guangzhou from 1980 to 2009 and total electricity consumption of Guangzhou from 1980 to 2008 and electricity consumption of industry and urban and rural residents of Guangzhou from 1990 to 2007 and electricity consumption per month of Guangzhou from 2004 to 2009 for statistical data analysis.

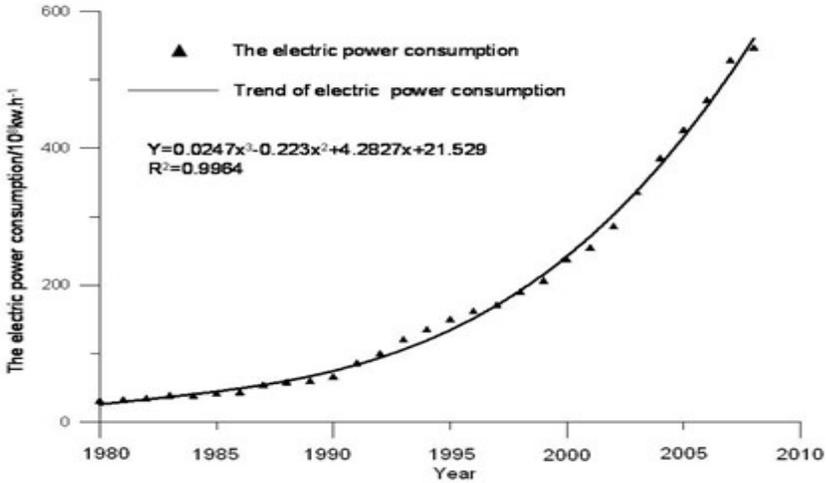


Fig. 1. The dynamics of power consumption in Guangzhou from 1980 to 2008

Model of Meteorological Electricity Consumption. There are many influencing factors of long-term changing of electricity consumption, one of the main influencing factors is social and economic development, and it represents the main trend of electricity consumption [4]. Meteorological factors make power consumption will produce fluctuations. Therefore, we first need to extract electricity consumption that meteorological factors influence, if we want to research the temperature changes on power consumption. Fig.1 shows the trend of actual electricity consumption was extracted by fitting cubic polynomial, according to using the least square method. We can ignore other random factors and establish the model of the actual electricity consumption for the decomposition.

$$y_w = y - y_i \tag{1}$$

y is actual electricity consumption, y_i is the trend of actual electricity consumption, y_w is the meteorological electric quantity that temperature influence[5].



Model of Cooling Degree-Day. Most analog power demand is function of seasonal climate factor. It uses two kinds of independent variables to develop different electricity statistical models. Primitive variables such as temperature, precipitation, relative humidity, wind speed, atmospheric pressure and solar radiation; and derived variables including heating degree-days (HDD), cooling degree-days (CDD).and enthalpy latent days. Most models show that the key weather variable is the outdoor air temperature both in its primitive and derived (HDD, CDD) forms [6]. Electricity demand is particularly sensitive to temperature changing [7]; especially high temperature of the summer on electric consumption has significant effects. So the model of cooling degree-days was used when we researched the electricity consumption that temperature changing influence. Computation formula of cooling degree-days is as follows:

$$T_{ci} = T_i - T_b, \quad D_c = \sum T_{ci} \quad (2)$$

T_{ci} is numerical value of cooling degree-days of the day i , T_b is basic temperature, T_i is average daily temperature of the day i , D_c is Annual accumulative value of cooling degree-days. Basic temperature is different in different countries. We adapt 25 as numerical value of basic temperature, because it was used as basic temperature of the models of cooling degree-days [8].

Results and Analysis.

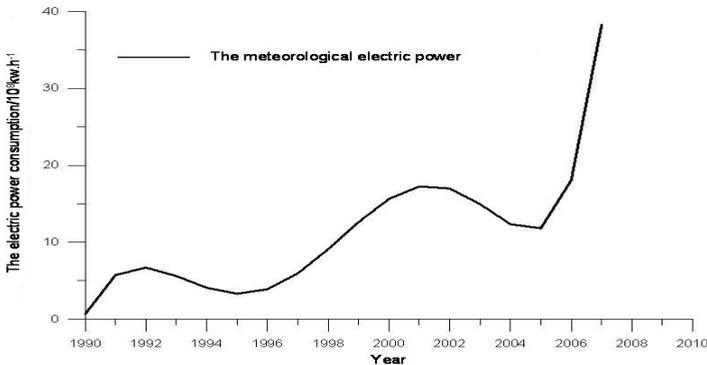


Fig. 2. The dynamics of the meteorological electricity in Guangzhou

Interannual Variations of Cooling Degree-day and Meteorological Electricity.

Fig.2 shows that meteorological power presented a fluctuating upward trend, and was divided into three periods: 1990-1996, 1996-2005, and 2005-2008.

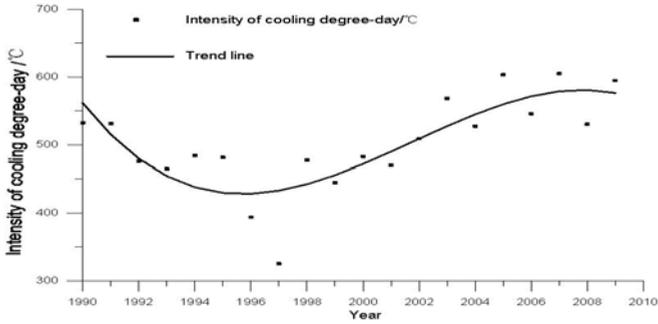


Fig. 3. The dynamics of cooling degree-day in Guangzhou from 1990 to 2009

Fig.3 shows that numerical value of cooling degree-day presented the downward trend from 1990 to 1996 and an up trend from 1996 to 2008. It reveals that variation of temperature and meteorological electric power consumption was significant correlation after 1996. Especially numerical value of cooling degree-day and meteorological electric power rose sharply after 2005.

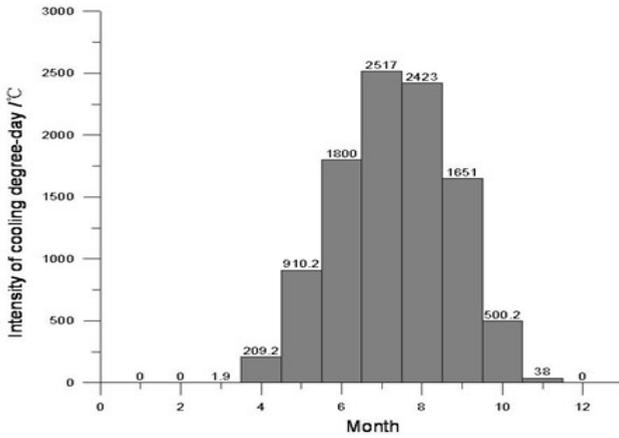


Fig. 4. The dynamics of cooling degree-day in Guangzhou from 1990 to 2009

Responses of the Electric Power Consumption to Monthly air temperature Change. Fig.4 shows that intensity of cooling degree-day began to increase from June to September, and it began to decrease in October. There is a period of high temperature from May to September in Guangzhou. It has great probability in July and August that frequent of high temperature appears.



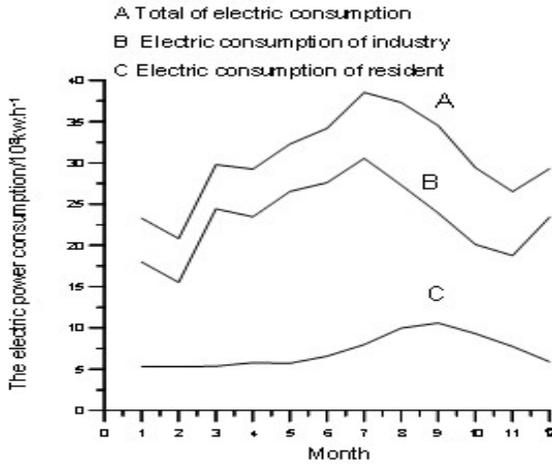


Fig. 5. The dynamics of power consumption in Guangzhou

Fig. 5 shows that electric consumption was maximal from June to September. Industrial electric power consumption was the main, and it began to increase in February, and achieved maximum in July, began to decrease in August. Electric consumption of residents' living achieved maximum in September, and it didn't change significantly in other months.

Correlation Analysis of Temperature and Electricity Consumption. We chose the data of average temperature of per month, the high (low) daily temperature of per month, and the data of industrial electric power consumption and electric consumption of residents' living of Guangzhou from 2004 to 2009 to make correlation analysis. The result shows that industrial meteorology electricity were strongly correlated with average temperature and the highest-temperatures from April to September, and it was not significant With the relevance of extreme high temperature, especially the correlation coefficient with average temperature of September reached 0.94, significant level achieved 0.05. This showed that temperature changing had stronger influence with industrial meteorology power consumption in late spring and summer and autumn. Meteorology electricity of residents' living was strongly correlated with the highest-temperatures of February, May, July, and September.

3 Conclusions

The dynamics of meteorological power of Guangzhou was divided into three periods: 1990-1996, 1996-2005 and 2005-2008. Because of restrictive factor of the social and economic development, it was different from the dynamics of cooling degree-day.

Electricity consumption was maximal from June to September. Meteorology electricity was strongly correlated with average temperature in Guangzhou, and it was not significant with the relevance of extreme high temperature.

Industrial meteorology electricity was strongly correlated with average temperature and the highest-temperatures from April to September. Meteorology electricity of residents' living was strongly correlated with the highest-temperatures of February, May, July, and September.

References

1. Liu, C., Chen, A., Ba, T., Meng, C., Yu, L.: Global Climate Changes from IPCC AR4 and Corresponding Countermeasures. *J. of Institute of Disaster-Prevention Science and Technology* 10, 140 (2008)
2. Mideksa, T.K., Bekken, S.K.: The impact of climate change on the electricity market. A review. *Energy Policy* 38, 3579–3580 (2008)
3. Lian, L., Zhang, Y.: Relationship of Electricity Supply and Weather Conditions —Qufu in Shandong. *Energy Research and Utilization*, 17 (2008)
4. Yuan, S., Qian, H.: Indices and Models Assessing Climatic Impacts on Energy Consumption. *Resources Science* 26, 1–2 (2004)
5. Zhang, X., Wang, Y.: Relationship between Power Consumption and Meteorological Conditions during the Summer in Beijing City and Its Forecast. *Meteorology* 28, 17
6. Pardo, A., Meneu, V., Valor, E.: Temperature and seasonality influences on Spanish electricity load. *Energy Economics* 24, 55–70 (2002)
7. Duan, H., Qian, H.: Responses of the Electric Power Consumption to Climate Change in Guangzhou City. *Journal of Applied Meteorological Science* 20, 81–82 (2009)
8. Wang, Z., Pan, A.: Characteristics and Damages of Summer High Temperature in Guangzhou City. *Journal of Meteorological Research and Application* 29, 26–29 (2008)

The Relationship and Strategy of Informatization on Farmers, Agriculture and Rural Areas and New Rural Construction

Li Quan-guo¹ and Kang Ling²

¹ Department of Geography, Xiangfan University, Xiangyang, China

² School of Management, Xiangfan University, Xiangyang, China

leesilee@sina.cn

Abstract. Informatization on farmers, agriculture and rural areas and new rural construction was an important subject on rural and agricultural development in China, and it was a long and arduous task. The connotation of informatization on farmers, agriculture and rural areas and new rural construction were described, the interaction on them were analyzed, the construction strategy of them were proposed, the purpose is the informatization of new countryside in China to provide reference.

Keywords: informatization on farmers, agriculture and rural areas, new rural construction, relationship, strategy.

1 Introduction

Informatization on farmers, agriculture and rural areas is not only the base and an important part of informatization on national economy but also the key to building a moderately prosperous society in China. Achieving informatization on farmers, agriculture and rural areas is quickly able to change the way and costs, in traditional agriculture, by which the farmers have access to technology information, and it also can broaden their horizons and increase the possibility of participating market activities, so informatization on farmers, agriculture and rural areas has become an important subject on rural and agricultural development in China, besides building a new socialist countryside is a long-term and arduous task to endeavor to carry on, therefore both have close connection. The paper conducted relevant researches on the development strategy and relationship between the informatization on farmers, agriculture and rural areas and the construction of new countryside in China, which can provide reference information for the informatization on new rural construction.

2 Informatization on Farmers, Agriculture and Rural Areas and New Rural Construction

Informatization on farmers, agriculture and rural areas is the full range of informatization about them, that is to say, informatization on farmers, agriculture and

rural areas is the extent and process for modern information technology to achieve universal application on rural production and management, public services, government's administration and other aspects of living consumption by way of strengthening the rural radio networks, telecommunication networks, computer networks and other information infrastructure which is to fully develop and utilize the information resources, build information service system and facilitate the exchange of information and the share of knowledge [1]. The content of informatization on rural areas should at least comprise the following six spheres: informatization on agricultural science and technology's education, informatization on rural resources and environment, informatization on rural life's consumption, informatization on rural productive management, informatization on rural government's management, and informatization on rural market's circulation [2].

Construction of new socialist countryside is a strategic decision which was addressed by the Party Central Committee to solve issues concerning farmers, agriculture and rural areas. In accordance with the requirements of building a new socialist countryside, for strengthening the construction of new socialist countryside it is essential to keep the use of modern technology to develop agricultural production, increase the income of farmers, ameliorate their life, improve their quality, promote the rural civilization, change its aspect, construct its environment and improve self-management ability. "productive development, affluent life, rural civilization, neat village and democratic management" the twenty-character policy is the basic requirement of new rural construction, and the application of information network technology accelerates to advance the rural economy construction, the political construction, the cultural reconstruction, the social construction and party's construction, and through planning the economic and society's development of the city and the countryside as a whole, it finally comes true for the coordinated development of the economic and society's development of the city and the countryside. Informatization on new rural construction regards the capacity of agricultural comprehensive production as the material foundation, in other words, it takes the economic construction as the center, the productive development as the primary task, and the capacity of agricultural comprehensive production as the emphasis of construction.

3 The Benign Interaction between Informatization on Farmers, Agriculture and Rural Areas and New Rural Construction

Informatization on farmers, agriculture and rural areas is an important content of the construction of new rural. As it is agriculture that is our industry with the lowest infomationized degree, it is farmers that are the hardest people to obtain information, it is countryside that is the place where information is the most occlusion of, information impeded seriously restricts the benign development of the construction of new countryside. Therefore, paying great attention to and accelerating informatization on farmers, agriculture and rural areas is an important content of new rural construction, simultaneously, it also is an effective measure to solve the outstanding problems in new rural construction, and an important way to promote the reform of rural economic system and achieve management and system's innovation. Informationization on new

rural areas is a complicated systematic project, whose essence is to apply information network technology to solve the issues on farmers, agriculture and rural areas, up to construct a new rural in the age of information. Informatization on agriculture and rural economy is the key to the construction of informatization on new rural areas, and the core of the construction is to achieve increasing agricultural production, farmer's income and the competitiveness of agricultural products by means of informatization. Informatization on new rural areas is the natural extension of informatization on city and the synthetical embodiment of the combination among current e-government, e-business, the social informatization, the exploitation and utilization of information resource and the characteristics of rural development. Informatization on farmers, agriculture and rural areas actually is informatization on the whole developing process of the rural economy and the development of social causes, and it takes information technology as the support to equip modern rural industries, takes digital countryside as the core and main manifestation to sustain rural management, takes information service system as the prop to exploit and utilize various agricultural resources more effectively and more reasonably, upgrade rural modernization level and promote rural economic development and social progress, which is an important content of the construction of new countryside.

Informatization on farmers, agriculture and rural areas is the inevitable requirement of new rural construction. Strengthening informatization on farmers, agriculture and rural areas is not only an important breakthrough and an inevitable demand in the construction of new countryside but also a starting point to solve the issues concerning farmers, agriculture and rural areas, and without it the requirement of the policy on new rural construction can not go, for the reason that if we want to transform and upgrade traditional agriculture quickly and effectively, and promote the process of agricultural industrialization and modernization, it is indispensable to apply modern information technology widely before, during and after agricultural harvest for each link to achieve "productive development" and attain "affluent life". Through the support of informatization on farmers, agriculture and rural areas, it can accelerate the process about which modern science and civilization and modern lifestyle penetrate in rural areas and impact it, besides it is also able to give impetus to rural areas to change customs, promote the spiritual cultural construction of countryside and accelerate the realization of "rural civilization", thereby achieving "neat village". Through the information technology it is realized to make village's affairs public, collect public opinions and display village's looks, which enables the out-villager to care and intervene village's management effectively, uphold rural stability and promote rural construction and management's democratization, thereby realizing "democratic administration" [3]. New rural construction should not only satisfy the twenty-character policy, but also make estimates about the agricultural natural resources, the ecology and environment, the meteorological and biological disasters, the cultivation and production of crops to achieve accurate and dynamic monitoring and real-time forecasts. However, informatization on farmers, agriculture and rural areas just provides the technical guarantee for this series of work.

The new rural construction is the ascription and the goal of the informatization on farmers, agriculture and rural areas. The new rural construction is actually a comprehensive project, it is either to construct new rural areas or to cultivate the new farmers, what's more, it also needs to develop new agriculture, which is the ascription

and the goal of the informatization on farmers, agriculture and rural areas. The new rural construction speeds up the process of the informatization on farmers, agriculture and rural areas, and it's also the concrete embodiment of it. The new rural construction aims at developing the public information service system completely, and makes further construction on agriculture's informatization. The new rural construction reforms the traditional agriculture to modern agriculture by means of information technology, links up the farmer's production and the market through the information service, lessens the differences between the city and the countryside on information by ways of improving the level of informatization, which is the urgent task of constructing the harmonious society and the new socialist countryside.

The new rural construction is the guarantee and the foundation of the informatization on farmers, agriculture and rural areas. The informatization on farmers, agriculture and rural areas requires not only good infrastructural facilities but also the great support and guidance from the government, and it still needs the reserve and accumulation of information-based talents. What's more, it needs the integration of all kinds of agriculture information resources. All these things can't do without a sound and harmonious environment on new rural construction. The new rural construction propels the informatization on farmers, agriculture and rural areas forward, totally speeds up the process of realizing it, and is the guarantee and the foundation of it.

4 The Research Strategies on the Informatization on Farmers, Agriculture and Rural Areas and the New Rural Construction

Use the successful experience of other countries for reference. In promoting the process on the informatization on farmers, agriculture and rural areas, other countries in the world have different features and they have different developmental models. The United States and Japan have perfect legislation and the good system on the informatization of the countryside, and the fund on investment is guaranteed. Countries such as Japan and Germany have perfect information installations in the countryside. They pay much attention to the construction of information system. France, Canada and others form a multi-level pattern on the agricultural information service, and the subjects of service vary. South Korea, India and other governments invest on the construction of the information installations. They establish the preferential policies on rural information service, pay great attention to the talents' training on rural informatization and the international cooperation. It is necessary to learn experience from countries which have done successfully on the rural informatization and new rural construction, and absorb the essence, in order to build a good developing environment for the informatization on new rural areas.

Strengthen the government's support. The government, as the manager and public organizer of the new rural informatization construction, links the service on information, the customer demands and providers of the policy together tightly. The government not only extends the service's scope to share the information service, but also provides the practical information in time, so that the information needs of the majority of rural users can be met, and therefore the construction on the rural informatization will be promoted. The government's support on the informatization on farmers, agriculture and rural areas and the new rural construction shows on two

aspects. On one hand, the government provides the fund for them. On the other hand, the government makes the rural informatization sustained develop on its own. The government, as the leading department, should base on the actual farmers' demands to develop all kinds of information products which are cheap, convenient and easy-adopted, and also provide the agricultural science and technology knowledge for the agricultural production, the farmers' lives and the rural economic development, besides they should train farmers' professional skills, which in the end provide the inexhaustible motivation for the new rural construction.

Strengthen the construction of information infrastructures. Most of our rural information infrastructures are quite weak, so we have to strengthen information infrastructures in order to realize the informatization on farmers, agriculture and rural areas and the new rural construction. Hardware is mainly about the construction of the basic network facilities and software is mainly the development and perfection of the agricultural information resource database and information search [4]. Concretely we may take all levels of agriculture department as a support on the base of information infrastructure foundation to build a central-province-city-county information backbone high speed wide network system, namely taking the provincial level agriculture message center as the core to link with the national agricultural information network, taking the around (city) as the hub and the county (city) as the net point. They have a unified data specification and shared standards, and seamless connectivity, which make it possible to attain roaming and zooming freely. We should make a scalable and comprehensive agricultural information network system with perfect function and outstanding performance, and link it to other net systems to form an all-round agricultural resources and economic information network system and then open up the digital agriculture market as the national agricultural service platform.

Strengthen the construction of information talents. The key of informatization on farmers, agriculture and rural areas is the farmer, and the fundamental cause for the rapid popularization on it is that the farmers don't have much knowledge. At present, many peasants don't possess the ability of operating computer. Thus, improving farmers' scientific and cultural quality, eliminating the mysterious sense for the computer are the key to solve the issues of the informatization on farmers, agriculture and rural areas and the new rural construction. On the one hand, it is necessary to strengthen the existing agricultural information network professionals' training to improve their business capability so that they can do them better on the information officers' inspection and the qualification, on the other hand, pay attention to the change of the concept on talents, not only students and technical workers are talents, but the average farmers too [5]. We should strengthen training the technical staffs, at the same time, we also should strengthen training the average farmers, who just are the main force in the countryside. Therefore we should hold various forms of farmers' training to improve people's information technology level, so that they can serve the new rural reconstruction better.

Create new demonstrations about the informatization on new rural areas. Creating new demonstrations about the informatization on new rural areas can be achieved by selecting individual rural areas whose information infrastructures are relatively completed, and can be mature gradually through a period of centralized construction. And they can reflect the advantage of agriculture informatization gradually, thereby being the pioneer of the informatization on farmers, agriculture and rural areas and the new rural construction and playing an exemplary role.

Establish comprehensive service platforms about "digital agriculture". "Digital agriculture" mainly includes: agricultural production management digitization, agricultural production management and service digitization, the rural macro management digitization, the farmers' education digitization, agricultural resources management digitization, etc. In order to promote rural industry development, enhance grass-roots management efficiency and strengthen government social service capacities by information technology, we can form a new countryside "Digital Agriculture" integrated information service platform for the rural grass-roots management departments, enterprises and residents through geospatial information, intelligence information processing, multimedia and virtual reality technology integration, and fully capture and management all kinds of basic information resources of rural, improve resource allocation efficiency, integrate rural economy and regional planning, serve farmers.

5 Summary

The informatization on farmers, agriculture and rural areas and the new rural construction promote each other, and are integral whole. The advancement of the informatization on farmers, agriculture and rural areas in the new rural construction plays a very positive role. In the new rural production development, the informatization on farmers, agriculture and rural areas plays a leading role in the market. It can build a platform for agricultural production, make agricultural instructions, solve the contradiction between the rural small-scale production and the large market, and also settle the agricultural production's low degree of organization and selling. What is more, it can impel the market scale expanding and improve the market efficiency, promote the formation and the development of information industry, form a new information economy and promote the development of the new rural construction. In turn, the new rural construction feeds the informatization on farmers, agriculture and rural areas, and provides a good development environment for it.

Acknowledgements. This work was supported by the Humanities and Social Sciences Project of Hubei Provincial Department of Education (2010b223), Research Projects of Xiangfan University (2009YB007 and 2009YB013).

References

1. Li, D.: Development report of rural information in China, pp. 19–21. Electronic Industry Press, Beijing (2008)
2. Liang, J.: The Current, Content and Countermeasure of New Rural Area Informationization. *Journal of Nanjing Agricultural University (Social Sciences Edition)* 9(3), 25–27 (2006)
3. Wang, Q., Huang, S.: New socialist rural construction and the rural informationization leapfrog development. *Rural Economy* (7), 5–8 (2007)
4. Qiu, S.: Promoting the construction informationization on farmers, agriculture and rural areas: a realistic approach to a new socialist countryside. *Journal of Fujian Party School* (10), 67–71 (2006)
5. Li, Z.: Research on the conditions and development strategies of Beijing new countryside informationization. Chinese Academy of Agricultural Sciences Dissertation, Beijing (2009)

Measuring the Efficiency of Rural Informatization in China 2007-2009: An Application of Data Envelopment Analysis*

Xiaochun Xu** and Jingdong Luan

College of Economics and Management, Anhui Agricultural University,
130 West Changjiang Road, Hefei, Anhui, China
chunchun474@yeah.com

Abstract. Coordinating development between urban and rural areas is the fundamental requirement of China's modernization, rural informatization is of great importance in the process. In this study, the Data Envelopment Analysis (DEA) method was applied to evaluate the relative efficiencies of rural informatization development in China. Empirical results indicate that rural informatization in the east, middle and west China shows different patterns of inefficiency. Suggestions for improving development efficiencies are also discussed.

Keywords: Data Envelopment Analysis (DEA), Rural Informatization, Efficiency, China.

1 Introduction

By the end of the 20th century, China faced seriously challenges in rural development, such as food security, rural poverty, as well as widening urban-rural inequalities. For improving the overall production capacity of agriculture, constructing a new socialist countryside, developing modern agriculture, achieving sustained income increases for farmers and building a moderately prosperous society, since the early 2000s, a new set of rural policies have been implemented to coordinate development between urban and rural areas, accentuating in particular urban-rural integration. Promoting the rural informatization is part of these policies. Since 2005, many informatization programs have been initiated, such as Three in One Agriculture Information Services Project (2005), Extend Telephone Coverage to Every Village Project (2005), Rural Comprehensive Agriculture Informatization Service Pilot Project (2006), especially Golden Agriculture Project (2007), government departments at all levels (from central to local), domestic and multinational enterprises are involved in these projects.

Despite a lot of difficulties including human resource shortages, technological constraints, limited funding, some of the projects have produced important

* This study was supported by AHU social sciences project (2010zs01zd), Department of Education of Anhui Province (2009sk120), and the national science and technology support program (2008BAD96B12).

** Corresponding author.

achievements. For example, the urban-rural Digital Divide Index (DDI) was 0.64 in the year 2007, 0.59 in 2008 and 0.56 in 2009 (China Information Almanac 2010), which means that from 2007 to 2009, the average information and communication technologies (ICTs) usage level in rural areas was 64%, 59% and 56% lower than that in urban areas, the urban-rural digital divide has remarkably decreased.

With increasing amount of resources allocating in rural informatization, more questions arising from the progress, could rural informatization eventually generate reasonable returns? Is the development sustainable? Given the scarcity of resources, it is important for socioeconomic growth to depend on production efficiency improvement to achieve sustainability. However, available studies have not provided satisfactory answers. This study applies DEA method to evaluate rural informatization development efficiency during the period of 2007-2009, to our knowledge, no national efficiency evaluation work has been presented, and this will help making the rural informatization development more efficient and sustainable.

2 Methodology and Model Specification

DEA is one of the most common techniques used in the analysis of efficiency. To keep this paper short, only a general description of DEA is given here, without mathematical background.

DEA is a non-parametric approach for measuring the relative efficiency of decision making units (DMUs), originally developed by Farrell (1957), introduced by Charnes, Cooper, and Rhodes in 1978 and further elaborated by Banker, Charnes and Cooper in 1984.

According to Farrell (1957), the concept of efficiency, might be characterized as the ability of a DMU to achieve maximum output given a set of inputs, or, alternatively, to consume minimum input given a set of outputs. An efficient DMU is one which achieves a maximum amount of outputs while utilizing a minimum amount of inputs, through the envelopment of the most efficient DMUs, a relative efficiency frontier are constructed, the rest DMUs are deemed inefficient and the distance of their input-output ratio from the frontier is the degree of their inefficiency.

In this paper, DEA-CCR (due to Charnes, Cooper and Rhodes, 1978) and DEA-BCC model (due to Banker, Charnes and Cooper, 1984) are used to evaluate technical efficiency (TE, also termed overall technical efficiency) and pure technical efficiency (PTE), then scale efficiency (SE) can be calculated.

The choice between input- and output- oriented is also a matter of concern. The input-oriented model aims at reducing the input amounts while keeping the present output levels, and the output-oriented model, aims at maximizing output amounts under the present input levels. Rural informatization in China has been identified as one key to achieving sustainable socioeconomic developments in the future, the present political and policy environment has provided excellent opportunities for rural informatization, no doubt that more investments of financial, governmental and technological resources from all concerned will be attracted. Hence, from a practical standpoint, it is more appropriate to use output-oriented DEA models which essentially address the question "By how much can rural informatization output quantities be proportionally increased without wasting input resources?"

3 Data and Variables

The input and output variables should reflect actual rural informatization development status as accurately as possible. Also, consider a rough rule of thumb in the envelopment model, suppose there are m input variables and s output variables, the number of DMUs (n) should be equal to or greater than $\max\{m \times s, 3 \times (m + s)\}$ (Cooper,2007), it is crucial for successful application of DEA. In this study, 31 provinces in China, which covers the whole country, are specified as DMUs, given the constraints of the available data, five input and four output variables are considered for the estimation of the DEA model.

In the traditional growth theory, factor accumulations (labor, capital, etc) are considered as the main driving forces of economic growth and social development. Five input variables are specified in this study, they are per capita consumption expenditure for transport, communications, education, cultural, recreation and services of rural households (I1), number of employed persons in telecommunications and other information transmission services (I2), government expenditure for education, culture, sport and media (I3), information, transmission, computer, services and software investment in fixed assets in the whole country(I4), and, rural household consumption expenditure(I5).

On the output side, referring to the International Telecommunication Union (ITU) ICT Development Index, four relevant output variables are specified to benchmark and assess rural informatization developments, they are number of internet users (O1), fixed and mobile telephones per 100 rural households (O2), TV sets per 100 rural households (O3), computers per 100 rural households (O4).

All data are obtained from the China Statistical Yearbook. The summary statistics for the input and output variables are presented in Table 1.

Table 1. Descriptive Statistics for Input and Output Variables

| Variable | Max | Min | Average | S.D. |
|--------------------|---------|--------|----------|----------|
| I1(Yuan) | 2155.14 | 209.47 | 729.49 | 413.14 |
| I2(Person) | 110581 | 2502 | 33926.91 | 20714.29 |
| I3(100millionyuan) | 914.71 | 40.8 | 301.94 | 174.81 |
| I4(100millionyuan) | 281.6 | 3.01 | 70.96 | 51.43 |
| I5(Yuan) | 13748 | 1895 | 4274.07 | 2286.06 |
| O1(10000persons) | 4860 | 36 | 959.43 | 854.84 |
| O2(Unit) | 320.66 | 44.58 | 176.62 | 52.95 |
| O3(Set) | 200.67 | 55.42 | 115.18 | 22.91 |
| O4(Set) | 54.33 | 0 | 7.44 | 11.91 |

Data source: China Statistical Yearbook (2008-2010)

4 Results and Discussion

Using the data which consisting of 31 provinces in China over a period of 3 years, both the output-oriented CCR and BCC models of DEA are applied to evaluate rural informatization efficiencies, the results are summarized in Table 2 and Table 3.

Table 2. Average DEA efficiencies 2007-2009

| Region | 2007 | | | 2008 | | | 2009 | | |
|-------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | TE | PTE | SE | TE | PTE | SE | TE | PTE | SE |
| East China Average | 0.975 | 0.989 | 0.985 | 0.969 | 0.984 | 0.983 | 0.967 | 0.986 | 0.980 |
| Middle China Average | 0.906 | 0.960 | 0.943 | 0.903 | 0.933 | 0.967 | 0.919 | 0.942 | 0.975 |
| West China Average | 0.949 | 0.970 | 0.976 | 0.961 | 0.971 | 0.987 | 0.951 | 0.966 | 0.982 |
| National Average | 0.943 | 0.972 | 0.968 | 0.944 | 0.962 | 0.979 | 0.945 | 0.964 | 0.979 |

Notes: TE - technical efficiency. PTE - pure technical efficiency. SE - scale efficiency.

It can be seen from Table 2 and Table 3 that during the period 2007-2009, the national average technical efficiency, scale efficiency and pure technical efficiency scores are 0.943, 0.944, 0.945(TE), 0.968, 0.979, 0.979(SE) and 0.972, 0.962, 0.964 (PTE) respectively. The national average TE score is about 0.945, which means that inputs almost can be reduced by 5.5 percent while retaining the same output levels. The national average technical efficiency was lowest (0.943) in 2007 and then an increasing trend can be observed during the study period when it reached highest (0.945) in 2009, which is positive sign indicating a substantial improvement. The results also show that the pure technical efficiency (PTE) is on average lower than the scale efficiency (SE) in 2008 and 2009, which means the main source of inefficiency is from the pure technical aspect, in other words, the scale efficiency contributes more than pure technical efficiency towards technical efficiency.

About one third of the DMUs (Beijing, Tianjin, Fujian, Shandong, Guangdong, Henan, Hainan, Guangxi, Guizhou, Shaanxi, Gansu, Qinghai and Ningxia) are relatively efficient every year. There are some DMUs (Jiangsu in 2008 and 2009, Zhejiang in 2007, Xizang in 2007, Chongqing in 2009 and Xinjiang in 2008), which have TE scores less than 1 but PTE scores equal to 1, their inefficiency comes mainly from the scale inefficiency, the scale of the rural informatization system should be adjusted in these provinces.

Furthermore, all 31 DMUs are divided into three regions: east, middle and west China, as shown in Table 3. The reason of this grouping is that the development levels, resource endowments, subsidy policies are different in these three regions. The three-year average TE scores of east, middle and west China are 0.970, 0.909 and 0.954, PTE scores are 0.986, 0.945, and 0.969, SE scores are 0.983, 0.962 and 0.981. It is interesting to find that the DMUs in middle China are more developed

Table 3. DEA efficiencies of 31 provinces in China 2007-2009

| Region | DMU | 2007 | | | 2008 | | | 2009 | | |
|--------------|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | | TE | PTE | SE | TE | PTE | SE | TE | PTE | SE |
| East China | Beijing | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Tianjin | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Liaoning | 0.815 | 0.902 | 0.904 | 0.759 | 0.854 | 0.889 | 0.846 | 0.874 | 0.969 |
| | Shanghai | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Jiangsu | 1 | 1 | 1 | 0.962 | 1 | 0.962 | 0.853 | 1 | 0.853 |
| | Zhejiang | 0.964 | 1 | 0.964 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Fujian | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Shandong | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Guangdong | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Middle China | Hebei | 0.861 | 0.988 | 0.871 | 0.952 | 0.998 | 0.954 | 1 | 1 | 1 |
| | Shanxi | 0.906 | 0.960 | 0.943 | 0.912 | 0.939 | 0.971 | 0.913 | 0.921 | 0.991 |
| | Jilin | 0.914 | 0.991 | 0.923 | 0.847 | 0.935 | 0.906 | 0.796 | 0.892 | 0.893 |
| | Heilongjiang | 0.871 | 0.922 | 0.944 | 0.842 | 0.865 | 0.973 | 0.852 | 0.853 | 0.999 |
| | Anhui | 0.952 | 0.982 | 0.969 | 0.923 | 0.933 | 0.989 | 0.852 | 0.905 | 0.942 |
| | Jiangxi | 0.910 | 0.965 | 0.943 | 0.893 | 0.904 | 0.987 | 0.939 | 0.970 | 0.968 |
| | Henan | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Hubei | 0.876 | 0.919 | 0.952 | 0.877 | 0.899 | 0.976 | 0.985 | 0.994 | 0.991 |
| | Hunan | 0.770 | 0.868 | 0.887 | 0.784 | 0.858 | 0.914 | 0.855 | 0.881 | 0.970 |
| Hainan | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | |
| West China | Neimenggu | 0.682 | 0.811 | 0.841 | 0.675 | 0.778 | 0.868 | 0.661 | 0.767 | 0.862 |
| | Guangxi | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Chongqing | 1 | 1 | 1 | 0.939 | 0.954 | 0.984 | 0.978 | 1 | 0.978 |
| | Sichuan | 1 | 1 | 1 | 1 | 1 | 1 | 0.955 | 0.999 | 0.956 |
| | Guizhou | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Yunnan | 0.822 | 0.846 | 0.972 | 0.920 | 0.925 | 0.995 | 0.851 | 0.857 | 0.993 |
| | Xizang | 0.925 | 1 | 0.925 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Shaanxi | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Gansu | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Qinghai | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Ningxia | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | Xinjiang | 0.956 | 0.987 | 0.968 | 0.998 | 1 | 0.998 | 0.965 | 0.973 | 0.991 |

Notes: TE - technical efficiency. PTE - pure technical efficiency. SE - scale efficiency.

than those in west China, however they are with the lowest average TE scores. The difference can be explained by the structures and the amounts of inputs and outputs of these regions, the different financial and other subsidy policies, especially the different rural informatization development bases and stages. As China's most economically developed regions, the DMUs in east China have the highest average TE, PTE and SE

scores, but all the three scores are obviously trending downward, almost 1 percent annual decline during the study period.

In DEA models, technically speaking, an inefficient DMU can become efficient by proportionally improve the values of controllable input or output variables, either reducing its use of inputs, or increasing its outputs, or changing both its inputs and outputs, until the efficient frontier is achieved. Table 4 shows the projections of the technical inefficient DMUs to the efficiency frontier, i.e., the adjustment strategies for the technical inefficient DMUs to be efficient.

It is observed that outputs can be averagely increased from 14.84%(O1) to 63.72% (O4) further. Although, DEA model is run in output maximization mode, still 2.92%(I3), 5.72%(I1), 11.10%(I2) and 15.84%(I4) average reductions were proposed for inputs. This implies that even for DMUs that are projected to efficient frontier, still there is excess usage of inputs. Reduce investment in fixed assets (I4), maintain a reasonable investment scale, optimize investment structures, will contribute directly to improve the efficiency for all DMUs. In middle and west China, change the status quo of high employment with low labor productivity (I2) in the related service sectors, is also an important efficiency improvement measures.

Table 4. Projections of Technical Inefficient DMUs to be Efficient

| Inefficient DMUs | | I1 | I2 | I3 | I4 | I5 | O1 | O2 | O3 | O4 |
|------------------|--------------|--------|--------|--------|--------|-------|-------|-------|-------|--------|
| East China | Liaoning | -9.62 | 0 | -2.50 | -19.85 | 0 | 24.21 | 25.77 | 29.57 | 52.88 |
| | Jiangsu | -12.25 | 0 | 0 | -5.37 | -0.65 | 7.07 | 15.40 | 7.07 | 29.75 |
| | Zhejiang | -6.81 | 0 | 0 | -7.94 | 0 | 1.24 | 1.24 | 1.60 | 1.24 |
| | Average | -9.56 | 0 | -0.83 | -11.05 | -0.22 | 10.84 | 14.14 | 12.75 | 27.96 |
| Central China | Shanxi | -17.68 | -28.66 | 0 | -22.61 | 0 | 9.86 | 14.77 | 9.86 | 9.86 |
| | Jilin | -3.32 | -22.43 | -1.23 | -7.21 | 0 | 17.65 | 17.65 | 31.09 | 31.44 |
| | Heilongjiang | -10.70 | -33.18 | 0 | -54.28 | 0 | 26.67 | 17.00 | 17.00 | 17.00 |
| | Anhui | -0.38 | -0.56 | -11.27 | -11.12 | 0 | 19.77 | 10.23 | 14.55 | 21.54 |
| | Jiangxi | 0 | -7.02 | 0 | -10.63 | -8.29 | 9.47 | 10.68 | 9.47 | 16.79 |
| | Hubei | 0 | -3.30 | 0 | 0 | -5.73 | 9.92 | 10.02 | 11.94 | 22.41 |
| | Hunan | 0 | -4.61 | 0 | -4.94 | -1.29 | 24.80 | 37.38 | 24.80 | 144.10 |
| Average | -4.58 | -14.25 | -1.79 | -15.83 | -2.19 | 16.88 | 16.82 | 16.96 | 37.59 | |
| West China | Neimenggu | -23.07 | -38.24 | 0 | -21.21 | 0 | 48.63 | 51.67 | 48.63 | 176.07 |
| | Chongqing | 0 | -7.43 | 0 | -10.08 | -9.67 | 2.92 | 5.09 | 2.92 | 81.82 |
| | Sichuan | 0 | -2.31 | -0.47 | -3.78 | -1.61 | 1.58 | 1.58 | 1.58 | 9.82 |
| | Yunnan | 0 | -18.72 | -4.40 | -24.69 | -0.13 | 15.94 | 19.90 | 15.94 | 159.92 |
| | Xinjiang | -2.05 | 0 | -23.90 | -33.84 | 0 | 2.86 | 5.198 | 2.86 | 18.119 |
| Average | -5.02 | -3.34 | -5.75 | -8.72 | -2.28 | 4.39 | 6.04 | 4.39 | 1.76 | |
| Average | -5.72 | -11.10 | -2.92 | -15.84 | -1.82 | 10.84 | 14.14 | 12.75 | 27.96 | |

Note: All scores are in percentage.



Drastically strengthening the development of infrastructure and basic facilities in rural areas, increasing the coverage of the communications networks, especially increasing the access to computers and internet, can be a key enabler for improving the rural informatization efficiency.

5 Conclusions

In this paper, a non-parametric data envelopment analysis (DEA) technique was applied to evaluate the rural informatization technical, pure technical and scale efficiency of 31 province-level units in China from 2007 to 2009. Empirical results show that, during this period, DMUs in east China obtains the highest DEA efficiency scores, this implies that the economic development levels are playing the most important roles in the implementations of rural informatization. However, compare to other regions, there is a clear decline in efficiency scores in east China. Furthermore, the projections of input and output variables that obtained from the DEA models indicate that reduce investment in fixed assets to a reasonable level, and greatly improve the access to computers and internet, would have significantly positive effects on the performance of rural informatization.

References

1. Cooper, W.W., Seiford, L.M., Tone, K.: Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software. Kluwer Academic Publishers (2000)
2. Cooper, W.W., Seiford, L.M., Zhu, J.: Handbook on Data Envelopment Analysis. Kluwer Academic Publishers (2004)
3. Qiang, C.Z.-W., Bhavnani, A., Hanna, N.K., Kimura, K., Sudan, R.: Rural Informatization in China. World Bank Publications (2009)
4. Chang, P.-L., Hwang, S.-N., Cheng, W.-Y.: Using Data Envelopment Analysis to Measure the Achievement and Change of Regional Development in Taiwan. *Journal of Environmental Management* 43, 49 (1995)
5. Cooper, W.W., Seiford, L.M., Zhu, J.: Slacks and congestion: a response to the comments of R.Fare and S. Grosskopf. *Socio-Economic Planning Sciences* 35, 205 (2001)
6. Chen, L.: Building the Scientific and Technological Information Network to Cover Rural China: The Current Situation and Trend in Future. *Issues In Agricultural Economy* 10, 56 (2004)
7. Liang, J.: The Current, Content and Countermeasure of New Rural Area Informationization. *Journal Of Nanjing Agricultural University (Social Sciences Edition)* 6-3, 25 (2006)
8. Chen, Y.: Analysis of the way of Span the "Last One Kilometer" of Agricultural Informatization. *Chinese Agricultural Science Bulletin* 9, 462 (2006)
9. Duan, Q., Yang, R., Feng, X.: Study on Information into Family in the Agricultural Informatization Construction. *Journal of Agricultural Mechanization Research* 7, 231 (2007)
10. Liu, S., Xu, S.: Study on Evaluation/Measurement Method for China's Rural Informatization. *Scientia Agricultura Sinica* 41-4, 1012 (2008)

11. Yu, Y., Du, Y., Peng, J.: Present Situation and Countermeasure of Rural Areas Informationization Construction in Western—Investigation and Analysis on 10 Counties of 5 Provinces/Municipality in Western. *Rural Economy* 25-4, 20 (2009)
12. Yang, C.: Study on the Level and Development of China Rural Informationization—Based on Second National Agricultural Census Data. *Library And Information* 1, 85 (2010)
13. Yang, Y., Tong, X.: Rural Agriculture Informatization Measurement in the Yellow River Delta Areas. *Shandong Agricultural Sciences* 9, 15 (2010)

Research on Software Architecture Base on Data Snap Technology

Dan Guo and Yu Zhang

School of Computer and Information Engineering, Harbin University of Commerce, Harbin, China
guozhang_123@tom.com

Abstract. The paper discussed a software architecture based on Data Snap technology, which is a multi-tiered distributed system's structure by technology of C/S .Using the software architecture , system can improve the software reusability, meanwhile, can solve the system load overweight cause by data access operations.

Keywords: software architecture, Data Snap technology, 3-tier application.

1 Introduction

Data Snap technology, which was introduced into Borland Delphi is a multi-layer physical structure technology. Before Delphi version 6.0 this technique is called "Midas", Multi-tier distributed application services suite. From the Delphi 6.0 began renamed Data Snap. Namely, the application's data module transplanted into a dedicated server (application server), and make all client program and the data module interaction. In the physical structure, 3-tier application need only in the application server installation and configuration database access software [1], client program only user interface code, installation is very simple.

2 Data Snap Technology Foundation

Data Snap request on the server (actually meddle layers computer) installation specific library, it can take from SQL server or other data extracted from the data provided to clients. Data Snap doesn't need a SQL server for data storage. Data Snap may retrieve data from many kinds of datasource , these Data sources including SQL, other Data Snap server or the data being computed[2].

1) IAppServer Interface

Data Snap application through the use of both ends of the IAppServer interface to communicate.

2) Link Protocol

Data Snap only defines the high-level structure, and use different technical Data from middle layer transplanted to the client. Data Snap support many different protocols, including:

Distributed COM(DCOM) and statelessness COM (MTS or COM+)

TCP/IP socket: most of the system can be use TCP/IP socket. By using TCP/IP, system can distribute client-side in Web needn't use DCOM, which will reduce the configuration of the trouble. In order to use sockets, intermediate computer must be running Borland provide ScktSrvr.exe application, the program can either as a application run, also can be used as service running. The application receives requests client-side and will these requests be transmitted to the remote data module using COM.

HTTP and SOAP: HTTP as Internet transfer protocol, simplified network with a firewall or proxy server (usually with custom TCP/IP socket different) connection.

Providing data packet: the whole Delphi multilayer data access structure thoughts center is data packets. Data packet is a data block move from application server to the client or move from the client to the server. Technically, data packet is data collection of a seed collection. It describes his contains data (usually some data record), lists the data fields and the name of the type. More importantly, a packet including restraint - applied to data collection of rules. Can usually settings these constraints in application server. The server will put them together with data sent to the client application. Through the packet switching, system realizes the communication between the server and the client. In order to quickly respond to user, the server-side supply components in a large data sets in managing multiple packet transmission..

3 Software Architecture

The overall design scheme of this lumber management system is established based on Data Snap technology. In the MIS framework , system determine the specific content involved in forestry industry , partition function for the system, analysis how to realize the function module, determine the effectiveness of the function of software for technology Lord route[3]. Overall structure using MIS framework thoughts, software development using Delphi realization, backend database using SQL Sever, make the treatment process and business Data separation, using Data Snap 3-tier application realize the interaction between system service and system application [4,5].

The Whole System Architecture Design. This system is based on forestry department's objective conditions, according to the user's needs, system has nothing to do with Internet application, therefore, no adopts B/S mode adopted C/S mode. For each department daily data access operations may cause the system load overweight adopted application middle-ware C/S three-layer structure [6].

Physical model

Physical model including application layer, business layer and data layer three layer structure can realize the interaction between system operation interface and the function of the system.

Data layer: system using the database system to store basic data, response data request, and comprehensive business layer management data, according to customers' demand generation customer need data sets.

Business layer: respond to user request, support authorized different users for the use of system function module.

Application layer: system users are need to use the browser access servers and system interact. Among them, the system administrator mainly completes system maintenance, code maintenance, the operation of the system monitoring. The user is completed by the browser daily work.

Logical model

Logical model include three different treatment levels, presentation layer, business logic layer and data storage layer.

Representation layer :It is located in client-side, for system user interface guide, receive user input, and to the application server to send service request, display treatment results.

Business logic layer: It is located in application server-side, for performing business logic, response representation layer queries and inventory, sales, transportation requests, processing the request to the database, sending requests, processing results back to the presentation layer.

Data storage layer: It is located in application database server-side, for performing data logic, based on the basic database and lumber data from a database, response business data request.

This system according to the three-layer system architecture model architecture, architecture process is to follow the principle of unified standards. Adopt three layer architecture model has the following advantages, module function clear and easy maintenance, and can be used in the late many deployment, has a higher security, accordingly, according to the actual situation of forestry management and to obtain the requirement analysis, this system chooses a 3-tier application. Overall system architecture as figure 1 shows.

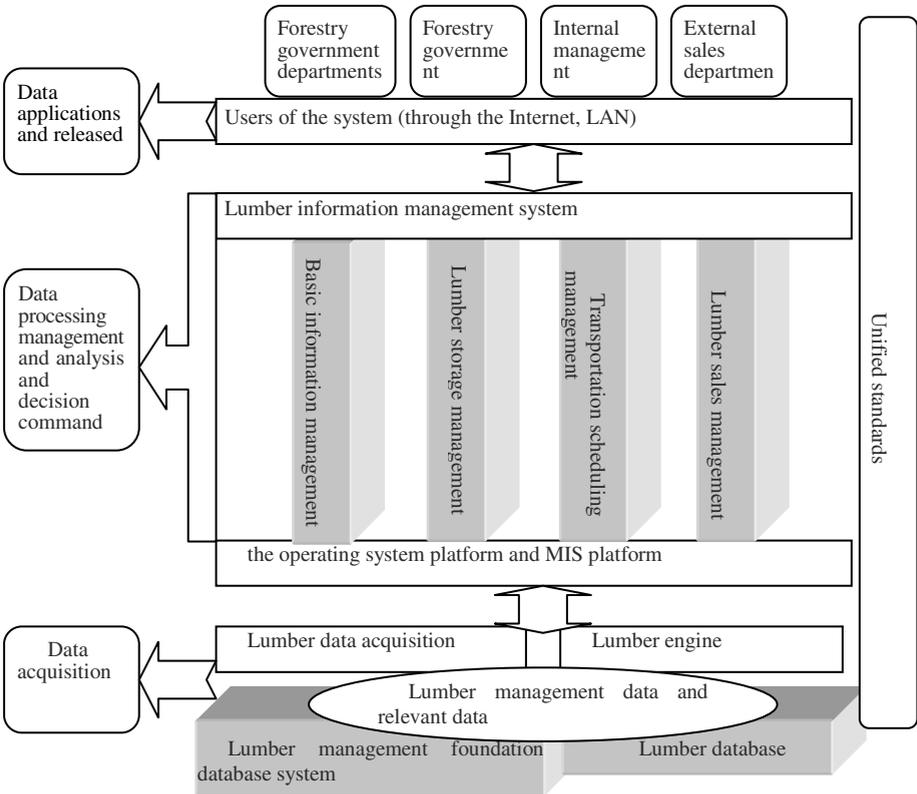


Fig. 1. This is a overall system architecture of a figure

The 3-tier application's bottom-layer based on database system, middle-layer is the operation control layer of the system, top-layer based on the network and users-oriented. Top-layer accept user requests, invoked forestry management system function. According to the system function request of top-layer user, middle-layer processing the required data, analyze the data provide decision-making, commanding system complete top-layer user requests. Bottom-layer response middle-layer's data processing request, complete data acquisition, send data and related data of forestry management to middle-layer.

Software System Structure. Client/Server mode is now the most common distributed computing environment. The traditional two layers of Client/Server structures exposed performance not ideal, expansion of poor, and difficult to transplant, safety, people according to lower malpractice hierarchy theory proposed three-layer Client/Server models, the large-scale application system is divided into three layers presentation layer , middle layer and data service layer. The presentation layer provides concise man-machine interface, finish data input/output. The middle layer complete business logic, for the presentation layer and database provides communication interface. The data service layer provides the data storage service. Three layer structure as figure 2 shows.

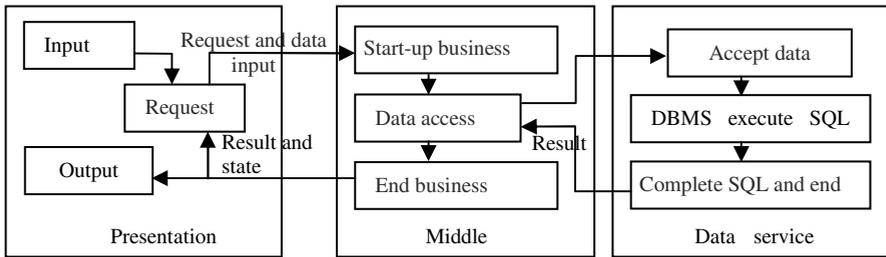


Fig. 2. This is a three layer structure of a figure

Data Snap technology is applied to three layers Client/Server model, figure 3 shows.

Presentation layer, namely the client user interface form proposed data connection request TDispatchConnection, front-end data to develop controls through data attributes, connected to the client data block DataModule of data set TClientDataset. Middle layer, in the client data module unified management all business logic required data source data set of data, data module of data sets TClientDataset in the client through TDataSetProvider controls in server remote data module is done with the server backend database connection. Data service layer, the database server extract data sets TDataSet send messages to client-side, completing the user data request.



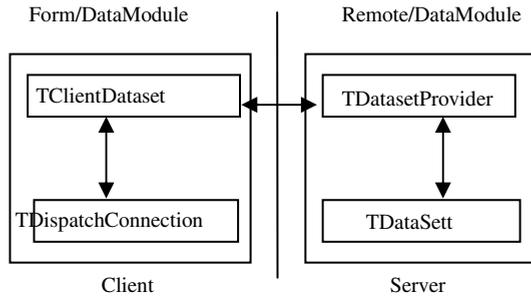


Fig. 3. This is a lumber management information system Data Snap applications of a figure

Data snap technologies as special key technologies for the development of multilayer distributed database, suitable for three layers Client/Server architecture system middle layer development[7]. It is a universal data interface provides good framework, facilitating implementation software reuse, distributed management, load balance, and other functions, can realize the three-layer model performance advantage. Low-level code reuse by object-oriented programming language has become the reality, but it has not suitable for specific field large software production requirement. In order to improve the production process of software reuse dynamics, not only to reuse the old code, but also to reuse similar analysis and design results and system structure, in order to reduce the cost of tectonic new software system and improve software reliability. This framework database using 3-tier application: data access logic is, the client used to interact with users of data acquisition and application server providing business logic and safety logic, database server for access to the data and processing. Middle layer data module, already can independent seal also more easily software expansion, has the good reusability.

4 Conclusion

This paper discusses the Data Snap technology of multi-layer software reuse, and the architecture framework used to lumber management system, which achieves both software engineering development data blocks reusability, also improved the forestry bureau of storage, transportation, sale the automation management process, increasing the reliability of the enterprise information, reduced the business department work and enhance the production efficiency.

References

1. Ruan, J., Yang, C.J.: Computer Engineering 32(14), 57–59 (2006) (in Chinese)
2. Schwerin, R., Xiao, M.: Meridian Data Snap Server. Computech. (11), 11 (1998) (in Chinese)
3. Zhang, X., Li, Z.Y., Deng, G., Chen, Y., Lei, Z.Y., Fan, D.P., Yang, Y.C.: Scientia Silvae Sinicae (S1), 40–43 (2006) (in Chinese)

4. Zhu, J.Z.: Computer Engineering (02), 245–247 (2003) (in Chinese)
5. Wang, Z.G.: Journal of Natural Science of Hunan Normal University (02), 15–17 (2001) (in Chinese)
6. Hong, Z.Q., Xu, L.: Application Research of Computers (01), 46–48 (2003) (in Chinese)
7. Jiang, C.H.: China Water Transport (Academic Version) (02), 98–100 (2006) (in Chinese)

Research of Agricultural Land Classification and Evaluation Based on Genetic Algorithm Optimized Neural Network Model

Liu TingXiang^{1,2}, Zhang ShuWen¹, Wu QuanYuan³, and Bao WenDong⁴

¹ Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130012, China

² Graduate School Chinese Academy of Sciences, Beijing 100039, China

³ College of Population, Resources and Environment, Shandong Normal University, Jinan 250014, China

⁴ Shandong Land Surveying and Planning Institute, Jinan 250014, China
Liu564@hotmail.com

Abstract. The common method used for agricultural land classification is weighted factors sum. Classification results are greatly influenced by subjective weights. Besides, the relationship between relevant factors and agricultural land grades would not be linear. This paper constructs the genetic algorithm optimized neural network model to calculate agricultural land physical quality value nonlinearly. This model is applied to agricultural land classification and evaluation work in Licheng district of Jinan city. It is prove to be effective and robust.

Keywords: Agricultural land classification, BP Neural Network, Genetic Algorithm.

1 Introduction

The agricultural land classification is the synthetical evaluation of the intrinsic properties of cultivated land that determining potential production and external environmental conditions that influencing potential production. Quantified factors are used to determine quality grades of agricultural land. The agricultural land classification is the foundation to utilize and manage limited agricultural land resources reasonably and realize the sustainable utilization of agricultural land [1, 2, 3, 4].

At present, the method in common use for agricultural land classification is weighted factors sum to calculate agricultural land grade value based on determination of factor system and weights [5]. Classification results are greatly influenced by subjective weights. Besides, the relationship between relevant factors and agricultural land grades would not be linear. Over the last decades, artificial neural network and genetic algorithm appeared based on nonlinear science and artificial intelligence. Both of them apply to nonlinear information problems process with multi-factors and imprecise conditions, which have been applied in many fields.

The main characteristics of neural network is learning and training, which is implemented by adjustment of network weights and thresholds [6]. The algorithm used in this paper is back propagation algorithm. BP algorithm has two limitations: One is

easy to fall into local minimum, and another is slow convergence. There are also two methods to solve these problems, improvement on algorithm and optimization of initial weights. The former is usually too complicated, so the latter becomes popular with ordinary user. One of effective optimization methods is genetic algorithm [7].

This paper establishes the Genetic Algorithm optimized Back Propagation Neural Network Model (GA-BPNN). Supported by GIS technology, agricultural land classification work is completed located in Licheng district in Shandong province.

2 Study Area

The case study area is Licheng district, Jinan city, Shandong province (Fig.1).

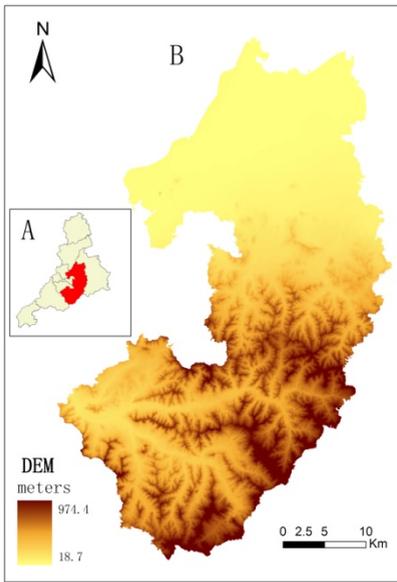


Fig. 1. Location of Licheng district (A)

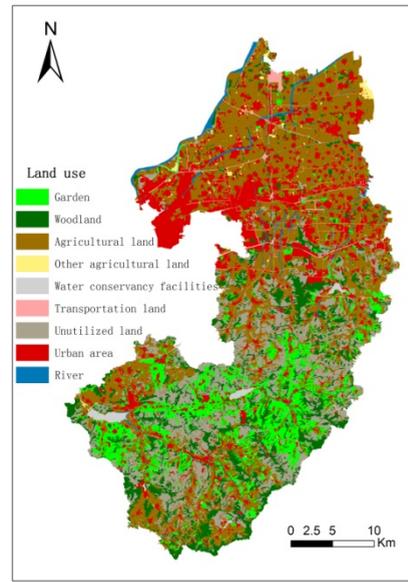


Fig. 2. Land use of Licheng district in 2007

It is located in the junction of Taishan Mountain area and Northern Shandong Plain area, and the junction of Jinan urban district and rural area. The altitude varies from less than 20m in the north to nearly 975m in the south. There is over 60000 ha. agricultural land in Licheng with the average annual precipitation of 665.7 mm and the average annual temperature of 14.3°C. The total sown area was 63300 ha. in 2007, including 47600 ha. of grain and 15700 ha. of economic crops. The total yield of grain was 259500 T. in 2007, with average of 5.45 T./ha..

Digital elevation model (DEM) of Licheng (B)

According to regional topography, climate, vegetation and agricultural zoning, study area is divided into two sections: Jiwei Piedmont (Section A) and Hilly Area of Tai-Lu-Yi-Meng-Ni (Section B), because agricultural classification should consider geographical region [8].

3 Method

Data pretreatment. According to principle of stability, dominant, productive, spatial variability, regional difference and simpleness, 8 factors are selected to classify the agricultural land: surface texture, slope, irrigation guarantee rate, drainage, soil depth, rock outcrops, organic matter content and barriers levels. These factors are digitized and normalized to 0-1 as model input data. Input data include training, testing and running data. Choose agricultural land physical quality value as output data, fig.3.

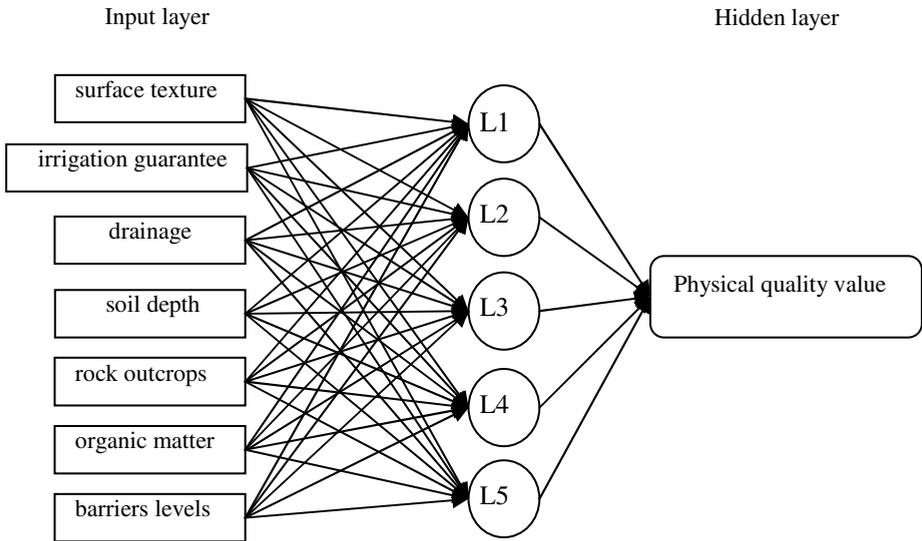


Fig. 3. Structure of neural network for agriculture land classification

Classification cells are determined by spatial overlay of agricultural land plot in 1:50000 map of land use and maps of factors. Besides, cells should be clipped by boundary of Section A and Section B, which would be classified respectively.

Training and testing data. Training plots are samples of BP neural network training, which should be representative and uniformly distributed. The number of training samples is usually determined by experience. According to the reference of other researches, the training sample number is 100 and testing sample number is 25 in Section A and Section B respectively. The agricultural land plot with best condition is selected as well in the sample.

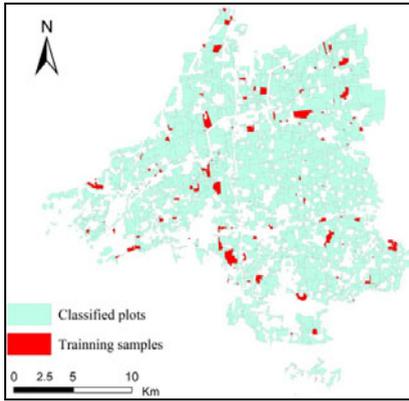


Fig. 4. Training samples of Section A

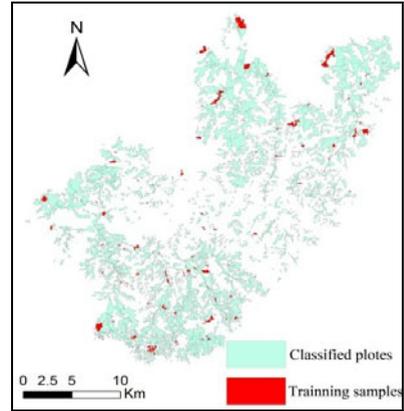


Fig. 5. Training samples of Section B

Physical quality value of training and testing data is determined by sample plots method. First, the best agricultural land was selected with highest physical quality value, and then physical quality value of other training and testing land plot is determined by factors vector euclidean distances to the best agricultural land.

Design a real number coding mode. Genetic algorithm can't operate on concrete problem space directly. Coding can convert the concrete problem to character string that genetic algorithm could use. We use real number as the object genetic operated, which is neural network weights and thresholds in this paper. The coding method is as follows:

$$S=R*S1+S1*S2+S1+S2 \quad (1)$$

Where S is the length of code, R is node number of input layer, S1 is node number of hidden layer, S2 is node number of output layer. R is 8, S1 is 5 and S2 is 1 in this paper.

Initialization and calculate the fitness function. The initial population is 100, and the code is stochastically created in [-1, 1]. They are the first generation. Take the reciprocal of squared error sum of BP neural network as fitness function of genetic algorithm. The smaller the squared error sum, the greater the fitness, and vice versa. Input the stochastic code and training data, and calculate the fitness function to evaluate precision of this code which is representation of neural network weights.

Selection. Select several codes with the best fitness as current generation. The aim of this step is to delivery the code with good fitness value to next generation directly or to generate new code by crossover or mutation. We choose roulette wheel selection as selection operator in this paper, that the probability to be selected of every code is proportional to the fitness value.

Evolution. This step uses crossover and mutation operation evolving current population to next generation. Crossover operator chooses two codes, so-called parents, and then creates their two descendants (children) using the following operation: it

selects a position inside the code string and starting from this position exchanges the remaining parts of the two codes. Mutation operation is an alteration of one or more bits in the code string; a parameter which gives a probability of performing this operation with a certain code is introduced. The code with best fitness in last generation would not be changed by crossover and mutation operation. So the next generation includes the code with largest fitness in last generation and other evolved codes. Some code could have larger fitness than the best code in last generation.

Iteration. Repeat selection, crossover and mutation. So the pre-determined code would continue to be evolved, until fitness is as good as expectation or training times meets the goal.

BP neural network training. The genetic algorithm considers the final best code as optimal solution. According to the coding mode, it is decoded to weights matrix of BP neural network. This weights matrix is input to BP neural network algorithm as well as training data finally.

Testing and application. After training, use testing data to test the model. If the test result is consistent with actual value, the model could be applied to other samples. Otherwise, adjust parameters and training again.

4 Result

Before genetic algorithm optimization, several parameters should be determined: probability of crossover (Pc), Probability of mutation (Pm) and Termination (T), which decide the result of genetic algorithm. Pc is 0.8, Pm is 0.1 and T is 100 in this study, which are determined by experience and experiment.

After 100 times iterations, genetic algorithm obtains the optimal code. It is decoded to weights and threshold of neural network. The training and testing results of BP neural network are as follows:

Table 1. Training and testing results of Section A

| Experiment times | | 1 | 2 | 3 | 4 | 5 |
|------------------|-----------------|------------|------------|------------|------------|------------|
| Training samples | Iteration times | 65 | 263 | 57 | 1071 | 169 |
| | MSE | 9.987E-010 | 9.871E-010 | 9.966E-010 | 9.995E-010 | 9.908E-010 |
| | Maximum error | 1.793E-004 | 1.770E-004 | 2.719E-004 | 2.850E-004 | 2.737E-004 |
| Testing samples | MSE | 5.318E-004 | 5.014E-004 | 5.492E-004 | 5.886E-004 | 5.298E-004 |
| | Maximum error | 6.183E-002 | 6.082E-002 | 6.108E-002 | 6.023E-002 | 6.103E-002 |

(training goal error of neural network is 1E-9)



Results show that precision of training and testing is good. So we can use trained BP neural network to calculate other agricultural land. The physical quality values of Licheng agricultural land are joined to spatial data to realize visualization, as followed in Fig 6:

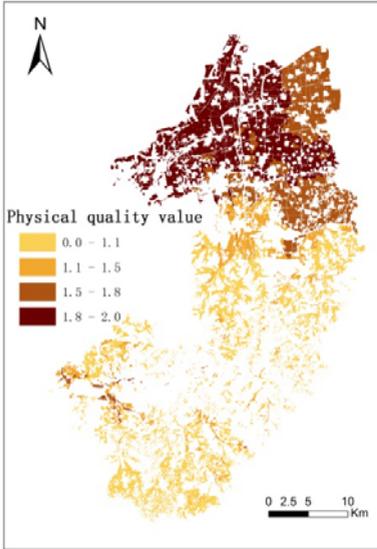


Fig. 6. Physical quality value

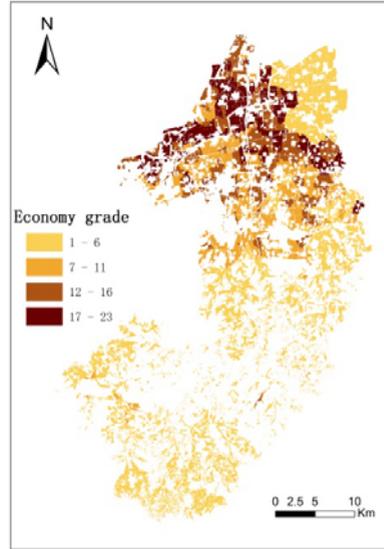


Fig. 7. Economy grade

Based on the calculation method of physical quality grade index, utilization grade index and economic grade index, agricultural grade of physical quality, utilization and economy are determined with equal interval. For further details on every grade calculation process, please see reference [8]. Fig. 7 shows the final economy grade map.

Table 2. Training and testing results of Section B

| Experiment times | 1 | 2 | 3 | 4 | 5 | |
|------------------|-----------------|------------|------------|------------|------------|------------|
| Training samples | Iteration times | 115 | 58 | 51 | 96 | 281 |
| | MSE | 9.699E-010 | 8.751E-010 | 8.826E-010 | 9.190E-010 | 9.976E-010 |
| | Maximum error | 1.142E-004 | 1.120E-004 | 1.494E-004 | 1.251E-004 | 2.139E-004 |
| Testing samples | MSE | 5.048E-004 | 8.684E-005 | 2.865E-004 | 4.692E-005 | 4.466E-005 |
| | Maximum error | 3.086E-002 | 3.109E-002 | 3.061E-002 | 3.020E-002 | 3.033E-002 |

(training goal error of neural network is 1E-9)



Some agricultural land not for growing crops was also classified in this paper. Classified agricultural land is 80986 ha., including 17 physical quality grades, 11 utilization grades and 23 economic grades. The area of 3 grade type has different distribution respectively, where physical quality grades are concentrated to 16 and 17 grade accounting for 70.2% of the total area, utilization grades are concentrated to 3,4 and 5 grade accounting for 47.6% of the total area, while economy grades are dispersed relatively, grade 2 to grade 6 are more concentrative. Comparing to agriculture land grade of entire Shandong province, agriculture land of Licheng district is better with high grade value because of fertile land and developed cultivation.

Discussion

Based on GIS technology, this paper integrated BP neural network and genetic algorithm to agriculture land classification and evaluation work. It brings a new idea for this work. Comparing to traditional weighted sum method, this method could obtain nonlinearity relationship between factors and quality value, and store it in the model structure. When other data were inputted, the model could produce interrelated results according to the training process. Because there is not standard result for agriculture land classification and evaluation, our method couldn't be compared to the weighted sum method on result. There are further researches ahead for the applicability of this method.

Acknowledgment. Financial supports from Farmland resources and Farmland productivity monitor and warning (KSCX1-YW-09-01) are highly appreciated. Two anonymous reviewers are also gratefully acknowledged.

References

1. Chen, M.Q., Zhao, X.M., Chen, A.Z., Cai, Z.M., Zhang, R.C.: Grade- Determination of Arable Land Based on GIS—A Case of Linhai City, Zhejiang Province. *Acta Agriculturae Universitatis Jiangxiensis* 21(3), 421–424 (1999)
2. Peng, J., Jiang, Y.J., Liu, S., Zhang, Q.C.: Research progress and prospect on classification and grading of agricultural land in China. *Chinese Journal of Eco-Agriculture* 13(4), 167–170 (2005)
3. Li, Y.X., Wang, J.M., Zhang, M.L., Zhang, Y.R.: Compare and application research of agricultural land classification and gradation methods. *China Land Science* 11(1), 33–37 (1997)
4. Song, J.B., Zhou, S.L., Peng, B.Z.: Study On Agricultural Land Gradation In Metropolitan Suburban—A Case Study In Tianhe District Of Guangzhou City. *Soils* 36(2), 168–172 (2004)
5. Yun, W.J.: Agricultural Land Classification and Its Application. *China Agricultural University* 6 (2005)
6. Cybenko, G.: Approximation by superposition of a sigmoidal function. *Mathematics of Control, Signals, Systems* 2(4), 303–314 (1989)
7. Chen, G.L., Wang, X.F., Zhuang, Z.Q., Wang, D.S.: Genetic Algorithm And Its Application, pp. 23–76. Posts & Telecom Press, Beijing (1996)
8. Bao, W.D.: Study on Land Use Dynamic Change Based on GIS. *Shandong University of Science and Technology* 6 (2007)

Evaluation of Agricultural Water Use Efficiency and Regional Irrigation Strategies in China Based on GIS

Li Yuyi¹, Pang Huancheng¹, Yu Tianyi², Li Hua², Li Yibing², and Ren Tianzhi²

¹ Key Laboratory of Crop Nutrition & Fertilization, MOA; Institute of Agro-resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, 100081, China

² Institute of Agro-resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, 100081, China
li25263@163.com

Abstract. Enhancing agricultural water use efficiency (AWUE) is the key way for water saving in China. In this paper, the evaluation method for AWUE was established based on the concept of agricultural standard yield and total consumed agricultural water. Through taking province as analysis unit, the relationship between AWUE and irrigation intensity (irrigation amount per acre), irrigation extent (ratio of effective irrigation area), standard yield and cropping index was analyzed. It was found that the difference of AWUE was great and changed from 1.70 kg/hm²·mm to 7.40 kg/hm²·mm in the different provinces of China in 2007. There was a significantly negative correlation between irrigation intensity and AWUE. A positive correlation between irrigation extent, standard yield and AWUE was also observed, but the correlation coefficient did not reach significant level, and no correlation between cropping index and AWUE was observed. According to the distribution of AWUE in China, three different types of regions were divided, the first type regions (AWUE≤3.0kg/hm²·mm) were located at Gansu, Ningxia, Yunnan, Guizhou, Hainan, Fujian, Guangdong and Guangxi. Second type regions (3.0 kg/hm²·mm≤AWUE≤5.0kg/hm²·mm) included Shanghai city, Heilongjiang, Jiangxi, Qinghai, Xinjiang, Shaanxi, Tibet, Hubei, Hunan, Anhui, Jiangsu, Zhejiang, Sichuan and Shanxi, and third type regions (5.0 kg/hm²·mm≤AWUE≤7.5 kg/hm²·mm) mainly included Chongqing city, Tianjin city, Beijing city, Liaoning, Jilin, Shandong, Hebei, Henan and Inner Mongolia. Lastly, different regional irrigation strategies were put forward.

Keywords: Standard yield, Irrigation intensity, Irrigation extent, Cropping index.

1 Introduction

At present, China was confronted with the severe water resources deficiency and the predicted amount of water resources shortage was 13 billion m³ in 2030 [1]. By far the biggest water user in China is the agricultural sector (including livestock), accounting for 61.9 percent of all withdrawals [2]. With the rapid development of social economy, more water will be diverted to other industry sectors and the problems of agricultural water resource shortage are becoming more and more urgent.

Therefore, it is very urgent to discuss how to utilize reasonably water resources and establish relevant development strategies.

The enhancement of agricultural water use efficiency (AWUE) was regarded as one of the important ways to ensure the double safety of grain and water resources in China [3, 4]. In last 30 years, AWUE was greatly enhanced since the increase of crop yield through crop variety improvement. But with the decrease of yield-increasing effect through crop variety improvement, the attention was gradually turned to the improvement of water resources management. Under current water resources constraint conditions, different irrigation development strategies always affect the AWUE of the whole agricultural system. Currently scientific evaluation of water resources and reasonable utilization was the research focus [5, 6, 7]. But now no new and detailed studies were reported on the macroscopic evaluation of AWUE and relevant regional irrigation strategies in China.

This paper quantitatively evaluated the AWUE of different provinces in China in 2007 and analyzed the relationship between AWUE and irrigation intensity (irrigation amount per acre), irrigation extent (the ratio of effective irrigation area), standard yield and cropping index, and then different irrigation strategies were put forward.

2 Methodology

Construction of AWUE Evaluation Model

AWUE represented the average standard yield of agricultural system per consumed agricultural water in a special region and it can be expressed by the following calculation formula:

$$P_{AW} = P_s / W_T \quad (1)$$

Where P_{AW} represents the AWUE (kg/hm²·mm), P_s represents standard yield per acre (kg/hm²), W_T represents total consumed agricultural water (mm).

Standard yield (P_s) represented the average yield level of agricultural system for special cropping system in a certain region. Since different crops possess different use value, economical value and chemical ingredients, so when calculating the average P_s in a certain region each crop product must be converted to standard yield. In this study grain energy was regarded as weight object and rice was regarded as standard object to calculate standard yield. Energy value and standard yield conversion coefficient of different crop products can be seen in Table 1 mainly referred to [6, 7]. So the standard yield can be expressed by the following calculation formula:

$$P_s = \sum_{i=1}^n C_i * P_i * A_i / \sum A_i \quad (2)$$

Where P_s represents standard yield per acre (kg/hm²), C_i is standard yield converting coefficient of i crop type, P_i is crop yield per acre of i crop type (kg/hm²), and A_i is sowing area of i crop type (hm²).

Table 1. Energy value and standard yield conversion coefficient of different crop products

| Variety | Grain energy(kJ·kg ⁻¹) | Standard yield converting coefficient (C _i) |
|------------|------------------------------------|---|
| Rice | 15491 | 1.00 |
| Wheat | 15742 | 1.02 |
| Corn | 16538 | 1.07 |
| Soybean | 20683 | 1.34 |
| Potato | 5008 | 0.32 |
| Peanut | 25916 | 1.67 |
| Rapeseed | 26377 | 1.70 |
| Cotton | 20013 | 1.29 |
| Sugar cane | 1932 | 0.12 |
| Vegetables | 3208 | 0.21 |

Total consumed agricultural water mainly included rainfall, ground water and soil water. Usually in a certain region a part of rainfall can be directly converted to inventory surface water, a part of rainfall was converted to soil water but other parts were lost by the runoff. When calculating total consumed agricultural water by crops, soil water can be replaced with rainfall since soil water was formed through long term calculation of rainfall, and the amount of utilized surface and ground water can be expressed by the irrigation water amount. So total consumed agricultural water in this study can be expressed by the following formula:

$$W_T = W_R - W_F + W_I \quad (3)$$

Where W_T represents total consumed agricultural water (mm), W_R is the amount of rainfall (mm), W_F is the amount of runoff (mm), and W_I is the irrigation amount (mm).

Data Sources

In this study, the data of rainfall, runoff and irrigation amount of different provinces were from China Water Resources Bulletin in 2007, and the index including cultivated land area, crop sowing area, crop yield and effective irrigation area of different provinces were collected from China Rural Statistical Yearbook in 2007.

3 Evaluation and Analysis of AWUE

Evaluation and Analysis of AWUE in Different Regions in China

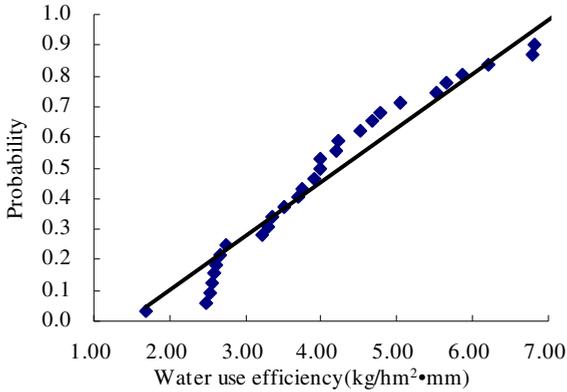


Fig. 1. Distributional probability of AWUE in China in 2007

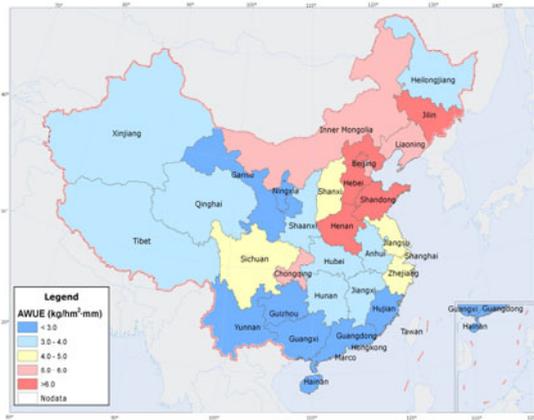


Fig. 2. Distribution of AWUE in the different regions in China in 2007

The AWUE in different provinces in China in 2007 was calculated by AWUE calculation model. As indicated in Figure 1, there was great difference of AWUE in different provinces of China. The value of AWUE was changed from 1.70 kg/hm²•mm to 7.40 kg/hm²•mm and the average value of AWUE was 4.24 kg/hm²•mm. It was also calculated that the dispersion coefficient of AWUE was 0.36 lower than the value of rainfall (0.52), this indicated that irrigation decreased the effect of rainfall difference on the AWUE. However, the dispersion coefficient of AWUE was appreciably higher than total consumed agricultural water (0.34) and it showed



that the comprehensive factors including irrigation technology, production condition, water resources allocation and irrigation and planting strategies enlarged the difference of AWUE.

According to the distribution of AWUE in different regions of China in 2007 (Figure 2), three different type regions were divided. First type region ($AWUE \leq 3.0$ kg/hm²·mm) mainly distributed in the Northwest dry regions, South high rainfall regions and Southwest plateau regions such as Gansu, Ningxia, Yunnan, Guizhou, Hainan, Fujian, Guangdong and Guangxi. In these provinces the standard yield remained at the level from 2900 kg/hm² to 5900 kg/hm² and the average standard yield was only 4308.8 kg/hm², the irrigation intensity (irrigation amount per acre) was relatively high and remained from 790 mm to 1700 mm and the irrigation intensity averagely reached 1165.7 mm per acre. Moreover, in these regions the irrigation extent (ratio of effective irrigation area) was lower than 40% except Guangdong (46%) and Fujian (71%). Second type regions (3.0 kg/hm²·mm $\leq AWUE \leq 5.0$ kg/hm²·mm) were mainly located at the regions such as Shanghai city, Heilongjiang, Jiangxi, Qinghai, Xinjiang, Shaanxi, Tibet, Hubei, Hunan, Anhui, Jiangsu, Zhejiang, Sichuan and Shanxi. In these regions the standard yield kept at the level from 3600 kg/hm² to 6400 kg/hm² and the average standard yield was increased to 4775.4 kg/hm², but the irrigation intensity changed greatly and remained from 320 mm in Shanxi province to 1137 mm in Xinjiang province and the average irrigation intensity was 673.0 mm per acre, and the irrigation extent also varied greatly and remained from 24 % in Heilongjiang province to 84 % in Xinjiang province. Third type region (5.0 kg/hm²·mm $\leq AWUE \leq 7.5$ kg/hm²·mm) mainly distributed in Chongqing city, Tianjin city, Beijing city, Inner Mongolia, Liaoning, Jilin, Shandong, Hebei and Henan. In these regions the standard yield kept at the level from 4200 kg/hm² to 6400 kg/hm² and the average standard yield was 5587.1 kg/hm², the irrigation intensity was relatively lower and remained from 270 mm in Henan province to 760 mm in Liaoning province and the average irrigation amount reached 446 mm. Moreover, the irrigation extent in these regions varied greatly and kept from 28 % in Chongqing city to 79 % in Tianjin city.

Effect of Irrigation Intensity on AWUE in China

Annual irrigation amount per acre was regarded as the representative index of irrigation intensity. Since the whole cropping system as the research object, irrigation intensity always represented the average irrigation intensity of multiple crops. As indicted in Figure 3, with the increase of irrigation intensity, AWUE decreased gradually and there was a significantly negative correlation between irrigation intensity and AWUE in different regions of China in 2007. The fitting curve equation can be expressed by the following: $f(x) = 3 \times 10^{-6} x^2 - 0.0086 x + 8.6429$, in which $R^2 = 0.6841$, and the correlation coefficient reached significant level.

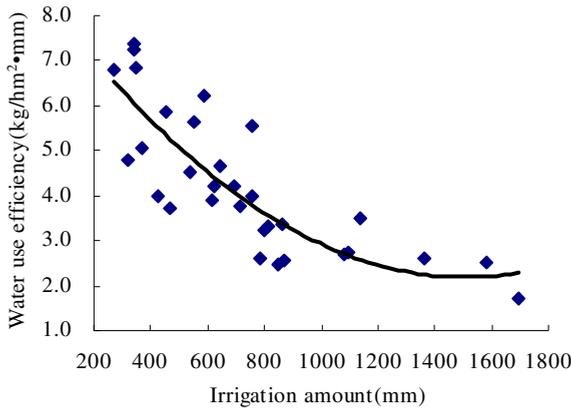


Fig. 3. Relationship between irrigation intensity (irrigation amount per acre) and AWUE in different regions of China in 2007

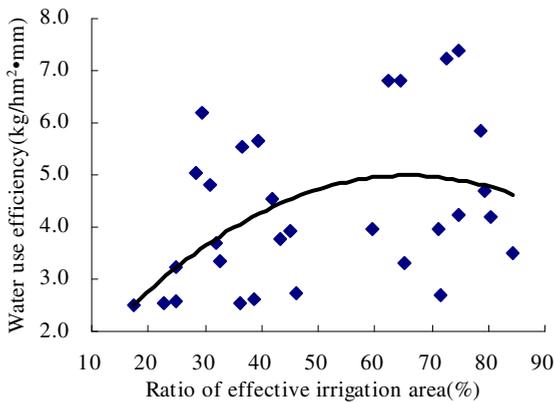


Fig. 4. Relationship between irrigation extent (ratio of effective irrigation area) and AWUE in different regions of China in 2007

Effect of Irrigation Extent on AWUE in China

Irrigation extent also affected AWUE of whole cropping system from another side. In this study, the ratio of effective irrigation area was used to represent irrigation extent since it always influenced the extent of yield increase and stability for drought resistance in the whole cropping system. From the Figure 4, it was observed that with the increase of the ratio of effective irrigation area, AWUE increased gradually and there was a positive correlation between irrigation extent and AWUE, but the correlation coefficient did not reach significant level. The fitting curve equation can be expressed by the following equation: $f(x) = -0.001x^2 + 0.1383x + 0.4296$, in which $R^2 = 0.2108$.



Effect of Standard Yield on AWUE in China

Standard yield also affected the AWUE of whole cropping system in each region. As indicated in Figure 5, with the increase of standard yield, AWUE increased gradually. There was a positive correlation between standard yield and AWUE in the different regions in 2007, but the correlation coefficient did not reach significant level. The fitting curve equation can be expressed by the following equation: $f(x) = 7 \times 10^{-8} x^2 + 2 \times 10^{-4} x + 1.4536$, in which $R^2 = 0.3334$.

Effect of Cropping Index on AWUE in China

Cropping index in a certain region always affected crop production and then affected AWUE of the whole cropping system. But from the Figure 6, no significant correlation between cropping index and AWUE was observed in the different regions of China in 2007.

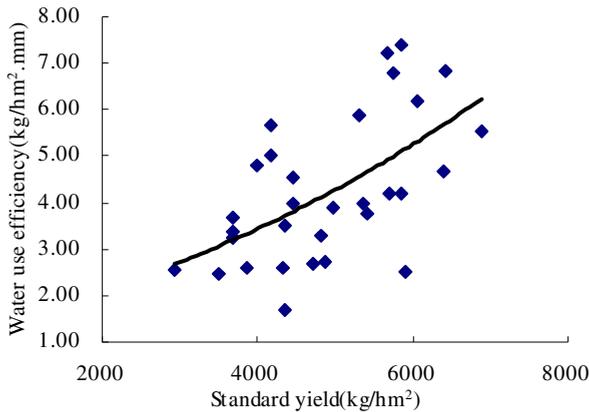


Fig. 5. Relationship between standard yield and AWUE in different regions of China in 2007

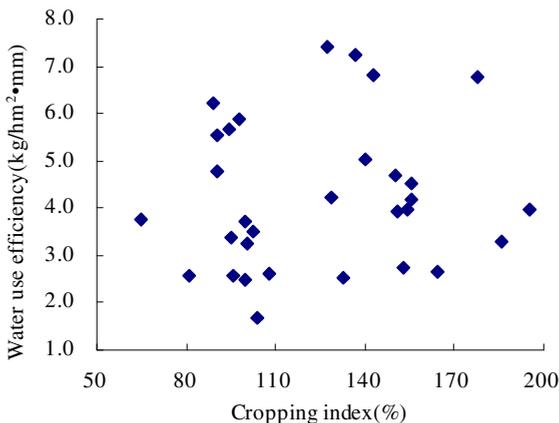


Fig. 6. Relationship between cropping index and AWUE in different regions of China in 2007

4 Regional Irrigation Strategies in China

Irrigation Strategies in the Regions of First Type

First type regions included Gansu, Ningxia, Yunnan, Guizhou, Hainan, Fujian, Guangdong and Guangxi provinces. In these regions AWUE was lower than $3.0\text{kg}/\text{hm}^2\cdot\text{mm}$ and the standard yield was also low since the poor cultivated land condition and farming management. Because of high field evapotranspiration, irrigation intensity was generally at a high level in these regions. Large amount of runoff and field leakage also resulted in large irrigation demand although rainfall is abundant in certain regions. In addition, the irrigation extent was lower except Guangdong (46 %) and Fujian (71 %) provinces. In these regions, we should extend water-saving irrigation techniques and actively reduce irrigation intensity. And at the same time we should strengthen water conservancy construction to further enlarge the irrigation extent and gradually enhance standard yield in these regions to obtain higher AWUE.

Irrigation Strategies in the Regions of Second Type

Second type regions mainly included Shanghai city, Heilongjiang, Jiangxi, Qinghai, Xinjiang, Shaanxi, Tibet, Hubei, Hunan, Anhui, Jiangsu, Zhejiang, Sichuan and Shanxi provinces. Compared with the regions of first type, these regions possessed higher AWUE and the irrigation intensity was relatively lower in these regions, but was still at a high level. In these regions, we should improve irrigation facilities and methods to reduce irrigation intensity. Additionally, the potential of standard yield increase was higher in these regions, so in order to obtain high AWUE increasing the standard yield in these regions was needed. It was need to indicted that in Heilongjiang, Shaanxi, Qinghai and Shanxi, the ratio of effective irrigation area was very low, this significantly resulted in lower AWUE, so in these regions the irrigation extent needed to further enlarge to obtain higher AWUE.

Irrigation Strategies in the Regions of Third Type

Third type region included Chongqing city, Tianjin city, Beijing city, Liaoning, Jilin, Shandong, Hebei, Henan and Inner Mongolia provinces. In these regions the AWUE and standard yield was higher, the irrigation intensity was relatively lower and irrigation extent was also high since the higher economic level and farming management. In these regions, due to large amount of industrial and domestic water demand and excessive exploitation of groundwater, the ratio of effective irrigation area should not be enlarged except Chongqing city, Liaoning and Jilin provinces, we should emphasize the improvement of traditional irrigation technologies, and continuously extend modern water saving technologies such as sprinkler, micro-irrigation and low pressure pipe irrigation to reduce the irrigation intensity as far as possible and at the same time to optimize layout for agricultural production and develop agriculture with distinctive features such as facilities agriculture in some regions to reduce the excessive use of groundwater.

Acknowledgment. This manuscript is based on the work financially supported by the Special Fund for Public Welfare Industrial (Agriculture) Research of China (201103001).

References

1. Liu, C.M., Chen, Z.K.: Water resources present situation evaluation and supply and demand development trend analysis. China Waterpower Press (2001)
2. Ministry of Water Resources, People's Republic of China. China water resources bulletin. China Water Power Press, Beijing (2007)
3. Peng, S.Z., Bian, L.M., Zhu, C.L.: Research and development of crop water production functions. *Advances in Science and Technology of Water Resources* 20, 17–20 (2000)
4. Xu, D., Gong, S.H., Li, Y.N.: Overview of recent study on improvement approaches and methods for crop water productivity. *Shuili Xuebao* 41, 631–639 (2010)
5. Hu, H.P., Lei, Z.D., Yang, S.X.: Agricultural water resources effective utilization and sustainable development. *China Rural Water and Hydropower*, 13–17 (1999)
6. Zhou, L.G., Cao, W.X., Zhou, Z.G.: Knowledge model based agricultural water productivity ratio evaluation and regional irrigation strategy analysis. *Agricultural Research in the Arid Areas* 24, 117–121 (2006)
7. Wang, R.F., Liu, J.H., Qin, D.Y., Liu, M.: Evaluation of Agricultural Water Production in Shaanxi Province. *Water Saving Irrigation*, 6–10 (2009)

The Research on Network Courses Education Resources Integration Based on P.E.

Xu Xiaoyang^{1,2}

¹ Sports, Changzhou University, Changzhou, Jiangsu, China

² Sports Engin. Research Inst., Wuhan Inst. of P.E., Key Sports Engin. Lab,
General Administration of Sport of China, Wuhan, Hubei, China
xu_keping@21cn.com

Abstract. With the computer technology and network communication technologies, networks have become the majority of people learning and living in the main channel to get information, online teaching has become of the times and social development of modern means of education. Powerful computer-based information dissemination network capacity, excellent performance of resource sharing, advanced two-way interactive and multimedia technologies for teaching modern distance education can provide a wealth of teaching resources and flexible forms of teaching, improve student learning initiative and flexibility conducive to individualized teaching and learning. Teaching Physical Education Curriculum for space conflicts, learning and training conflicts and other issues, Physical Education reform should make full use of the advantages of modern distance education, integrated curriculum resources, teaching time and space to expand, enrich teaching ways to make physical education to meet the requirements of social development.

Keywords: Network Courses, Education Resource, Integration, P.E.

1 Introduction

With the computer technology and network communication technologies, networks have become the majority of people learning and living in the main channel to get information, online teaching has become of the times and social development of modern means of education. Powerful computer-based information dissemination network capacity, excellent performance of resource sharing, advanced two-way interactive and multimedia technologies for teaching modern distance education can provide a wealth of teaching resources and flexible forms of teaching, improve student learning initiative and flexibility conducive to individualized teaching and learning. Teaching Physical Education Curriculum for space conflicts, learning and training conflicts and other issues, Physical Education reform should make full use of the advantages of modern distance education, integrated curriculum resources, teaching time and space to expand, enrich teaching ways to make physical education to meet the requirements of social development.

2 Modern Distance Education Concepts and Features

Modern distance education is that students and teachers, students and educational institutions using a variety of media, the main means of distance education system, teaching and communication links of education. It is with the development of modern information technology, a new form of education took place, is to build a knowledge economy people are the primary means of lifelong learning system. China has appeared in TV video teaching, TV, correspondence and other traditional forms of distance education, one-way, there is study time, space and other limitations, and as a demonstration tool for teaching media are not the cognitive tools can not achieve interactive , independent of the study. Modern distance education is associated with multimedia technology and network, development of communication technology and the formation of a new two-way interactive new way of education is the era of knowledge economy Gouzhu people the main means of life-long learning system, with space-time adaptive and education, open , interactive, it has retained the traditional teaching of its advantages, but also broke the existing form of classroom teaching is conducive to the establishment of student-centered learning, and enrich the teaching content, to increase the students interest in learning and self , scalability of the knowledge, science and technology to develop students ability to work, and to timely feedback, in large part to address cross-regional allocation of teachers Optimization. With its advanced communication network to transmit information, can be achieved across time and space, interactive, and independent learning. Modern distance education is to fully utilize and optimize the resources of education, popularization and improvement of the quality of education, the formation of open education network and build life-long learning system in a major project.

Distance education can be a large capacity, high speed, wide range of information dissemination, the help of technical means to build rich multimedia educational resources information, to break the constraints of time learning to improve student learning autonomy and initiative. PE teaching, the learner may at any time and place to learn from the network theories of knowledge, motor skills, overcome by space, time or climate impact. Meanwhile, sports skill teaching techniques involved in the existence of various body movements is difficult to explain, understand difficult issues such as difficult and repetitive, you can also multi-media video, animation or Web site to collect major sports events in the sports stars of the standard action to resolve assisted instruction.

3 Physical Education Online Education Resources Construction Status and Problems

Network Course Computer Network Technology is being applied to the concrete embodiment of teaching, is based on multimedia, computer networks and the Internet on the basis of a new teaching model is extended to the classroom, campus network, the Internet, of resources in the school or even the country, around the world to share teaching [1]. Sports Network courses include all levels (national, provincial and university level) courses of quality course Web sites and specialized sports programs, the Ministry of Education, Ministry of Finance issued the "teaching quality and

teaching reform project" file under the guidance of the whole country off course under construction boom, after a decade, our universities have built a surprising number of quality courses in the institutions of the campus network can be seen on the home page course, network curriculum shadow, however, the survey found , many course Web sites have become nobody cares "on schedule." Resource old, site updates lag, and even some in name only. Compared to other professional sports network course more worrying situation. The reason, the network policy when course construction, financial support, lack of maintenance after the completion of funding and policy to encourage students to use is the most important factor. Internet courses follow the development trend of information technology education and lifelong learning needs, for people to acquire new knowledge at any time provides a convenient and powerful support [2]. If we can curriculum resources in the existing sports network sports on the basis of distance education, not only can solve a lot to learn but who can not learn in school (rural middle and primary school teachers, social workers, sports, athletes and coaches) the study problems, while distance education course fees can also be used for follow-up development, so that the network curriculum resources for sustainable development.

4 PE Curriculum Resources for Teaching Online Education Advantages

Network Teaching Physical space can be expanded to address the shortage of teachers, inadequate facilities and other issues. Professional and other professional sports the biggest difference is that many courses require a specific time and place, and there are theoretical and technical subjects of the points. In schools, technical teaching of traditional sports venues will often be in short supply, weather and other factors, the theories are also often the students to participate in training and competition time, be affected by the conflict. Outside the school, PE teachers, athletes, coaches, sports instructors and other sports, and social workers to take advantage of study leave or interrupt their work to the university, to pay a lot of time and some learning, cost of living, to achieve the desire to learn. Construction of sports online course, course content through the speaker teachers will publish to the Web, teacher-student interaction to take questions, answer questions and other forms of extra-curricular learning and after school, to expand teaching space and time, can make up for school hours, the site is inadequate; but also can overcome the external many difficulties learners. Network Teaching Physical Education to update the resources to help students master the latest and most cutting-edge knowledge. The rapid development of world sports, sports competition rules (such as soccer, volleyball, table tennis, etc.) constantly updated and improved, teachers need to keep learning and care about sports, hot issues, and the most advanced teaching methods used in teaching , the latest sports knowledge to the students. Physical Education Site for timely collection of media information on the latest events, abundant teaching resources, teaching theory with practice, to resolve difficulties. The traditional textbook publishing cycle for a long, slow inevitably delay the update of knowledge. Teaching physical education curriculum through the website, teachers can broadcast directly from the Internet or download the latest sports knowledge to the students and enrich the teaching content, the students according to their needs and characteristics of the "Order" independent study, an increase of interest

in learning, fast access to the forefront of sports knowledge, thereby enhancing the level and quality of teaching.

Network Teaching PE PE teachers can overcome the age, physical factors to cause barriers to the teaching .In the traditional physical education, movement model is the most intuitive, most important teaching method. However, the "Sports is not old age," physical education teachers as their own quality, age and other reasons, some technical movement model in place, and not standardized, not coherent, so that students in the beginners action correct action is difficult to grasp the concept, but through video, video and other development and utilization of network resources, students can view the course website by elite athletes world-class standard action, or action sports videos young teachers themselves to learn the concept of establishing the correct action. Moreover, the physical education curriculum website information over time, physical education teachers to record and preserve the valuable teaching resource.

An open network resources, sharing, interaction, collaboration and independent features such as network data access is the process of data mining, information retrieval needs of those who possess the necessary skills and memorizing capacity of the search. Network teaching physical education that will help students to learn the process of self-awareness and self-educational function; help train students to think independently, ask questions and practical ability to solve problems; help students according to their own knowledge structure, time and place suitable arrangements for Learning plans and learning progress, to achieve individualized. Materials through the Web site other than the knowledge of the latest action sports technology, sports dynamic. Both expand their horizons at the same time, through the course website with teachers, students interactive discussions, both for the students and teachers, life and communication, but also develop self-learning ability of students.

5 Modern Distance Education and Online Education Resources, Physical Education Curriculum and Practice the Concept of Integration

Internet-based distance education is an important part of teaching is the implementation of modern distance education project, the formation of open education network, the establishment of education for an important part of teaching resources. "National long-term development plan for education reform and the" proposed "to strengthen the network system of teaching resources, develop e-learning courses, innovative online teaching mode, distance education to carry out high quality education." "Teaching of Physical Education Guidelines "clearly states:" To attach importance to theory and practice, practice teaching attention in the sports-related theoretical knowledge penetration and use of various forms and modern teaching methods." Physical Education Teaching Resource Network is an open, interactive, sharing, economy, collaboration and autonomy and so on, can promote the formation of lifetime sports consciousness of students but also to expand PE curriculum and curriculum resource development and effective way. Resource-based, teaching resources is the modern distance education can be sustained and healthy development of the fundamental [3]. Physical Education teaching resources through the network (such as courses, web courses, etc.) to carry out distance education is quality education in school physical

education to adapt to the necessary way to school physical education in the future to construct the primary means of lifetime sports, but also to integrate existing sports boutique courses and online course resources to ensure quality construction of the national sustainable development results. Modern Distance Education and Physical Education Network Teaching is conducive to building a learning society, to facilitate autonomous learning and lifelong learning. Hunan Institute of Physical Education courses place more emphasis on the construction of sports networks and use of course, network course construction have begun to take effect, the current project, "college sports", the "School Sports", "exercise physiology" and "Sports Anatomy" 4 Hunan Province, courses, project "University Sports", "School Sports", "Exercise Physiology", "Sports Anatomy," "volleyball", "Basketball", "Gymnastics", "track" and "Volleyball" 8 field grade quality courses, into a "sports multimedia courseware," "School Sports", "Exercise Physiology", "Sports Economics" and "computer software used to learn" and five network programs. Theoretical courses such as "School Sports", "exercise physiology" and "Sports Anatomy" and because of the limited hours of classroom teaching can not meet the learning needs of students, school teachers using the course website will be used in courseware, videos and other resources open to students and to provide reference materials, online testing section, an effective solution to the outstanding student class "enough", the extra time to facilitate the students in a dormitory, classroom, or vacation home in the independent study. Technical Courses like "College Sports", "Gymnastics", "Volleyball", "Athletics", "basketball" and other courses timely collection of sports trends, expose students to the latest events, watching sports stars of the movement technique, the teacher timely explanations, students in a relaxed pleasant atmosphere learned knowledge. Speaker of teachers through the course website to update to enable the students to preview, review after class, some students for participating in competitions, training and other reasons do not get the class, you can use the course website tutorial. Network teaching physical education resource utilization high, the school "outstanding online course" and a number of multi-gate time Teaching Award and praise from students. Also project a number of teaching practice at the provincial level, university education reform issues, teaching experience, sum up the network, to promote the formation of theoretical results.

6 Summary and Outlook

In the context of Quality Physical Education in China built a large network of educational resources, but most of the sports courses, online course heavy construction, light maintenance, low resource utilization. Distance education can take advantage of educational resources and optimize and improve the quality of education, the formation of open education network and build life-long learning system. Distance education and physical education curriculum integration network, able to expand the teaching time and space to address the sports learner's "learning and training contradictions" and "engineering contradictions", but also contribute to the existing network of physical education curricular and course of sustainable development, there conducive to physical education resource sharing is the future school physical education and lifelong physical education an important way.

Acknowledgment. The work is supported by Humanities and social science research projects in Changzhou University No. JW201009.

References

1. Wang, Z.-Y.: Study of Construction of Network Teaching of Physical Education. *Journal of Qiqihar University (Philosophy & Social Science Edition)* (2), 171–172 (2010)
2. A, Y.-G., Y, S.-P.: Present Situation and Characteristics of Construction of Network Courses of China's Physical Education Specialty. *Journal of Shanghai Physical Education Institute* 32(4), 68–72 (2008)
3. Lin, H.: Construction, Reorganization and Utilization on Teaching Resources in Distance Education. *Journal of University of Electronic Science and Technology of China (Social Sciences Edition)* 7(3), 88–91 (2005)
4. Joyce, K.M.: Enhancing Social Presence in Online Learning: Mediation Strategies Applied to Social Networking Tools. *Online Journal of Distance Learning Administration* 12(4) (2009)

Design of Field Information Collection System Based on uC/OS-II

ZhiLai Zhang¹, Yan Li², and ZhiAn Wang³

¹ HanDan College, HanDan, China 056005

² HanDan No. 25 middle school 056000

³ HanDan College, HanDan, China 056005

zhang18309@sina.com

Abstract. In this paper, a portable field information collection system is designed, according to the actual needs of field information collection. System is composed of processor C8051f021 and uC/OS-II operating system as the platforms of hardware and software, System working processes and the using of GPS modules are emphasized. This system, with simple operating method and smart expanding interfaces, is meaning, full in updating agricultural automatic equipments with low cost.

Keywords: UC/OS, GPS, C8051F020, NMEA0183.

1 Introduction

Precision planting, also known as precision agriculture. Field information collection is the key to implementation of precision agriculture. Generally, data is collected by sensor, and sent to Control center through the network. Public Telephone Network and mobile phone network are used to Transfer data. However, when measured in a wide range, cost too high [1].

To solve these problems, wireless sensor network is used by domestic and foreign researchers to collect field information [2]. Through the method data's intermittent collection and uploading layer upon layer, system cost is cut down to some extent. But wireless transmission is constrained by the communication distance and power supply mode; the distance between nodes can not be too long. When need temporary research or monitor the area with large area but small information changes in local area WSN (wireless sensor network) cost too high, and difficult to promote at the same time.

Based on the latest research of field information collection system, from the actual needs of data processing, reasonably design the hardware and the system's working arrangements, and meet the design requirements

2 Hardware Design

Portable field information collection system consists of six parts, controller module, sensor module, GPS module, touch screen module, power module, SD card driver module. The block graph of is illustrated in figure 1.

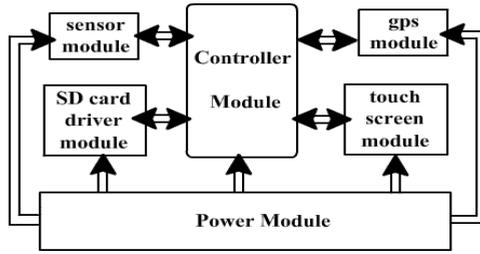


Fig. 1. System block diagram

Considering the practical needs of the information collection system, controller module is C8051f020 [3], which produced by Silicon. C8051f020 consists 64k Bytes FLASH,4352 Bytes Internal Data RAM, hardware SMBus, SPI, two UART Serial Ports Available Concurrently, five General Purpose 16-bit Counter/Timers , low power mode support. The data acquisition of relative humidity and temperature is through chip SHT11 [4], made by Sensirion in Switzerland. The data acquisition of environment light intensity is through TSL2561 [5] converters, made by TAOS company. Specific technical parameters can refer to table 1.

Table 1. Sensor technical parameters

| name | humidity and temperature (SHT11) | light intensity (TSL2561) |
|-------------------|--|---------------------------|
| Operating voltage | 2.4~5.5[V] | 2.7~3.6[V] |
| Range | Humidity : 0~100%[RH] Temperature : -40~123.8[°C] | 0 to 40,000 [Lux] |
| Accuracy | Humidity : ±3% Temperature : ±0.4°C | |

For ease of data processing and Field mesh, GPS module is Gn-80-h GPS-OEM Board [6], which produced by FURUNO.Gn-80-h is the GPS Receiver with 16 parallel channels, tracking performance is superior, and quickly locate. The Communication Protocol of Gn-80-h follows the NMEA0183 (V. 2.3).

NMEA0183 is the most generally international standard about Marine electronic communications interface and protocol. Under the NMEA-0183 standard, all characters used are printable ASCII text (plus carriage return and line feed). NMEA-0183 data is sent at 4800 baud. The data is transmitted in the form of "sentences". Each sentence starts with a "\$", a two letter "talker ID", a three letter "sentence ID", followed by a number of data fields separated by commas, and terminated by an optional checksum, and a carriage return/line feed.

If data for a field is not available, the field is simply omitted, but the commas that would delimit it are still sent, with no space between them. Since some fields are



variable width, or may be omitted as above, the receiver should locate desired data fields by counting commas, rather than by character position within the sentence.

The optional checksum field consists of a "*" and two hex digits representing the exclusive OR of all characters between, but not including, the "\$" and "*". A checksum is required on some sentences.

Table 2. Data structure

| | | | | | | | | | |
|---------|------------|----|-------------|----|--------|----|----|-----|-------|
| \$GPGLL | ,3444.0000 | ,N | ,13521.0000 | ,E | 123456 | ,A | ,A | .43 | CR LF |
| Field# | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |

This select \$GPGLL statement format, GPS-OEM'S Instruction \$PFEC, GPInt, GLL00. Output statement format is as follows table 2. And the domains' mean are listed in table 3.

Table 3. Domains' mean

| #. | Description | Range | [Bytes] |
|------|-------------------------|---|---------|
| 1-2. | Latitude | | |
| | “34”: degree | 00-90 | [2] |
| | “44”: minute(integer) | 00-59 | [2] |
| | “1234”: minute(integer) | 0000-9999 | [4] |
| 3-4. | “N”: North/South | N or S | [1] |
| | Longitude | | |
| | “035”: degree | 000-180 | [3] |
| | “21”: minute(integer) | 00-59 | [2] |
| | “0000”: minute(integer) | 0000-9999 | [4] |
| 5. | “E”: East/West | E or W | [1] |
| | UTC | | |
| | “12”: hh 00 | 00-23 | [2] |
| | “34”: mm 00 | 00-59 | [2] |
| 6. | “56”: ss 00 | 00-59 | [2] |
| | Status | A or V | [1] |
| | | “A”: Data Valid (Stand-alone or DGPS) “V”: Navigation receiver warning | |
| 7. | Position | A: Autonomous D: Differential mode N: Data not valid | [1] |
| | | | |
| | | | |
| 8. | Checksum | | [2] |

3 Embedded Operating System: μ C/OS-II

μ C/OS-II is a low-cost priority-based pre-emptive real time multitasking operating system kernel for microprocessors, written mainly in the C programming language. It is mainly intended for use in embedded systems. It has ports for most popular processors and boards in the market and is suitable for use in safety critical embedded systems such as aviation, medical systems and nuclear installations.

μ C/OS-II is a fully preemptive real-time kernel. This means that μ C/OS-II always runs the highest priority task that is ready. The design is divided into six tasks. Task are listed in table 4

Table 4. Task partitioning

| Task No. | Task name | Real-time | Priority |
|----------|-------------|-----------|----------|
| 1 | WatchDogStk | Low | 7 |
| 2 | KeyTaskStk | High | 6 |
| 3 | WriteSDStk | High | 5 |
| 4 | LCDSStk | Higher | 4 |
| 5 | ADStk | Higher | 3 |
| 6 | GPSSStk | Higher | 2 |

All tasks and mailboxes, queues, semaphores, fixed-sized memory partitions, etc. are created in main function.

```
void main (void)
{
    WDT_Init();           // Initialize WatchDog
    Timer0_Init();       // Initialize Timer
    Uart_Init();         // Initialize UART
    Lcd_Init();          // Initialize LCD
    GPS_Init();          // Initialize GPS
    OSInit();
    //Create mailboxes, queues, semaphores
    GprsQ = OSQCreate(&GprsMsg[0],10);
    SMsgQ = OSQCreate(&SMsg[0],5);
    // Create fixed-sized memory partitions
    Mem20 = OSMemCreate(Part1,20,50,&err);
    Mem50 = OSMemCreate(Part2,100,10,&err);
    // Create tasks;
    OSTaskCreate(WatchDogTask,(void*)0,&WatchDogStk[0],7); // WatchDog task
    OSTaskCreate((void*) KeyTask,(void*)0,&KeyTaskStk[0],6); // key task
    OSTaskCreate((void*)WriteTask,(void*)0,& WriteSDStk[0],5); // SD card write
tasks
    OSTaskCreate((void*)GPRSTask,(void*)0,& LCDSStk[0],4); // LCD Display task
    OSTaskCreate((void*)SMsgTask,(void*)0,& ADStk[0],3); // AD task
    OSTaskCreate(ReadTask, (void *)0, & GPSSStk[0],2); // GPS task
    OSStart();
}
```

Concurrent multi-task scheduling does not consider the low-power applications. It is difficult to play a role in portable monitoring environment limited resources. Therefore, it needs to make some improvements. In μ C/OS-II, idle task is running when the system does not have other tasks to be scheduled execution. The task to perform a busy wait operation, continuous loop waiting for interrupt. Processor has been running and

Consume more energy. To reduce power consumption, idle task is modified to the form.

```
void OSTaskIdleHook (void)
{
  PCON |=1; /* Enter low power mode */
}
```

4 Conclusion

Based on the actual needs of precision agriculture, combined with GPS technology and embedded technology, this paper designed a portable field data acquisition system based on C8051F series C8051f020 and GPS OEM Board. The system has high performance, low cost, compact and low energy consumption, easy to carry, etc. suitable for outdoor operation in a long time to carry and use.

References

1. Bogen, H.R., Huisman, J.A., Oberdor: Evaluation of low cost soil water content sensor for wireless network application. *Journal of Hydrology* 344(1-2), 32–42 (2007)
2. Zhang, Z., Wang, Z., Li, Y.: Design of Wireless Monitoring Networks for intelligent Irrigation. In: *Proc. SPIE*, vol. 7752, p. 775207 (2010)
3. C8051F02x data sheet (rev. 1.4) Silicon Laboratories (2003)
4. SHT1x/SHT7x Relative Humidity & Temperature Sensor System Sensirion (2004)
5. TSL2560, TSL2561 data sheet. TAOS059K (April 2007)
6. GN-80 Series GPS Receiver. Furuno (2003)
7. GPS RECEIVER GN-80 series NMEA version PROTOCOL MANUAL. Furuno (2003)

Analysis and Prediction of the Total Number of Harbin Ice-Snow Tourism Based on Times Series

Yu Zhang

Computer & Information Engineering College, Harbin University of Commerce Harbin,
Heilongjiang 150028, China
zhangyj1@yeah.net

Abstract. This paper constructs unary linear regression model of time series and predicts the total-arrival indicator manifesting the Harbin ice-snow tourism market demand by adopting the method of integrating normative research & empirical research, qualitative analysis & quantitative research, which has significantly reference meaning to exerting great efforts to develop the ice-snow tourism industry and realize the sustainable development of the ice-snow tourism in Harbin.

Keywords: Ice-snow tourism, Tourism arrivals, Time series, Unary linear regression.

1 Introduction

Tourism is a complicated social activity and economic activity [1]. The tourism market is generated and developed with the tourism economic activity emerged and developed [2]. The research into the tourism market can grasp the content, constitution and change trend of the tourism demand [3]. Further it is beneficial to defining the target market, expanding the visitor source and determining the tourism development planning [4]. In the meantime, the change in the tourism demand is usually manifested by the tourism market. It is directly associated with the development depth and breadth of the tourism industry. Obviously, it can be seen that grasping the change rule of the tourism market demand is fully significant to the rise & fall of each tourism reception country or tourism destination [5].

Tourists should finish the experience of their own traveling and realize the purpose of traveling through spatial mobility as the personnel mobility is one of the substantive characteristics of the tourism activity in the research into the tourism market [6]. Consequently, the tourist flow with certain mobility direction and mobile personnel quantity is formed. It manifests the general state and development trend of the tourist spatiotemporal distribution and constitutes the base and driving force of the scale development of the tourism market.

When the market demand scale and change and development rules of the ice-snow tourism in Harbin are measured, the “tourism arrival” is one of the most vital economic indicators. It forms and directly reflects the potential market demand of the visitor

source. Analysis of tourism arrival is the crucial foundation and of the analysis of the tourism market. The scale of the tourism arrival decides the total quantity market demand. It is an imperative to have certain quantity of tourist source in continual increase trend as the foundation and guarantee in order to form certain scale tourism market. This paper establishes quantitative relationship among tourism arrival and relevant influencing factors by selecting “Total tourism arrivals” as the vital economic-technical indicators for analysis to measure the demand intensity of ice-snow tourism market and predict the future development trend of the total ice-snow tourism arrivals.

2 Unary Linear Regression Model

The representation of unary linear regression model:

$$y_t = b_0 + b_1 \times x_t + u_t \quad (1)$$

Eq. 1 indicates the real relationship between variable y_t and x_t in which y_t is called explained variable (or contingent variable and dependent variable), x_t is called explanatory variable (or independent variable), u_t is called random error term, b_0 is called constant term (intercept term) and b_1 is called regression coefficient.

Changes of y_t , b_0 and b_1 are called regression parameters, which are usually unknown, required to be estimated. t represents ordinal number. As t indicates time ordinal number, x_t and y_t are call time series data. When t indicates non-time ordinal number, x_t and y_t are called cross-section data. u_t includes numerous micro-factors that influencing changes of y_t except x_t . The changes of u_t are uncontrollable. The abovementioned model can be divided into two parts: ① $b_0 + b_1 \times x_t$ refers to non-random part; ② u_t refers to random part.

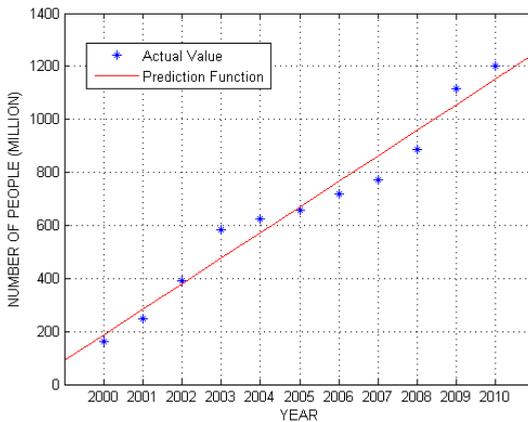
3 Establishment and Prediction of Unary Linear Regression Model of Total Tourism Arrivals

3.1 Data Analysis

Time series forecasting method refers to a method for establishing forecasting model by adopting scientific mathematical method according to the time series of historical data of the market phenomenon to extend the quantity of market phenomenon into the future forecast the future development change trend of the market phenomenon and determine the predictive value of the market [7]. With the international and domestic ice-snow tourism arrivals in Harbin as the predicted object, the relevant data between 2000 and 2010 is collected as shown by the Table 1. Draw the scatter plot represented by the Fig. 1 in the Table 1.

Table 1. 2000-2010, the total number of ice-snow tourism

| Year | T | Total number (million) |
|------|----|------------------------|
| 2000 | 1 | 160.80 |
| 2001 | 2 | 249.60 |
| 2002 | 3 | 391.23 |
| 2003 | 4 | 584.85 |
| 2004 | 5 | 626.30 |
| 2005 | 6 | 658.37 |
| 2006 | 7 | 719.02 |
| 2007 | 8 | 770.64 |
| 2008 | 9 | 886.49 |
| 2009 | 10 | 1114.50 |
| 2010 | 11 | 1202.70 |

**Fig. 1.** 2000-2010, the scatter plot and simulation linear model of the total number of Harbin ice-snow tourism

From observation and analysis of scatter points in the Fig. 1, it can be seen that: the change rule of the overall time series data basically shows up the feature of long-term rising change trend except there is a bigger increase in the ice-snow tourism in Harbin due to the influence of 2003 SARS incident and convening of 2009 World University Winter Games. Therefore, this paper describes changes in phenomenon to further perform forecasting by adopting the linear trend model in the trend extrapolation.

3.2 Establishment of Unary Linear Regression Model of Total Tourism Arrivals

Linear type change trend indicates the data of time series increases or reduces according the same quantity of each phase and shows up approximate linear rising or declining trends, that is, time series data has long-term linear change trend. Through observation and analysis of Fig.1, the distribution of 11 data of time series is like a linear approximately. Thereby, this paper adopts unary linear trend prediction model.

① Using diagram method to establish linear model

The result of established linear model is as shown by the Fig.1, with analyzing the time series trend by using the data in the Table 1.

② Using analysis tool to establish linear model

The results of analysis of trend of the time series, regression result calculation and parameter test are as shown by the Table 2. The linear model from the Table 2:

$$\text{RENSHU} = 96.53800 \times T + 90.27200 \tag{2}$$

Table 2. Times series parameter calculation and test results of tourism total number

| Dependent Variable: RENSHU | | | | |
|----------------------------|-------------|-----------------------|-------------|----------|
| Method: Least Squares | | | | |
| Date: 04/05/11 Time: 20:30 | | | | |
| Sample: 2000 2010 | | | | |
| Included observations: 11 | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| T | 96.53800 | 6.245860 | 15.45632 | 0.0000 |
| C | 90.27200 | 42.36149 | 2.130992 | 0.0619 |
| R-squared | 0.963695 | Mean dependent var | | 669.5000 |
| Adjusted R-squared | 0.959661 | S.D. dependent var | | 326.1556 |
| S.E. of regression | 65.50714 | Akaike info criterion | | 11.36516 |
| Sum squared resid | 38620.66 | Schwarz criterion | | 11.43751 |
| Log likelihood | -60.50838 | Hannan-Quinn criter | | 11.31956 |
| F-statistic | 238.8977 | Durbin-Watson stat | | 1.029030 |
| Prob(F-statistic) | 0.000000 | | | |

③ Model analysis

From the perspective of linear model, time series shows the positive linear trend according to time fluctuation. It is explained that the No. of people participating in ice-snow tourism in Harbin is on the rise year on year, which is in line with the actual situation. The fit effect of the entire model is very good and the coefficient of determination reaches 93.3965%. F test statistic of overall model linear and T test statistic of regression coefficient of the time term is statistically remarkable.

Additionally, from the line graph (Fig.2) of actual data of tourism arrivals, model fitting data and residual series, it can be seen that the fit effect is relatively good. The residual data of the rest years changes within one time standard deviation except the residual data of 2003.

From the foregoing analysis, it can be seen that time series shows up positive linear trend according to time fluctuation, which is in line with the actual situation. The fit effect of the Eq. 2 is pretty good. It can be used to predict the total ice-snow tourism arrivals in the following years in Harbin.

3.3 Prediction of the Total Number of Harbin Ice-Snow Tourism

① Point prediction

Using prediction model, the respective quantitative values (unit: million) of the total tourism arrivals in Harbin between 2011 and 2014 predicted by the Eq. 2 is:

$$Y_{2011} = 96.5380 \times 12 + 90.2720 = 1248.728, Y_{2012} = 96.5380 \times 13 + 90.2720 = 1345.266,$$

$$Y_{2013} = 96.5380 \times 14 + 90.2720 = 1441.804, Y_{2014} = 96.5380 \times 15 + 90.2720 = 1538.342.$$

Thereupon, the obtained development trend of total ice-snow tourism arrivals in Harbin between 2000 and 2014 is as shown by the Fig.3.

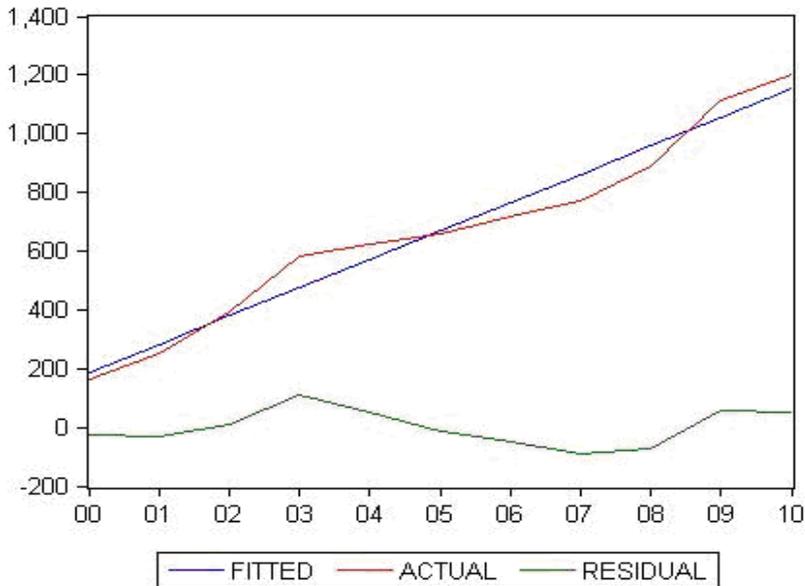


Fig. 2. The line graph of the actual data, fitted data and residual serial about the tourist arrivals



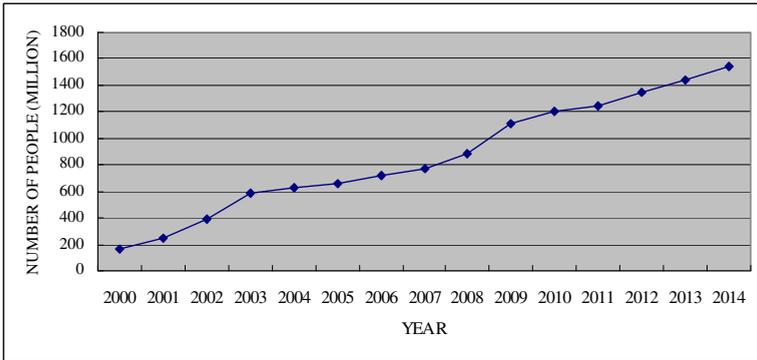


Fig. 3. 2000-2014, the trend of the total number of Harbin ice-snow tourism

② Interval prediction

The approximate confidence interval with confidence level of 95% obtained by using the standard deviation that predicts point estimate \hat{y} and regression analysis is: $(\hat{y} - 2S_y, \hat{y} + 2S_y)$.

For S. E. of regression $S_y=65.50714$ and under the condition of 95% confidence level, namely $\alpha = 0.05$, the respective approximate prediction value (unit: million, namely, the prediction interval of the total ice-snow tourism arrivals in Harbin between 2011 and 2014) of the confidence interval of the predictive values of years between 2011 and 2014:

$$\begin{aligned}
 Y_{2011} &= 1248.728 \pm 2 \times 65.50714 = 1117.714 \sim 1379.742, \\
 Y_{2012} &= 1345.266 \pm 2 \times 65.50714 = 1214.252 \sim 1476.280, \\
 Y_{2013} &= 1441.804 \pm 2 \times 65.50714 = 1310.790 \sim 1572.818, \\
 Y_{2014} &= 1538.342 \pm 2 \times 65.50714 = 1407.328 \sim 1669.356.
 \end{aligned}$$

The schematic diagram of upper and lower limits of the prediction value of total ice-snow tourism arrivals in Harbin between 2011 and 2014 is as shown by the Fig.4.

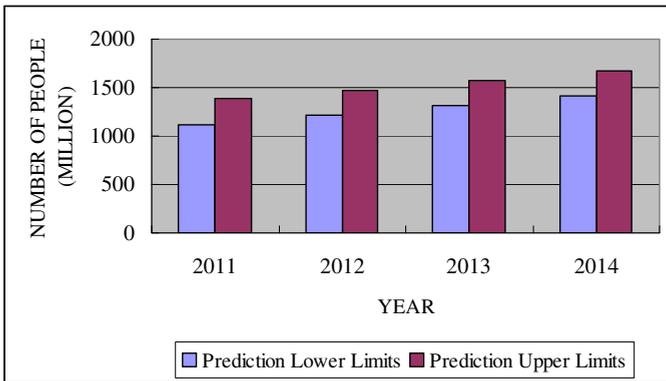


Fig. 4. 2011-2014, predicted value of the total number of Harbin ice-snow tourism

4 Summary

This paper analyzes the data change tend of the market demand of ice-snow tourism between 2000 and 2010 in Harbin by using analysis of time series, describes the changes of phenomenon by selecting the linear trend model in the trend extrapolation, establishes unary linear regression model of the total ice-snow tourism arrivals in Harbin and predicts the ice-snow tourism arrivals between 2011 and 2014 in Harbin. In accordance with the prediction result, it explains that the model has quite good prediction accuracy and it is realistically & theoretically significant to the development of ice-snow tourism market, rational formulation of tourism policy and regulation of the tourism development.

References

1. Ioannides, D., Apostolopoulos, Y.: Political Instability War and Tourism in Cyprus: Effects Management and Prospects for Recovery. *Journal of Travel Research*, 51–56 (1999)
2. Greenidge, K.: Forecasting Tourism Demand: An STM Approach. *Annals of Tourism Research*, 98–112 (2001)
3. Chu, F.-L.: Forecasting Tourism Demand: A Cubic Polynomial Approach. *Tourism Management*, 209–218 (2004)
4. Durbarry, R., Sinclair, M.T.: Market Shares Analysis: The Case of French Tourism Demand. *Annals of Tourism Research*, 927–941 (2003)
5. Nada, K., Witt, S.F.: Leading Indicator Tourism Forecasts. *Tourism Management*, 503–510 (2003)
6. Cayey Goh, R.L.: Incorporating the Rough theory into Travel Demand Analysis. *Tourism Management*, 511–517 (2003)
7. Pye, S.S.: A Linear Expenditure Model of Tourism Demand. *Annals of Tourism Research*, 443–454 (1991)

ZPD Incidence Development Strategy for Demand of Internet in Business – Teaching of New and Old Comprehensive Regional Higher Education Institutes

Hong Liu

School of Business, Linyi University, Linyi 276005, P.R. China
Hong7386@126.com

Abstract. Comprehensive regional higher business education institutes need ICT systems to facilitate the exchange of ideas and information about agriculture. Many regional higher education institutes (CRHEIs, mainly be regional comprehensive university) have been founded from 1998 in China. Now many students study in those CRHEIs. How to improve the equality of those CRHEIs become more and more important. In this paper, we analyses internet-based teaching and learning issues at CRHEIs level and old comprehensive higher education institutes (OOCNHEIs) level. Using normative Delphi method, we discussed internet-based issues and get that internet integration in CRHE need to be in lined with proper strategy in order to get their true benefits.

Keywords: ZPD incidence development strategy, Internet, Delphi, comprehensive regional higher education.

1 Introduction

Comprehensive regional higher business-science education institutes need internet systems to facilitate the exchange of ideas and information about agriculture[1]. It is recognized by many researchers that use of internet and application is an important business-skill for any worker. Internet usage at institute helps students to continue their learning beyond the classroom[2]. Internet -skilled teachers in comprehensive regional higher education institutes should adopt right pedagogical tools and practices in their teaching and enable their students embrace these new technologies. With time passed, internet has changed the whole process of teaching and learning. But some teachers lag far behind the others in adopting internet innovation[3]. However, change in professional domain, internet innovation, and teaching methods is ineluctable.

In china, many regional higher education institutes (mainly be regional college) have been founded from 1999 in China. From 2006, most of those regional higher education institutes promoted to be comprehensive regional higher education institutes (CRHEIs) in succession. Students in those CRHEIs become more and more. Those CRHEIs should improve their quality of teaching. In this study, teaching /learning with and without the help of internet in CRHEIs and comprehensive nationwide higher education institutes (OCNHEIs) is explored. Their ability gaps in tackling and solving problems are recorded, and therefore proper strategy or mechanism can be figured out to reduce these ability gaps to a minimum.

In the following, we propose college of business in Linyi University(In the past, Linyi University be Linyi Normal University) as an example on our research of CRHEIs in PR China. In this research, the personnel of college of business science in Linyi University are asked to devise development strategy with the help of normative Delphi technique. The purpose of these development strategies is to increase teachers/learners experiences by the use of internet in CRHEIs, and therefore improve output of CRHEIs. The School of Business Science in Linyi University now has 130 faculty members, of whom there are 11 Professors, 47 Vice Professors, 32 with doctoral degrees, 10 in the process of completing doctoral degree programs, and 73 with master's degrees. The school now has 6030 full-time undergraduates and non-degree students.

We propose Qingdao University as an example on our research of OCNHEI in PR China. The School of Business Science in Qingdao University now has 118 faculty members, of whom there are 39 Professors, 40Vice Professors. 88% of the Professors and Vice Professors had achieved a PhD degree. The organization of this paper is as follows. In section I , we propose the introduction of internet in CRHEIs. In section II, we discuss the data analysis of college of Business science in Linyi University and Qingdao University. In section III, we present our conclusion that development strategy for requirement of internet proposed by panelists.

2 Data Analysis

In this paper, we get the following data of college of Business science in Linyi University and Qingdao University as the way of [4]. According to normative Delphi technique, a questionnaire was prepared and hand-delivered to the 230 members of staff in college of business science in Linyi University, and 180 of panelists answered the questionnaire. There are 126 members of staff in college of business science in Qingdao University. 78 of panelists answered the questionnaire. Those data showing ZPD gaps obtained through the questionnaires is shown in Table 1 and Table 2.

The concept of Zone of proximal development (ZPD)was coined by [5], ZPD gap is the difference between future/maximum and current state of any development/use of internet. In this research, teaching/learning with and without the help/use of internet at CRHEIs and OCNHEIs level are explored and their ZPD gaps are recorded, so that a strategy that can reduce these gaps to a minimum can be devised. Please refer to Table 1 and Table 2 for issues.

1. A teacher prepare/develop class lecture by reading online, searching information from internet before his class lecture. ZPD gap(3.04) of CRHEIs and ZPD gap(2.15) of OCNHEIs are recorded which shows levels of teachers using internet tools and applications for these tasks.

2. For developing course material, sharing educational content, communication between teachers and outside using internet and applications, ZPD gap (1.68) of CRHEIs and ZPD gap(2.06) of OCNHEIs are obtained.

3. Checking exam papers, recording grades, and announcing results takes a lot of teachers' time. ZPD gap of teachers of CRHEIs and OCNHEIs are 1.39 and 0.51 respectively.

Table 1. ZPD gaps of Linyi University(New Comprehensive Regional Higher Education Institutes) and Qingdao University(Old Comprehensive Regional Higher Education Institutes)

| | issues | ZPD gaps of Linyi University | ZPD gaps of Qingdao University |
|---|--|------------------------------|--------------------------------|
| 1 | Prepareng/Developing Class Lecture | 3.04 | 2.15 |
| 2 | Presenting/Sharing Material | 1.68 | 2.06 |
| 3 | Assessing Student's Learning | 1.39 | 0.51 |
| 4 | Academic Research | 1.73 | 0.93 |
| 5 | Social Networks | 1.86 | 1.02 |
| 6 | Use of Common internet | 0.08 | 0.08 |
| 7 | Use of research and educational internet | 1.18 | 1.12 |

4. For finding research information, communicating with researchers, and sharing ideas with other teachers, ZPD gap (1.73) of CRHEIs and ZPD gap(0.93) of OCNHEIs are obtained.

5. Teachers quest for knowledge using social networks and learner forums. ZPD gap of teachers of CRHEIs and OCNHEIs regarding these tasks are 1.86 and 1.02 respectively.

6. Very small ZPD gap(0.08) of both CRHEIs and OCNHEIs are recorded for use of common internet such as MS office, web browsers, e-mail, search engines etc.

7. ZPD gap(1.18) of CRHEIs and ZPD gap(1.12) of OCNHEIs are recorded for use of research and educational internet such as translation software, online mapping, demonstration and scenario tools, web conferencing tools, and course management tools etc.

Now we discuss the above data as the way of [6]. In comprehensive regional higher education institutes, teachers usually perform a number of tasks. It takes a teacher a lot of time to prepare a class lecture for his day-to-day teaching task. A teacher must develop course material and sharing educational content using internet and applications. If a teacher is effective at checking exam papers, recording grades, and announcing results, his keeping record tasks will become much easier. Administrative tasks require a teacher to spend time keeping student records, issuing book, and supporting students. There are many internet and applications that a teacher can use while finding research information and sharing ideas with other teachers.

The spread of internet and applications is considered as necessary in CRHEIs of developing countries; thus they can drive in pedagogical challenges coming from latest development. However, few strategies have been devised to solve these issues in developing countries (especially PRC). Accordingly we try to devise a strategy including some important measures for internet enhancement in this study.

3 Conclusions

Through Comparing the ZPD gaps of Linyi University and Qingdao University, some development strategy for requirement of internet are listed below, most proposed by panelists in college of Business science in Linyi University.

1. Internet for teaching

RHE teachers need to be proficient in usage of internet and applications to work effectively. Some of the recommendations suggested by our panelists in this regard are: (1) teams comprising of teachers, educational advisors, and library staff need to be developed which develop course-based education content. (2) design of a persistent training program for faculty/staff in the use of internet.

2. Internet for gaining proper attainment of students

Some of the actions suggested by our panelists in this dimension are: (1) enable high speed internet access for management, faculty and administrative staff. (2) local internet needs to be developed and consummated. (3) students progress between key stages may be measured through management information system. (4) database of students records for data analysis should be designed to increase the current level of data sharing.

3. Internet for teachers' development

Some of the recommendations suggested by our panelists in this dimension are: (1) Internet training centers that fulfill training needs of teachers needs to be established and developed. (2) skills of internet should be trained to enable them to train faculty with major development and technological changes occur with time in the world of internet.

References

1. Information and Communication Technology Strategic Plan, 2005-06 to 2009-10. University of Oxford, http://www.ict.ox.ac.uk/strategy/plan/ICT_Strategic_Plan_March2007.pdf (retrieved August 2009)
2. Argyll and Bute Community Services: ICT Strategy for Education, <http://www.argyllbute.gov.uk/pdf/store/infotechnologystrategy> (retrieved August 2009)
3. Deshpande, P.: Connected – where next?: A strategy for the next phase in the development of education ICT in Bournemouth, http://www.bournemouth.gov.uk/Library/PDF/Education/Education_ICT_Strategy_2004_to_2009.pdf (retrieved August 2009)
4. Shaikh, Z.A.: Usage, acceptance, adoption, and diffusion of information & communication technologies in higher education: a measurement of critical factors. *Journal of Information Technology Impact* 9(2), 63–80 (2009)
5. Vygotsky, L.S.: *Mind in Society: The development of higher psychological processes*. Harvard University Press, Cambridge (1978)
6. Ahmed, Z.: ZPD Incidence Development Strategy for Demand of ICTs in Higher Education Insitutes of Pakistan. In: *The 3rd IEEE International Conference on Intelligent Information Technology Application*, vol. (11), pp. 170–174 (2009)

Analysis of Changes of Agricultural Tax System

Zhang Huiqing

XinZhou Vocational & Technical College, XinZhou, Shanxi, China
zhanghere@yahoo.cn

Abstract. Agricultural tax has a long history in China, for China's agricultural economy played an important role for economic development in China has made outstanding contributions. In this paper, developed by China after the founding of the agricultural tax system and its implementation, and change to do a sort of agricultural tax reforms summarized the content and features to the construction of new agricultural tax system, the realization of urban and rural integration Enlightening.

Keywords: Agricultural tax, Special agricultural product tax, Value added tax, Tax reform.

1 Introduction

Agricultural tax salary and the feudal monarch's military spending provided the main financial support, the agricultural tax as the feudal regime, the material basis for survival. After the founding of new China, agricultural industrialization and the construction of China's accumulation of money, food security and social stability, military people have played an important role.

2 Transformation of the Ancient Agricultural Tax

In slave society, the summer generation of "tribute" is the prototype of the agricultural tax, to the feudal society, Lu implementation of "early tax mu", began to levy land tax in kind specialty products, a sign of maturity of the agricultural tax, agricultural tax in China from behind with their mode of production has been 5,000 years of history, we have experienced several changes. Tang implementation of the "two-tax" created a precedent for tax reform in China, "two-tax" includes household tax and land tax, household tax levy refers to the household as a unit of land by each household to determine the number of taxable property; land tax is Rebellion Rebellion by the number of taxes levied. Collected each year with summer and fall. "Two-tax" merge the tax items, focused on the time, charges to tax all together into the two taxes among the central control of tax collection powers. Introduced in the late Ming Dynasty, "a whip", "a whip" the various lord, land tax and a variety of charges, and a land tax, a levy targeted to Rebellion, and be converted into silver to pay. "A whip" reforms to simplify, to a unified tax system, the role of social productivity and stability. Early implementation of the Qing Dynasty "farmland," the tax system, the poll tax and the

Rebellion tax integration, according to the number of unified collection liturgical Rebellion. "Farmland" simplify the tax principle of a single standard only on the land tax, and stability of the Qing government revenue.

Three times in the history of agricultural reform, basically to simplify, unify the tax system, strengthening the central financial authority, weakened local forces, began to reduce the burden on peasants have played, the role of promoting agricultural development, but it does not last long. Mainly due to the agricultural tax reform with a strong personal nature of the reform is not institutionalized.

3 After the Founding of New China, the Agricultural Tax System and Its Transformation

The liberation war, the agricultural tax base, and the liberated areas known as the "land tax", "public grain," began after the founding of a unified agricultural tax is called. June 3, 1958, the NPC Standing Committee passed the "Republic of China Agriculture Ordinance," established the basic agricultural tax policy principles, this is the first legal system of agricultural taxation. China's agricultural tax system is different from other industries, the establishment of separate taxes, including agricultural tax and special agricultural product tax.

Overview of the agricultural tax. Agricultural tax, is engaged in agricultural production, agriculture and the production of income units and individuals a tax levied. It is the country's participation in agriculture as a major form of income distribution. Its own characteristics in two ways, one is taxable income of a particular agricultural tax based on annual output of agricultural land tax as a tax basis, while others generally sales tax revenue, operating income as a tax basis; the other is to impose strong seasonal storage, due to the seasonal harvest of various crops is different from the general points the summer and fall of agricultural tax collection, tax year means a calculation, and other taxes generally on a monthly basis tax returns.

Basic elements of agricultural taxpayer who is the agricultural crops and to agricultural income units and individuals, primarily of agricultural production and management, then contracting households, farms and other enterprises, administrative units and four other citizens; agricultural tax levied on agricultural income, including crop income, crop income and other crops of economic income; agricultural implement the proportional tax on agricultural income, the average rate of top-down regulations, taxpayers under the provisions of the average tax rate applicable tax rates; agricultural areas with an additional tax is imposed along, it belongs to village collective funds, the implementation of Village by village, township finance department is responsible for the accounting of special account, all expenses incurred for the village; agricultural tax is in production, should be collected in crop harvest season, sub-summer and fall collection, January 1 to August 31 levy for the summer period September 1 to December 31 as the sign of autumn; agricultural tax has been imposed food-based, the standard calculation of the local staple food, staple food in accordance with the different ratio other than the equivalent calculated as staple food.

Tax payable = taxable agricultural land area * yield * Standard annual rate of food price *

Additional amount of agricultural areas = the proportion of agricultural added tax payable *

Agricultural land tax is levied according to tax laws should be agricultural land, the land tax to be careful to verify the implementation of relevant policies and regulations. Annual production is based on the natural conditions of land and local assessment of general business conditions in normal years crop yield, assessed after a period of time should be kept relatively stable.

The role of agricultural. The accumulation of agricultural tax in construction funds for the state to regulate the rural economy, promoting agricultural development, consolidate the worker-peasant alliance, is of great significance, its role is mainly manifested in two aspects, one is the accumulation of capital, as early in the Revolutionary War, the agricultural tax revenue in fiscal revenue 80%, and is spending the funding regime, the revolutionary base areas and the main source of people's army; the early days, the average agricultural taxes accounted for 18.3%, is still the main source of state revenue; With the acceleration of industrialization, agricultural tax in the state finance declining proportion of income, but in some economically underdeveloped agricultural areas, agricultural tax is still the main source of local revenue; the other is to adjust the income, the role of agricultural adjustment of income, mainly through the implementation of regional differences in tax rates and tax incentives to be embodied. For a good natural conditions, income and more areas, more than the tax burden for the poor natural conditions, in areas with little revenue, tax less, for the minority, mountainous area, giving relief to take care of. Bridging the gap between rich and poor, has played an important role in social stability.

The agricultural tax reform. Current "Republic of China Agriculture Ordinance" has been running for over 40 years, the whole is formed under the planned economic system, is associated with the natural economy in rural areas, with the deepening of China's reform and opening up the rural economy has undergone tremendous change, making the current agricultural tax system and market economy is not suited, not only in the form behind, the burden is also not reasonable. State agriculture industry and other industries of national economy, a distinction between taxpayers treated separately, the aim is to reduce the burden on farmers, but the actual implementation of the policy process, there are many problems, the agricultural tax is the tax basis and tax Rebellion annual output, after approval, generally do not change, but in fact the peasants for the land of the Rebellion in the change, and farmers can not respond to a tax increase or decrease, addition, food inputs for the farmers to grow seeds, fertilizers, pesticides, agricultural machinery and equipment production and labor costs would be given compensation, agricultural policies have not taken into account. Grain farmers regardless of profit or loss, how much profit, how much loss, we must pay a fixed agricultural tax. Shows, agricultural tax reform is necessary.

The current rural tax reform, but in the present stage of reform measures under the conditions, it is difficult to solve some deep-rooted contradictions and problems. Farmers have long borne the burden of a large number of taxes, especially a variety of names of non-tax burden on many kinds of tired, unabated. Data indicate that in 2000 compared with 1990, farmers bear the total taxes and fees increased to 135.9 billion yuan from the 46.9 billion yuan, of which: agriculture, "four tax" from 88 billion yuan to 35.2 billion, 21.6 billion yuan from the village-level retention increased to 35.2 billion, townships fee from 11.7 billion to 268 billion, other charges from 48 billion

yuan to 27.4 billion, per capita commitment of farmers increased from 55.8 yuan taxes increased 168.4 yuan. Facts show that an excessive burden on farmers the problem is not fundamentally resolved. In rural areas to take tax cuts and tax policy is an important means to promote rural development, the current reform program, we will accelerate the abolition of agricultural tax, China from 2001 to gradually reduce the agricultural tax rate, abolition of agricultural tax within five years, and in Jilin and Heilongjiang provinces to implement exempt the agricultural tax reform; second is to continue to levy value added tax included in the scope of agricultural products, and implement tax policy; third is included in the income of agricultural production and operation scope of personal income tax collection. Engaged in agricultural production and management from the current income situation of farmers, the vast majority of farmers income of less than a threshold, the basic will not increase the burden on farmers, personal income tax is limited to a relatively small number engaged in the kind of planting, breeding, etc. large income of the farmers.

Overview of special agricultural product tax. Special agricultural product tax is the state in our country engaged in agricultural production or acquisition of special products, special products to agricultural income units and a tax levied on individuals, it is attached to the agricultural tax is part of the agricultural tax. November 12, 1983, the State Council promulgated the "specialty agriculture and forestry revenues levied on agricultural tax provisions," marked the formal introduction of agriculture and forestry product tax. This stage is characterized by the development of agriculture and forestry around the specialty tax levy means of taxation, collection exploring the work of a certain trial, the central government was not to the tax levy over the task assigned specialty agriculture and forestry. March 13, 1989, the State Council issued the "Circular on Further Improving agricultural tax levied specialty agriculture and forestry work," requiring local comprehensive specialty tax levy forestry, agriculture and forestry product tax collection methods and made several important improvements, expanding the scope of taxation improved tax measures, the provisions of the tax rate, levy a separate task assigned, included in the local budget, agriculture and forestry product tax into a comprehensive collection phase. January 30, 1994, "the State Council on the agricultural tax levied on agricultural specialty income requirement" marks the reunification of special agricultural product tax, tax on special agricultural products of this phase of the scope of taxation by the unified income scale. March 2000, the CPC Central Committee State Council issued the "Regulations on the work of rural tax reform," which called special agricultural product tax policy adjustment, according to the agricultural tax and special agricultural product tax levied on the principle of non-overlapping, non-agricultural tax in the land on the production of agricultural and special products, the continued imposition of special agricultural product tax, land tax on agricultural production in the agricultural special products, collected only agricultural tax.

The basic elements of agricultural specialty tax. Taxpayers of special agricultural product tax is in China in the production or acquisition of taxable agricultural specialty units and individuals, special agricultural product tax is levied on agricultural special product revenue, including its tobacco revenue, income horticulture, aquaculture revenue, forest income, livestock income, Shijun income, precious food income, other income.

Special agricultural product tax payable = real income (acquisition value) * tax rate

Added Tax on special agricultural products tax payable = taxpayers - the original burden of agricultural tax

The role of agricultural specialty tax. Special agricultural product tax collection help to change the agricultural product and food and other basic agricultural tax burden imbalance, unreasonable conditions and promote the rational allocation of agricultural resources; levy taxes on special agricultural products which helps to balance the structure of agricultural production, through the tax rate development, to guide and promote agricultural in favor of special products to meet the needs of the people, in line with the direction of national industrial policy; levy taxes on special agricultural products is an important local revenue source.

Reform of agricultural specialty tax system. Since the introduction since the special agricultural product tax has become an important local taxes, especially in the rural revenue account for a large proportion of local finance has played a positive role. To further reduce the burden on peasants, and promote stability and development of agriculture, the state decided to cancel the special agricultural product tax, a step can be difficult to cancel the first reduction of items, lower taxes, reduce the scope, phasing out of special agricultural product tax. Rural tax reform since 2000, on the part of the collection of special agricultural product tax, adjusted taxable items and tax rates, introduced in 2003 and further initiatives to abolish taxes on special agricultural products.

4 Current Agricultural Tax

December 29, 2005 the Nineteenth Meeting of the Tenth National People's Congress voted and passed the Abolition of the "PRC Agriculture Ordinance," the decision to have played an important role in the agricultural system in the implementation of the nearly half a century later, and finally from the historical stage. According to the law abolished the agricultural tax regulations, which is China's agricultural development and the world practice and the landmark event, the Chinese agricultural tax, to take "less, more, deregulation" policy, no doubt adapt to the times, adapt to the development of world economic integration trend. The abolition of agricultural taxes, marking the unification of urban and rural taxation taken to the landmark first step.

After phasing out the agricultural tax on individuals engaged in agricultural production and operation of personal income tax generally does not have the conditions, China has more than 200 million farmers, decentralized operation, the majority of low income households, very few large agricultural operation with higher income, personal income tax the high cost of operation is difficult. Therefore, China's tax policies and regulations of the individual operating income achieved in agricultural production, "during" personal income tax temporarily.

Gradual abolition of agricultural tax, agricultural products included in the scope of VAT, the state agriculture-related frequency distribution of a series of preferential policies to further reduce the burden on peasants. VAT is the sales of goods in our country, provision of processing, repair, replacement services, and imported goods units and individuals to a tax levied on the value added. Belongs to the Chinese value-added agricultural products tax, current tax policy, agricultural production units and individuals of primary agricultural products sold, are exempt from VAT, is

essentially excluded from the value-added agriculture outside the system. Countries has been the introduction of VAT from the current practice of situation, integration of agriculture into value-added tax is a more specific range of thorny issues. Solve the double taxation from the point of view, value-added tax should be included in the scope of agriculture; but collection and management point of view, agricultural producers households, widely distributed, accounting is not perfect, there are a lot of their value added tax difficult.

VAT is currently the country, the processing of agricultural products mainly include the following several ways, each with advantages and disadvantages. (1) sharp zero rate for agricultural products, this approach solves the problem of double taxation, but bring a lot of tax, the operation is cumbersome; (2) low tax rates on agricultural products, agricultural production inputs to be with their input tax deductible agricultural output tax equal to zero offset between the two. Although this approach avoids the double taxation, but a uniform low tax rate can not be done so that the tax burden on agricultural producers of each zero. (3) EU regulations on taxation according to the normal taxation of agricultural producers have difficulty to take "increase compensation law" provides tax farmers have difficulty selling their produce to the buyer to receive a percentage increase, from compensation, This ratio is formulated by the fare increase as the purchaser's input tax to be deductible. (4) most countries adopted a VAT exemption for agricultural producers directly to the practice, in order to support agricultural production and reduce the difficulty of tax collection, China has adopted this approach.

5 Conclusion

China system is an important part of the economic system is one of development and change of economic system and closely linked with the social, political and economic conditions change, the tax system also constantly adjust and reform, tax policies in different periods have different for the then social stability and economic development level has played a positive role. Since reform and opening, China's agricultural tax system has made great progress, but in the face of the socialist market economic situation and the challenges of joining WTO, the existing tax system but also reflect some aspects are not perfect and problems need to further reform and improvement so as to promote the health of agriculture and rural development.

After the abolition of agricultural tax, China entered the era of non-agricultural tax, agricultural tax was rescinded just does not make all of the "three rural" issue solved, agriculture and rural work must also address the current problems, how to build a new agricultural tax system is currently urgent need to address the problem, China's accession to WTO, agricultural reforms are bound to face a broader international environment.

Construction of new agricultural tax system should be considered in two ways, first, the design of the agricultural tax system with international standards, independent of the agricultural tax abolition of agricultural and other industries, impose the same tax and preferential policies designed to eliminate the taxation of real existence urban and rural isolation, and fundamentally change the dual tax structure, will rural areas and farmers, agricultural taxation into a unified tax system in the country; the second is

fiscal adjustment in the new environment policies on agricultural subsidies, as supporting tax reform measures.

References

1. Wang, X.: The direction of tax reform. Taiyuan China Agricultural University (June 2009)
2. Key reform of agricultural economy and Kay Law (October 2010)
3. Li, R.: Tax assessment of agricultural technology start-ups (September 2009)

Application Research of Smart Water Systems Based on I2C Bus

Shi Binbin¹, Hu Zhifen¹, and Dai Minli²

¹ Dept.of Electrical Engineering, Suzhou Vocational University, Suzhou, China

² Dept.of Computer Engineering, Suzhou Vocational University, Suzhou, China
smn549@21cn.com

Abstract. A highly accurate smart water system with low power consumption is realized based on Inter-Integrated Circuit. It can drive digital temperature sensor TMP275 with I2C module. The actual temperature is showed by the decimal numerical code tube. The hardware circuit and software design process are given. It provides a good reference for the smart home systems with low power, low cost and excellent performance.

Keywords: Smart Home, Inter-Integrated Circuit, TMP275, Temperature Measurement.

1 Introduction

With the fast development of the electronic industry since the dawn of the 21st century, the smart home systems are entering people's lives.

With the gradual development of sensor technology in recent years, the smart home systems have received more and more attention from the research community. With the help of temperature sensors, the smart water system can detect the water temperature coming out of a tap and indicate the water temperature level with different colors based on the measured data. As a result, better living conditions can be provided with less energy consumption and more security assurance. In this paper, a smart water system with low power consumption, low cost and high accuracy is designed based on the digital temperature sensor chip TMP275[1].

2 System Design Outline

Currently, most smart water systems available on the market are expensive and energy consuming. In this project, functional modules and controlling modules with good performance price ratio and low power consumption are used to design the smart water system. The MSP430F247 from TI is used as the master chip. The highly accurate temperature sensor is used to measure the temperature. The I2C bus is adopted for the communication between the two chips[2,3]. The measured temperature is classified into three classes, which are then indicated by an LED with three different colors. The smart water system has very low power consumption and a low cost. It is simple to implement and easy to maintain. It also has a very high accuracy.

3 System Hardware Design

Temperature detection circuit. The TMP275 chip, constituting the sensor module, acts as the slave module. It can measure the temperature and then transmit the data to the master module through the I2C.

The TMP275 chip was introduced by TI as a low power consuming temperature sensor with digital output. It's accuracy is around $\pm 0.5^{\circ}\text{C}$. It is well suited for temperature measurement in areas such as communication, consumer electronics and industrial applications[4].

The high accuracy of TMP275 enables more efficient heat ventilation and power management. It's low power consumption can elongate the battery life and minimize self-heating. It's accuracy is within $\pm 0.5^{\circ}\text{C}$ when the temperature is ranged is $-20^{\circ}\text{C} \sim +100^{\circ}\text{C}$. The dual line serial port of the chip is compatible with I2C. The chip comes in a eight-pin MSOP package. Other features include: $50\ \mu\text{A}$ low current flow, nine to twelve bits programmable resolution, $0.1\ \mu\text{A}$ power down current mode, excellent stability across the whole temperature range and a broad operating temperature range from -40°C to $+125^{\circ}\text{C}$. Additionally, the chip supports up to eight different addresses, so that high flexibility is made possible in the design of the interface bus. The pinout diagram of TMP275 is shown in figure 1.

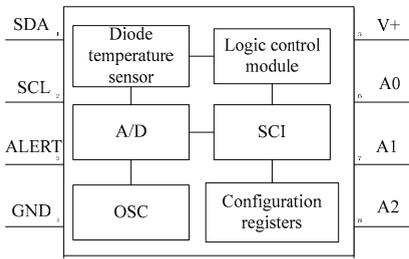


Fig. 1. TMP275 pinout diagram

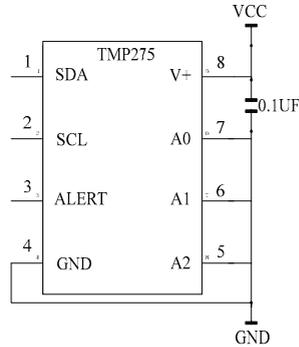


Fig. 2. The peripheral circuit schematic of TMP275

The dual line serial port of TMP275 (pins SCL and SDA) are compatible with the I2C bus. Therefore, it can be directly connected to the I2C bus. When the measured temperature exceeds the highest temperature or drops below the lowest temperature as set by the user, the pin ALERT can output a high or low signal. The above mentioned three pins need pull-up resistors in implementation. The pins A0, A1, A2 can be connected to either the ground or the VCC, there by determining the chip address. The peripheral circuit of TMP275 is shown above in figure 2. The chip is powered through the V+ pin and the GND pin. The $0.1\ \mu\text{F}$ capacitor is used to filter the power supply. The pins A0, A1, A2 are connect to the ground, then by determining the chip address as 1001000.



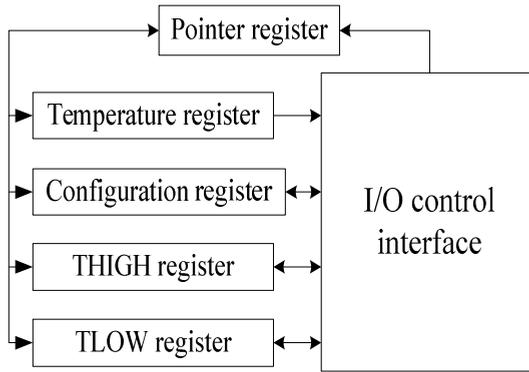


Fig. 4. The internal register structure diagram

The operation mode of TMP275 is determined by the configuration register. The configuration data format is shown below in table 1.

Table 1. Configuration register tables

| D7 | D6 | D5 | D4 | D3 | D2 | D1 | D0 |
|----|----|----|----|----|-----|----|----|
| OS | R1 | R0 | F1 | F0 | POL | TM | SD |

Under normal operation conditions, when the measured temperature falls outside the user-defined range, the TMP275 would choose either the comparison operation mode or the interruption mode depending on the status of TM in the configuration register. Under the comparison mode, if the measured temperature exceeds THIGH for n consecutive times, the comparator asserts the ALERT signal to notify the master that the operating temperature is abnormal, the ALERT signal deasserts only when the temperature falls below the TLOW for n consecutive times. The comparison mode is the default mode under normal conditions. Under the interruption mode, the ALERT signal deasserts only when a register operation occurs or the device enter the shut down mode[6]. Temperature tolerance test can be carried out under this mode.

Additionally, a shut down mode can be chosen for the device to save power. When this mode is chosen, the device shuts down automatically when the current temperature conversion is completed. The current under this mode is only 1µA[7]. The next temperature conversion is initiated only when the OS bit is set in the configuration register. This mode is configured by the SD bit in the configuration register.

I2C communication interface subroutine. The two data lines of TMP275 are compatible with I2C protocol. It can only act as a slave device. Both the fast mode(1~400 kHz) and the high speed mode (1 kHz~3.4 MHz) are supported.

I2C bus overview. The master device initiates a data transfer where as the slave device is controlled by the master. The master generates the serial clock(SCL) to control the bus and the start and stop conditions. Data can be transferred only when the



bus is not busy. During a transfer, the SDA must stay unchanged when the SCL is high. The SDA can change only when a start or stop signal arrives.

As a slave device, the TMP275 can start operating only when a start signal has been received. If the received address is correct, an acknowledge signal is sent out and a read/write operation is initiated depending on the R/W bit. When the stop signal arrives, all the operations are terminated.

Slave device receive mode. Under the receive mode, the master device sends to the TMP275 its address and status information first(R/W=0), and sends out the data to be written in the address pointer register. The next few bytes are written into the registers according to the content of the pointer register. For every data successfully received, the chip would send out a confirmation signal. The master terminates the data transfer through sending out an end signal.

Slave device transmit mode. Under this mode, the master device sends to the TMP275 its address and status information first(R/W=1), and then reads the data designated by the address pointer register. For every data successfully received, the chip would send out a confirmation signal. The master terminates the data transfer through sending out an end signal.

Display module subroutine. Upon receiving the water temperature signal from the sensor, the master chip classifies the water temperature into three class. The master then determines the port to be driven. The ports P6.6, P6.5, P6.4 drive blue, green and red lights respectively. The three lights are all installed at the water tap and are water-proof. When different lights are lit, the water would show different colors. By parts of the codes to control the lights are as follows.

```
P6DIR |= 0x70;    //The control port is set to output
// LED control program
if (RxWord < 0x1400) // The current water temperature < 20°C?
P6OUT = 0x40;    // Green light
else if (RxWord < 0x1E00) // 20°C < The current water temperature < 30°C?
P6OUT = 0x20;    // Blue light
else if (RxWord > 0x1E00) // The current water temperature > 30°C?
P6OUT = 0x10;    // Red light
```

Additionally, the current temperature can be shown directly on the four bits eight segment digital tube display. In order to ensure proper delay and making use of the “persistence of vision” effect of the human eye, the tube display software is based on an interruption process driven by a timer. Each interruption would refresh the display.

5 Summary

As part of a smart home system, the smart water system can improve people's living condition. In this paper, a smart water system is proposed and implemented with low power consumption and low maintenance cost. Field practice has shown that the chip TMP275 is highly efficient and can function as expected. Using the system as a reference design, a master chip along with several temperature sensor chips can be used to build other environment surveillance and controlling system in the family, and very high measurement accuracy can be achieved thereby.

References

1. Marwedel, P.: Embedded System Design, pp. 10–50. Springer, Heidelberg (2010)
2. Nagy, C.: Embedded systems design using the TI MSP430 series. Newnes, 95–100 (2003)
3. MSP430X24X Data Sheet Rev2 (2009), <http://www.ti.com/>
4. TMP275 Data Sheet Rev1 (2008), <http://www.ti.com/>
5. Li, Z.-X., Li, T.-Z.: Development of A Soil Temperature Monitor System Based on MSP430. Hubei Agricultural Sciences 49, 2241–2243 (2010)
6. Qiu, Z.: Temperature measurement instrument based on MSP430F247 and TMP275. Electronic Design Engineering 17, 3–6 (2009)
7. Wouters, P., De Cooman, M.: Injectable biotelemetry CMOS chip for identification and temperature measurement system. In: Proc. Solid-State Circuits Conference, pp. 45–48 (June 2010)

Evaluation of Mountain Road Influences on Driving Fatigue

Xiao Jinjian^{1,2,3}, Huang Simin¹, and Wang Yunsong²

¹ Department of Industrial Engineering, Tsinghua University, Beijing, China

² China National Institute of Standardization, Beijing, China

³ Chang'an University, School of Automobile, Xi'an, China

xiao.m.li@sohu.com

Abstract. Driving fatigue was caused by complicated road alignments in mountains. In order to evaluate road environments on driving fatigue, the driver subjective feeling of tiredness, driving operating and psychophysiology records, and their sub-factors were adopted as evaluation factors. on the basis of group multicriteria decision making and fuzzy analytical hierarchy process (AHP) method, the driving fatigue hierarchic fuzzy evaluation model was established, and a simplified scale method was adopted to generate the judgment matrices meeting the consistency checking firstly. the professional drivers' opinions were aggregated to form triangle fuzzy numbers with fuzzy Delphi method. Finally different influences on driving fatigue were ranked with total utility values of fuzzy numbers and the optimal alternative could be selected. An example was given, under this approach, show that the decision making process was a systematic and practical method for evaluation of the driver selection.

Keywords: driving fatigue, group multicriteria decision making, fuzzy analytical hierarchy process (AHP), fuzzy ranking.

1 Introduction

Road environment in mountain area is complicated, there are more curve-slope section, those reasons like rather small curve radius, bigger gradient, short and changeable sighting distance cause drivers heavier working load, relevantly it leads to driving fatigue phenomena [1,2]. Driving fatigue is leading factor and key reason that leads to big traffic accident in passenger transportation enterprises. Traffic accidents caused by driving fatigue are 6.8% of more than 10-passengers bus accidents and 7.2% of the death-toll in recent years [3]. Hence research in different changeable rules of driving fatigue is meaningful, and it is of application value in decision of driver selections in the special operating roads section. In order to fully evaluation driving fatigue of passenger-bus drivers, multi index fuzzy evaluation integrated with multi layers and factors must be established.

In regards of problems above mentioned, ranking with utility values of fuzzy number [4,5] and group multicriteria decision making [6,7] are integrated with driving fatigue evaluation in mountain driving. Hence it assures justification and credibility of

decision process, obtains accurate evaluation order result and provides reference for comparable priority selection of professional driver in specific running roads sections.

2 Structure Model of Driving Fatigue Fuzzy Evaluation

This paper adopts fuzzy evaluation sequential structure model, shown in fig. 1.

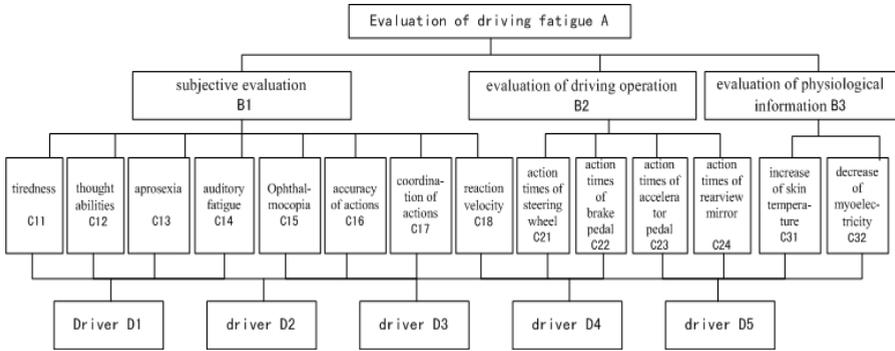


Fig. 1. Structure model of driving fatigue fuzzy evaluation

Target layer of subjective evaluation is drivers’ subjective feeling of driving fatigue, and Target layer of objective evaluation mainly consists of evaluation index of driving operating records and physiological information monitoring. In detail, it covers sub index forming criteria layer and sub criteria layer, in which thought abilities, accuracy and coordination of actions and ophthalmocopia are subjective evaluation index in driving, and adjustment times of steering wheel, action times of brake pedal, observation times of rearview mirror, and adjustment times of accelerator pedal are objective evaluation index through drivers’ control frequency during driving. In addition, drivers’ skin temperature and myoelectricity objectively reflect affected falling levels by driving fatigue, it is of reference value[8,9]. The last layer is selected drivers to meet requirements of the operating road and running bus type.

3 Judgment Matrix of Criteria and Integrating of Fuzzy Weights

Firstly relative importance of criteria levels is evaluated, secondly each expert compares criteria in pairs, and then builds comparison matrix of priority relationship.

$$B^t = (b_{ij}^t)_{m \times n} \quad (t = 1, 2, \dots, T)$$

, T is number of selected evaluators, where:

$$b_{ij}^t = \begin{cases} 1.0 & b_i = b_j \\ 0.5 & b_i < b_j \\ 0.0 & b_i > b_j \end{cases} \quad (1)$$

In order to transfer comparison matrix B' to the normal relationship judgment matrix $\overline{B'}$, elements of B' are summed in rows, it is written as:

$$r_i^t = \sum_{j=1}^n b_{ij}^t \quad (i = 1, 2, \dots, n) \tag{2}$$

$$\overline{b_{ij}^t} = \frac{|r_i^t - r_j^t|}{2n} + 0.5 \tag{3}$$

Thus matrix B' is transferred to normal relationship judgment matrix $\overline{B'}$. Factors of $\overline{B'}$ are summed in rows, and then normalize them in column, to get important evaluation W_i^t of each expert. Detailed decision methods of criteria weight are as written:

$$W_i = (O_i, P_i, Q_i) \tag{4}$$

In formula (4), W_i is i rule weight of rule layer, where: $O_i = \min_i \{W_i^t\}$, $Q_i = \max_i \{W_i^t\}$, $P_i = \sqrt[T]{\prod_{i=1}^T W_i^t}$. Utilizing similar method and step, fuzzy weight of sub rule criteria under each rule criteria could be got.

$$W_{ij} = (o_{ij}, p_{ij}, q_{ij}) \quad (j = 1, 2, \dots, n_i) \tag{5}$$

In formula (5), W_{ij} is fuzzy weight of j sub rule criteria under i rule criteria; W_{ij} is number of sub rule criteria under i rule criteria.

4 Integration of Sub Driving Fatigue Evaluation

Each factor language variable in driver subjective fatigue evaluation adopts same fuzzy number, while drivers operating records and myoelectricity monitoring adopt different fuzzy number respectively to reveal difference that different sub factors play role in drivers fatigue influence, shown in the table 1.

In the meanwhile, in order to integrate experts' opinions and reduce decision making information loss, linear weight sum method is adopted, but not evaluation of average decision makers. Every selection evaluation method is as followings:

$$V_{ijk} = (W^1 \otimes V_{ijk}^1) \oplus (W^2 \otimes V_{ijk}^2) \dots (W^T \otimes V_{ijk}^T) = \left(\sum_{t=1}^T W^t e_{ijk}^t, \sum_{t=1}^T W^t f_{ijk}^t, \sum_{t=1}^T W^t g_{ijk}^t \right) = (e_{ijk}, f_{ijk}, g_{ijk}) \tag{6}$$

In formula (6), \oplus and \otimes stand for fuzzy plus and minus, V_{ijk} is fuzzy evaluation of $D_k (k = 1, 2, \dots, m)$ in j criteria; W^t is weight displaying expert P_t decision making;

V_{ijk}^t is evaluated value by expert P_t in sub criteria, that is $V_{ijk}^t = (e_{ijk}^t, f_{ijk}^t, g_{ijk}^t)$.

Table 1. Fuzzy number of language variable value

| factors | fuzzy number of each language variable value | | | | |
|-----------------------------------|--|----------------------|----------------------|----------------------|----------------------|
| | NB | NS | ZE | PS | PB |
| subjective fatigue | (0.00,0.0 0.0.25) | (0.00,0.2 5,0.50) | (0.25,0. 50,0.75) | (0.50,0. 75,1.00) | (0.75,1. 00,1.00) |
| action times of steering wheel | (0.00,0. 00,0.30) | (0.00,0 .30,0.66) | (0.30,0 .66,0.82) | (0.66,0 .82,1.00) | (0.82,1.0 0,1.00) |
| action times of accelerator pedal | (0.00,0. 00,0.20) | (0.00,0. 20,0.33) | (0.20,0. 33,0.66) | (0.66,0. 83,1.00) | (0.83, 1.00,1.00) |
| action times of brake pedal | (0.00,0. 00,0.33) | (0.00,0 .33,0.50) | (0.33,0. 50,0.83) | (0.50,0. 83,1.00) | (0.83,1 .00,1.00) |
| action times of rearview mirror | (0.00,0. 00,0.25) | (0.00,0.2 5,0.50) | (0.25,0. 50,0.75) | (0.50,0. 75,1.00) | (0.75,1 .00,1.00) |
| increase of skin temperature | (0.00,0 .00,0.17) | (0.00,0.1 7,0.33) | (0.33,0. 0,0.83) | (0.60,0.8 3,1.00) | (0.83,1. 00,1.00) |
| decrease of myoelectricity | (0.00,0 .00,0.25) | (0.00,0.2 5,0.66) | (0.25,0. 66,0.75) | (0.66,0. 75,1.00) | (0.75,1. 00,1.00) |

5 Evaluation of General Driving Fatigue

After getting every criteria, sub criteria weight and other optional drivers' fatigue evaluation, Integral evaluation can be calculated on hierarchy structure model (as shown in figure 1). Firstly normalize weight of n_i criteria under no. i criteria to obtain its relevant weight:

$$W_{ij} = w_{ij} \otimes \left[\sum_{j=1}^{n_i} w_{ij} \right]^{-1} = (o_{ij}, p_{ij}, q_{ij}) \tag{7}$$

Through integrated sub value V_{ijk} and sub criteria weight W_{ij} , evaluation value of criteria B_i of the driver D_k can be got.



$$V_{ik} = \sum_{j=1}^{n_i} (V_{ijk} \otimes W_{ij}) = (R_{ik}, S_{ik}, T_{ik}) \tag{8}$$

In formula (8), $R_{ik} = \sum_{j=1}^{n_i} (e_{ijk} o_{ij})$, $S_{ik} = \sum_{j=1}^{n_i} (f_{ijk} p_{ij})$, $T_{ik} = \sum_{j=1}^{n_i} (g_{ijk} q_{ij})$.

Then, based on every evaluation value V_{ik} and criteria weight W_i , general fuzzy evaluation of selected drivers can be got.

$$V_k = \sum_{i=1}^n (W_i \otimes V_{ik}) = (X_k, Y_k, Z_k) \tag{9}$$

In formula (9), $X_k = \sum_{i=1}^n (O_i \otimes R_{ik})$, $Y_k = \sum_{i=1}^n (P_i \otimes S_{ik})$,

$$Z_k = \sum_{i=1}^n (Q_i \otimes T_{ik}).$$

6 Ranking of Driving Fatigue Evaluation

Through introducing fuzzy max and mini sets and relevant membership function, fuzzy utility value and then rank triangular fuzzy number can be gotten, that is, general fuzzy evaluation. Through 2.3, relative general fuzzy evaluation of n drivers with fuzzy number N_1, N_2, \dots, N_n can be got. Membership function is $\mu_{N_i}(x), x \in R$; relevant fuzzy max set is fuzzy set \tilde{M} , membership function is:

$$\mu_{\tilde{M}}(x) = \begin{cases} \frac{x - x_{\min}}{x_{\max} - x_{\min}} & x_{\min} \leq x \leq x_{\max} \\ 0 & x_{\max} = \sup S \end{cases} \tag{10}$$

In formula (10), $S = \bigcup_{i=1}^n S_i$, $S_i = \{x | \mu_{N_i}(x) > 0\}$, shown in figure 3, right utility relevant to each fuzzy number N_i can be got:

$$U_{Ri} = \sup_x (\mu_{\tilde{M}}(x) \wedge \mu_{N_i}(x)) \tag{11}$$

Similarly, mini set \tilde{G} , relevant membership function $\mu_{\tilde{G}}(x)$, and left utility value U_{Li} can be got, as shown in figure 2, further general utility value U_{Ti} can be got, that is:



$$U_{Ti} = \frac{U_{Ri} - U_{Li} + 1}{2} \tag{12}$$

After obtaining left and right utility values of each selected driving fatigue of selected drivers, each driver’s utility value can be ranked through calculation. If selected driving fatigue evaluation is relevant fuzzy number N_1, N_2 , relevant utility value is $U_{T1} < U_{T2}$, then $N_1 < N_2$; if $U_{T1} = U_{T2}$ (vertex not overlap), fuzzy number close to right is bigger, and it shows general utility value relative with fuzzy number is bigger, hence road surrounding affects driving fatigue bigger.

7 Applying Analyses

Driving fatigue difference evaluation is realized through road test. Test road is typical mountain road section, and collection system of test data is as shown in fig.2. Tested vehicle is designated operating passenger bus of this road section. Selected drivers are five male professional drivers with good health, and normal physiology and psychology.

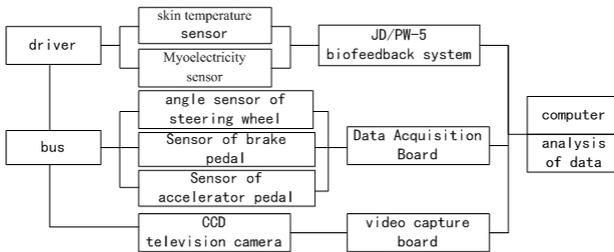


Fig. 2. Structure of test data acquisition system

Decision making group consists of 5 experts, that is P_1, P_2, \dots, P_5 . Utilizing scale method, each criteria of rule layers in pairs are compared, and relationship comparison matrices are built. Utilizing formula (1) to formula (6), experts opinions are integrated, flowingly fuzzy weight of each criteria $B_i (i = 1, 2, \dots, 4)$ is got, each sub criteria fuzzy weight under each criteria is achieved, shown in the table 2. Weight value of each expert is 0.15、0.15、0.20、0.20、0.30.

Table 2. Integrated weights of B layer

| B | P1 | P2 | P3 | P4 | P5 | Fuzzy weight |
|----|--------|--------|--------|--------|--------|------------------------|
| B1 | 0.3333 | 0.1905 | 0.1563 | 0.2812 | 0.2131 | (0.1563,0.2264,0.3333) |
| B2 | 0.3889 | 0.2857 | 0.3125 | 0.3437 | 0.2500 | (0.2500,0.3126,0.3889) |
| B3 | 0.2778 | 0.2857 | 0.3125 | 0.2178 | 0.2295 | (0.2178,0.2622,0.3125) |

Each expert respectively evaluates driving fatigue influence by utilizing fuzzy language value (as shown in the table 1), and obtains evaluated value of each sub criteria under each criteria. Utilizing formula (7) to formula (9), general evaluation of each driver is obtained basis of layer calculation from inside to outside, shown in the table 5. Using formula (12), calculation general utility value of U_{T1}, \dots, U_{T5} can be got, shown in table 3 and fig. 3.

Table 3. General evaluation and utility value of each driving fatigue

| | general evaluation | left utility value | right utility value | general utility value |
|----|------------------------|--------------------|---------------------|-----------------------|
| D1 | (0.2214,0.5486,1.1298) | 0.7234 | 0.6150 | 0.4458 |
| D2 | (0.2076,0.5448,1.1177) | 0.7209 | 0.6117 | 0.4454 |
| D3 | (0.2192,0.5345,1.1215) | 0.7321 | 0.6072 | 0.4375 |
| D4 | (0.2013,0.5348,1.0865) | 0.7357 | 0.5980 | 0.4311 |
| D5 | (0.2258,0.5648,1.1135) | 0.7132 | 0.6175 | 0.4522 |

Selected drivers is ranked by utilizing general utility value of fuzzy number, obviously ranking of selected drivers is D5、D1、D2、D3、D4. The driver D5's driving fatigue is most affected by the operating road section. This evaluation result conforms with feelings of selected drivers. That is, AHP based on group multicriteria decision making integrated with fuzzy ranking can effectively evaluate driving fatigue influence and ranking.

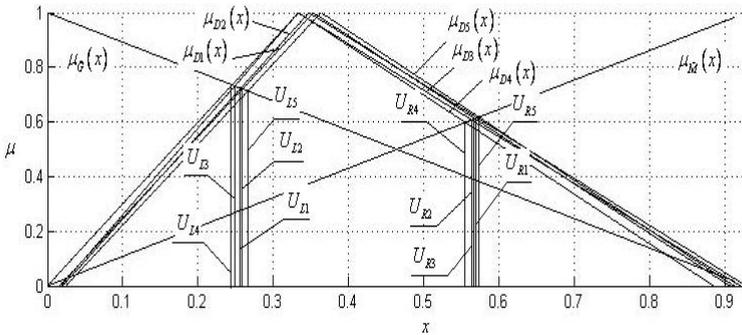


Fig. 3. Utility value of driving fatigue fuzzy number

8 Conclusions

In order to increase accuracy of decision making analytical process, integrating experts' evaluation, decision making result is integrated, decided and judged by using fuzzy utility value to rank fuzzy number relevant with general evaluation of selected drivers. Experts can obtain primary order of each driver, not only combine group opinions, but also expose significant evaluation items. Generally, evaluation result can

reflect driving fatigue influential levels expressed in detailed values by operating road sections, and can be convenient for ranking, therefore effective method for bus companies to select drivers in the special operating road sections is provided.

References

1. Wang, W., Shun, F.: *Driving Behaviour Theory and Its Application in Road Traffic System*. Science Press, Beijing (2001)
2. Zhang, D.: Determination of factors affecting driving fatigue by analytical hierarchy process. *Process in Safety Science and Technology* 3, 46–50 (2002)
3. Analyzes of characteristics and reasons of passenger vehicle super-large traffic accidents (June 2006), <http://www.cps.com.cn/ebook/ly.asp?id=55>
4. Kwong, K.C., Bai, H.: A Fuzzy FAHP Approach to the Determination of Importance Weights of Customer Requirements in Quality Function Deployment. *Journal of Intelligent Manufacturing* 13, 367–377 (2002)
5. Tanadtang, Park, P., Joo, D., Shinya, H.: Incorporating uncertain and incomplete subjective judgments into the evaluation procedure of transportation demand management alternatives. *Transportation* 32, 603–626 (2005)
6. Herrera, F., Martínez, L., Sánchez, P.J.: Managing non-homogeneous information in group decision making. *European Journal of Operational Research* 166, 115–132 (2005)
7. Mao, J.: Advances of Fatigue Detecting Technology for Drivers. *China Safety Science Journal* 15, 108–112 (2005)
8. Stephen, H.: Failrclough Monitoring Driver fatigue via driving performance. *Ergonomics and Safety of Intelligent Driver Interfaces* 130, 221–235 (1997)
9. Heitmann, A., Guttkuhn, R., Aguirre, A.: Technologies for the Monitoring and Prevention of Driver Fatigue. In: *Proceedings of the First International Driving Symposium on Human Factor in Driver Assessment, Training and Vehicle Design* (2001)

Multi-channel Feedback Data-Curtail Identification

Bo Zhao, Qinrang Liu, and Bo Yuan

National Digital Switching System Engineering & Technological R&D Center, China
bozhaow@163.com

Abstract. Single method may hardly satisfy fast and full-scale identification owing to the limitation of capability and speed, especially under the trend of new applications emerging and capacity expanding in Internet. An online identification architecture is proposed, in which identify units are assembled flexibly with various topologies to perform diverse tasks. Furthermore, the front end unit filters traffic according to the feedback of identifying results of back end units. Process pressure of system is reduced for core identifying units so that the system processes can achieve convenient capacity expanding and smooth capability updating. Low cost and high performance online identification is achieved.

Keywords: Application identification, Architecture, Online identification.

1 Introduction

Accurate identification and classification of network traffic is an important task to both bandwidth management and quality of service. Moreover, network engineering problems such as traffic shaping or workload modeling also rely on identification of network traffic.

An instinctive identifications is port number matching, since a port number is usually corresponding to certain application[1][2]. However, the method has become increasingly inaccurate, as users are now able to use random port number to transfer control and data information [3]. And new technologies emerge to solve this problem. For example, the identification based on DPI (Deep Packet Inspection)[4][5], the identification based on DFI (Deep Flow Inspection)[6][7], the identification based on host behavior[8][9] and the identification based on active networking[10]. Usually, one method is usually applicable to one or some certain applications, and has poor accuracy to other applications. New method has to be developed in order to satisfy new application. An integrated and expandable identification system is required to provide high accuracy and high capacity identification.

From the traffic identification perspective a Multi-channel Feedback Data-curtail Identification (MFDI) is pursued, which is easy to expand new identify unit and facilitates the integrated identification of traffic of multiform applications. At the same time, the capacity is enhanced by discarding redundant data while identify processing according the previous identify result feedback.

2 Previous Work

DPI and DFI are two leading technologies used in application identification, and port number matching is included in DPI.

Moore et al. represented a classification scheme utilizing the full packet payload[4]. Packets were grouped into flows and the classification operated on flows instead of packets, so that the information can be processed more efficiently. Moore compared nine distinct identification methods according to complexity and the amount of packets required. Methods with less complexity and less data demanding were used first. And the trade-off was made between system complexity and identification accuracy. However, the selection of methods was not automatically, and may require human intervention. Moreover, seriation is the only one organization form of different methods, which limited expanding ability.

ZHANG Guangxing et al. represented a multi-phases identification methodology (MPI), which measured traffic flow for different applications through both packet level and flow level[11]. A flow is identification with the protocol port-based, signature-based, protocol semantic-based and flow character-based identification methodology. However, similar with literature ^[4], the organization form of MPI limited its capacity.

Karagiannis et al. represented an application identification approach based on observing and identifying patterns of hosts behavior at the transport layer, named BLINC[8]. BLINC analyzes patterns at three levels: i) the number of distinct hosts communicates with single host and communities of hosts those interact with the same set of hosts, ii) the functional role of particular hosts by using the number of source ports and iii) the application level. The limitation of BLINC is that only application types rather than specific applications can be identified. Further more, it is time expensive to analyze the host statistical behavior so that BLINC can be difficult applied in real time measurement.

3 Methodology

3.1 MFDI

MFDI is composed of DV module, AS module and ID module, while ID module is composed of controller, reconfigurable full-interconnected backboard and several identify units. Fig 1 shows the structure. Real line with arrow indicates raw data and transport direction. Broken line with arrow indicates result data and transport direction. Hollow line with arrow indicates control data and transport direction. Traffic identification is processed by DV module and ID module together. ID module includes several identify units those are connected to the full-interconnected backboard. Connecting topology of identify units are adjusted dynamically by the controller.

Process of flows has two steps. The first step is to divide and to filter, and the second step is to identify. The first step has filtered some data so that lighten the burden of the second step. At the same time, results of the second step dynamically influence filter rules of the first step as feedback.

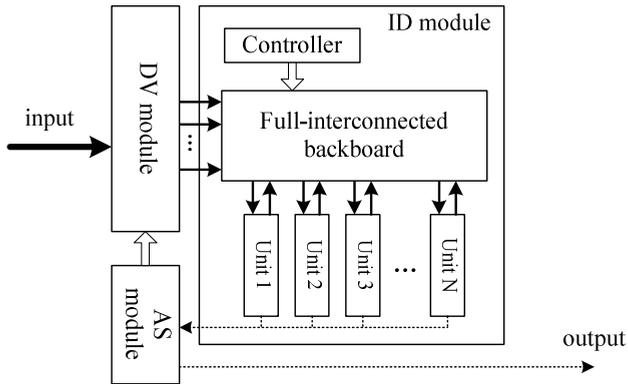


Fig. 1. MFDI architecture

The packets transmitted in the link are captured and mapped into flows by 5-tuple: source IP, destination IP, source port, destination port and protocol. Flows are input into DV module, in which flows are replicated and filtered according to the flow type and filter rules and then output to full-interconnected backboard through several distinct channels. In ID module, data from one channel are processed by one or more identify units. Multi channels process simultaneously helps reduce the time delay and increase the accuracy. Identify units in serial are arranged according to complexity and data requirement. The topology of identify units adjust dynamically with reconfiguration technology to enhance process performance and expanding ability.

Full-interconnected backboard performs flows data exchanging. Flows are directed to certain identify units in serial, parallel or iterative to reduce data amount, expand length of key words and increase speed. Further more, new identify unit may be added into system conveniently and merged with current identify units to achieve function extending.

Results of every identify unit gather to AS module. Final result is output after integration and analysis in AS module. At the same time, new rules are produced as feedback to DV module.

3.2 DV Module

DV module of MFDI performs data division, filtering and degression by mapping packets into flows, replicating to several channels and filtering flows according to certain rules. DV module does not perform identification. Filtering is performed by 5-tuple matching so that can be used on high speed links. Distinct filtering rules are created by AS module for distinct channels.

Abnormal traffic and long term flows occupy quite significant percentage of network traffic. Abnormal traffic includes traffic composed of packets with no payload which usually come from DoS (Denial of Service) attack flow, and simplex flows consist of packets exchanged between a particular port/protocol combination in only one direction between two hosts, and worm and virus attacks flows. Bases 20%-80% law, long term flows consist of 20% in number and 80% in amount of Internet.

The removal of abnormal traffic and long term flows allowed the number of unidentified flows that needed further processing to be significantly reduced.

Simplex topology of identify units in serial has several disadvantages. First, the identification results of DFI are usually probabilities rather than yes or no. And accuracies vary a lot among different types of traffic. Consequently, filtering flows according results of DFI in serial may lead to error cumulation. Furthermore, identification in serial will increase time delay, which is sensitive in online identification. To solve the problem, we separate the identification to several channels. Distinct channels filter by independent rules. DFI identify units identify parallel rather than serial, so that accuracy is increased with no error cumulation.

Filter rules are set according to functions. For instance, set rules to filter traffic that has been identified for channels connected to identify units, set rules to filter abnormal traffic for channels connected to supervise system.

3.3 Identify Units

Every identify unit connect to full-interconnected backboard and construct a certain topology to perform identification. Full-interconnected backboard may change topology dynamically according to the requirement to expand function.

One or more identify units in serial or in parallel corresponding to a data channel. Identify units in serial are arranged according to complexity and data requirement. The topology of identify units adjust dynamically with reconfiguration technology to enhance process performance and expanding ability. System updating may achieved by connecting new identify unit to the full-interconnected backboard and reconfiguring topology in backboard without modifying any other modules and units.

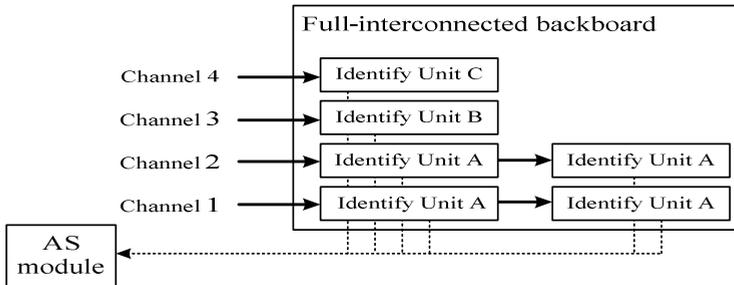


Fig. 2. Dynamical assembling of identify units

In Fig 2, packets pass through Channel 1 with DPI identify units in serial to achieve amount and length increasing of keywords. Packets pass through Channel 1 and 2 with DPI identify units in parallel to achieve speed increasing. Complex algorithms executing by hardware, even software function simulation can be achieved by making packets passing through identify units iteratively.

Results of identify units gather to AS module. Then filter rules are created or updated and downloaded to DV module. At the same time, AS module output final conclusion based on results from every identify units.



4 Results

Four DPI identify engines and two DFI identify engines are connected to full-interconnected backboard as identify units. The topology relationship of them is outlined in Fig 2. Identify unit A is DPI identify engine, identify B and C are DFI identify engines.

Four DPI identify engine are designed based on TCAM components. Set the capacity of TCAM component is M bit, and support identify speed is S bit/s. While the keyword length is configured to L bit, then a single identify engine will support M/L keywords. And two identify engines in serial can support $2M/L$ keywords with length of L bit. Otherwise, the length of keywords may be configured to $2L$ bit, then two identify engines in serial can support M/L keywords with length of $2L$ bit. DPI engines in serial expand the number or the length of keywords. On the other hand, identify engines in parallel can process data separately. So DPI engines in parallel expand the speed of identification.

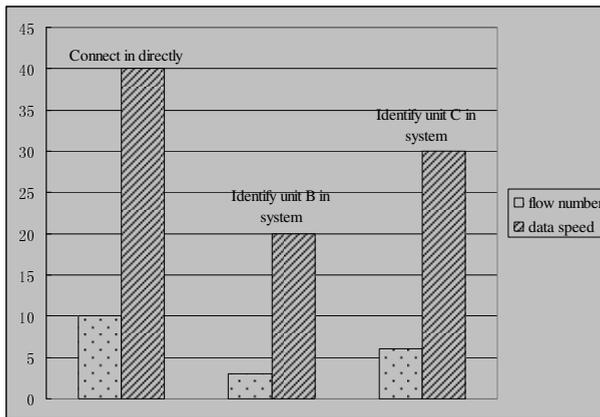


Fig. 3. Load compare of identify units

In this work we designed an experiment system that equips DPI identify engines working at 1Gbps, and with DFI engines identifying P2P and Skype traffic. DFI engines supported maximum number of concurrent flows at 40k, and working at 1Gbps. The filter option is that Channel 1 and 2 filter according to the result of Identify Engine A, Channel 3 filters according to the result of Identify Engine A and B, Channel 4 filters according to the result of Identify Engine A and C. The system was deployed in backbone link at 2.5Gbps. Measurement at steady state shown that the speed of Channel 3 was 0.3Gbps, the maximum number of concurrent flows was 20k. And the speed of Channel 4 was 0.6Gbps, the maximum number of concurrent flows was 30k. Every identify engines were working well. Fig 3 compares the load change of identify unit B and C.

Two metrics are used to quantify performance: maximum number of flows and data speed. Obviously, though the system afford data speed at 2.5Gbps, the load of identify

units in system is less than the identify units solely afford data speed at 1Gbps. MDFI reduces load of identify units evidently and provide better performance.

5 Conclusions

An online application identification architecture MDFI is proposed to the problem of online identification with low cost and high performance. Compared to other traffic identification architectures, MDFI assembles identify units flexibly and efficiently according to the feedback of identify result, and provides the ability of convenient capacity expanding and smooth capability updating, which are very desirable advantages for a potential implementation of an traffic identify system.

References

1. Sen, S., Wang, J.: Analyzing peer-to-peer traffic across large networks. In: Proceedings of the ACM SIGCOMM Internet Measurement Workshop 2002, Marseille, France (November 2002)
2. Fraleigh, C., Moon, S., Lyles, B., et al.: Packet-level traffic measurements from the Sprint IP backbone. *IEEE Network* 17(6), 6–16 (2003)
3. Alshammari, R., Nur Zincir-Heywood, A.: Generalization of signatures for SSH encrypted traffic identification. In: Computational Intelligence in Cyber Security, CICS 2009 (2009)
4. Moore, A.W., Papagiannaki, K.: Toward the Accurate Identification of Network Applications. In: Dovrolis, C. (ed.) PAM 2005. LNCS, vol. 3431, pp. 41–54. Springer, Heidelberg (2005)
5. Karagiannis, T., Broido, A., Brownlee, N., et al.: Is P2P dying or just hiding. In: IEEE GlobeCom (2004)
6. Moore, A.W., Zuev, D.: Internet traffic classification using bayesian analysis techniques. In: ACM SIGMETRICS, Banff, Canada, pp. 50–60 (June 2005)
7. Zuev, D., Moore, A.W.: Traffic Classification Using a Statistical Approach. In: Dovrolis, C. (ed.) PAM 2005. LNCS, vol. 3431, pp. 321–324. Springer, Heidelberg (2005)
8. Karagiannis, T., Papagiannaki, K., Faloutsos, M.: BLINC: Multilevel traffic classification in the dark. In: ACM SIGCOMM (August 2005)
9. Karagiannis, T., Broido, A., Faloutsos, M., et al.: Transport layer identification of P2P traffic. In: Proc. of the 4th ACM SIGCOMM Conf. on Internet Measurement, pp. 121–134. ACM (2004)
10. Dedinski, I., De Meer, H., Han, L., Mathy, L., Pezaros, D.P., Sventek, J.S., Zhan, X.Y.: Cross-Layer Peer-to-Peer Traffic Identification and Optimization Based on Active Networking. In: Hutchison, D., Denazis, S., Lefevre, L., Minden, G.J. (eds.) IWAN 2005. LNCS, vol. 4388, pp. 13–27. Springer, Heidelberg (2009)
11. Zhang, G., Xie, G., Yang, J., et al.: Accurate online traffic classification with multi-phases identification methodology. In: Proc. of the IEEE International Conference on Consumer Communications and Networking, pp. 141–146 (2008)

Research on Parametric Design of Hydraulic Retarder Cascade

Zou Bo¹, Wei Wei², and Yan Qing-dong²

¹ School of Mechanical engineering, Beijing Institute of Technology,
Beijing 100081, China

² National Key Lab of Vehicle Transmission, Beijing Institute of Technology,
Beijing100081, China
ZOU7364@126.com

Abstract. A model of vehicular hydraulic retarder circular circle and blade parametric drawing method was designed. And using the UG/Open API as the secondary development tool, the software of the hydraulic retarder cascade parametric design and 3D modeling was compiled, and then the automatic modeling of series vehicular hydraulic retarder cascade was successfully completed. It was proved that this method could accomplish the work of automatic design and modeling series hydraulic retarder cascade accurately and quickly.

Keywords: hydraulic retarder, blade cascade, parametric design.

1 Introduction

Hydraulic retarder is a kind of secondary braking component which is more reliable and safer than any other assistant braking device. As it is more compact, powerful and less noisy, it is widely used in modern heavy vehicles. The equation of the braking torque when hydraulic retarder works can be expressed as:

$$T_R = \lambda \rho g n^2 D^5 \quad (1)$$

Where the TR is braking torque(N·m), λ is torque coefficient($1/((r/\text{min})^2 \cdot \text{m})$), ρ is density of the transmission oil(kg/m^3) ; g is acceleration of gravity(m/s^2) ; n is speed of rotation(r/min) ; D is diameter of the hydraulic retarder(m).

Torque coefficient λ is the key factor to evaluate the working efficiency and the power density of the hydraulic retarder, and it could be improved by optimizing the structure of the cascade. Traditional design method of hydraulic retarder cascade was retrofit design. In the process of the retrofit design, a series of similar models with different size will be built[1-3], so the design efficiency was decreased by the heavy modeling work.

Parametric design method controls the relative connection and size of the model by a couple of parameters[4-5]. The size and the parameters were corresponding with fixed relationship. When the parameters changed, it will update the model to a new one (see Fig.1).

Series hydraulic component products are similar structures but different size. Compare with the traditional design method, the parametric design method store the whole process of the design. And the process could be changed at any step to get new series similar products. So it is an advanced method for series hydraulic component products design. This paper will using the UG/open API as the secondary development tool, study on parametric design method of the hydraulic retarder cascade.

2 Circular Circle Parametric Design

Hydraulic retarder cascade consists of circular circle and blade. Generally, Design of hydraulic retarder circular circle borrows the idea form the hydraulic coupling[6]. Long circle type is usually used in the hydraulic retarder circular circle design[7]. Fig 2 showed the long circle type of circular circle, the geometry feature of which was analyzed to realize parametric drawing and build the parametric model.

The width B, external diameter D1 and internal diameter D2 is the overall dimension of hydraulic retarder, so these parameters could be specified in the vehicular integrated design. It is also could be specified by the empirical equation. Form the equation (1) we can get the external diameter D1 as:

$$D_1 = \sqrt[5]{\frac{T_R}{\lambda_M \rho g n^2}} \quad (2)$$

Server empirical equations had been concluded form statistic data of hydraulic retarder. The internal diameter D2

$$D_2 = (0.25 \sim 0.6)D_1 \quad (3)$$

The distance Δ between rotor and stator:

$$\Delta = (3 \sim 4)\text{mm} \quad (4)$$

The distance $\overline{O_1O_2}$ between upper and lower center of circle:

$$\overline{O_1O_2} = (0.01 \sim 0.02)D_1 \quad (5)$$

The radius of upper and lower arc is

$$R = \frac{(D_1 - D_2) - \overline{O_1O_2}}{4} \quad (6)$$

To insure the curve of circular circle smooth, the mid arc is tangential with the upper and lower arc. The value of circular circle wide B is determined by the mid arc. As Fig 3 showed, when the center of circle is at the middle of $\overline{O_1O_2}$, the tangent point is at the top with the upper arc and at the bottom with the lower arc of the circular circle.

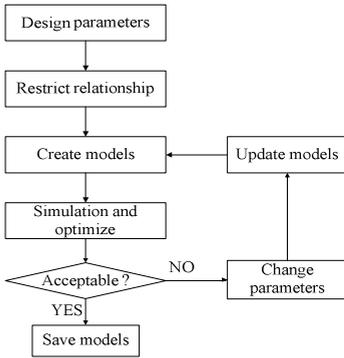


Fig. 1. Process of parametric design

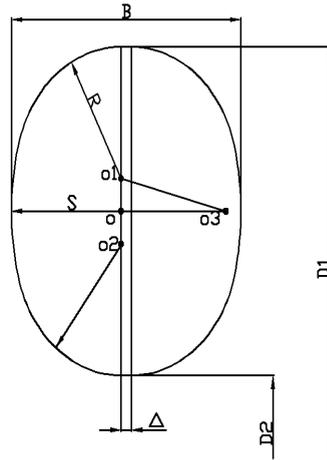


Fig. 2. Hydraulic retarder circular circle

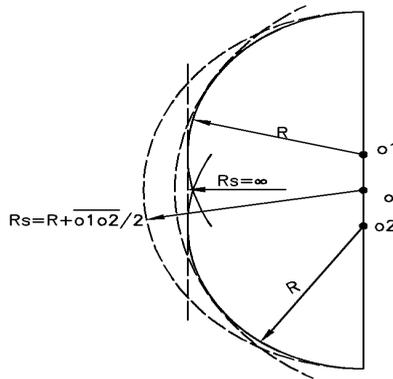


Fig. 3. Geometrical relationship

Then the value of the mid arc radius R_s :

$$R_s = R + \overline{O_1O_2} / 2 \tag{7}$$

Then circular circle wide B get the maximum value:

$$B_{max} = 2R + \overline{O_1O_2} + \Delta \tag{8}$$

When the value of mid arc radius $R_s = \infty$, the mid arc became a straight line. Then the circular circle wide B get the minimum value:

$$B_{min} = 2R + \Delta \tag{9}$$



So the value spans of circular circle wide B:

$$B_{\min} < B < B_{\max} \tag{10}$$

As Fig 2 showed, from Pythagorean proposition we can get equation as:

$$(\overline{O_1O_3})^2 = (\overline{OO_1})^2 + (\overline{OO_3})^2 \tag{11}$$

Then simultaneous equations (4), (10) and (11) we can get the value of mid arc radius Rs:

$$R_s = \frac{4d^2 - 4R^2 + (\overline{O_1O_2})^2}{8(d - R)} \tag{12}$$

The parameter d stands for single impeller wide:

$$d = \frac{B - \Delta}{2} \tag{13}$$

Set the axis x as the revolution axis and rotate the circular circle curve around the axis, then we get the internal surface of the hydraulic retarder cascade.

3 Blade Parametric Design

After the circular circle model had been built, we begin to build the blade model. As Fig 4 showed the blade of the impellers, S represents the stator, R represents the rotor, V represents relative movement direction of rotor, δ is thickness, β is blade angle and Z is blade number.

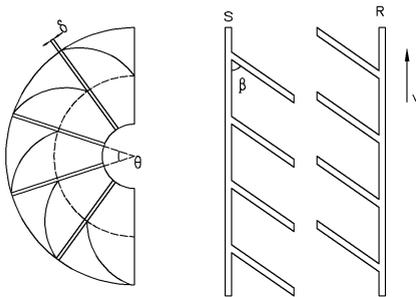


Fig. 4. Blade of impellers

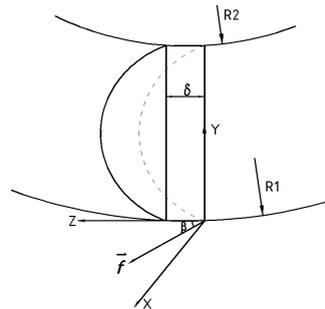


Fig. 5. Blade geometric model

The angular interval θ between blades is determined by the blade number Z:

$$\theta = 360^\circ / Z \tag{14}$$

As Fig 5 showed the blade geometric model ($R1=D1/2$, $R2=D2/2$), set the y-z coordinate planes as the reference plane in coordinate system, the blade parameter δ and Z had been included in the section model.

Set the direction vector \vec{f} to generate the blade with blade angle β .

$$\vec{f} = x\vec{i} + y\vec{j} + z\vec{k} \tag{15}$$

Set the x-z coordinate planes as the reference plane in coordinate system, and then the blade projection in the reference plane is the direction vector \vec{f} . We can get the blade extension direction vector F:

$$\mathbf{F} = [x, y, z]^T = [1, 0, \text{ctg}\beta]^T \tag{16}$$

Extend the section to the internal surface of the hydraulic retarder, and then we get the blade model.

4 Procedure and Example

UG/open API is a professional secondary development tool to realize modeling and design work automatically by program. To store all the design parameters and parameter relationship in the modeling tool, the whole process of modeling was compiled into the program. Fig 6 showed the flow process diagram of program.

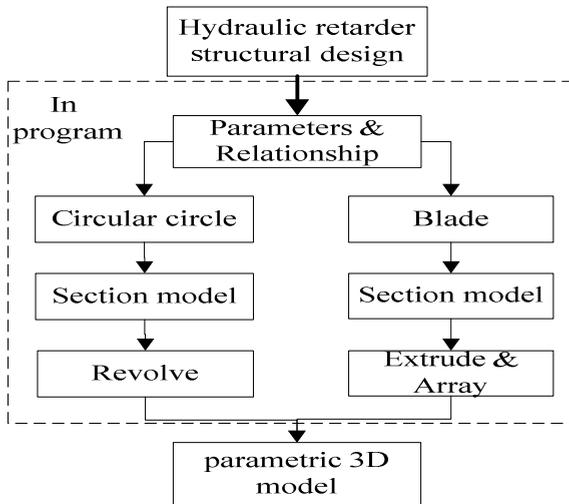
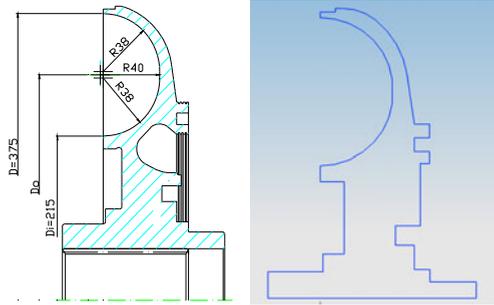


Fig. 6. Flow process diagram of program

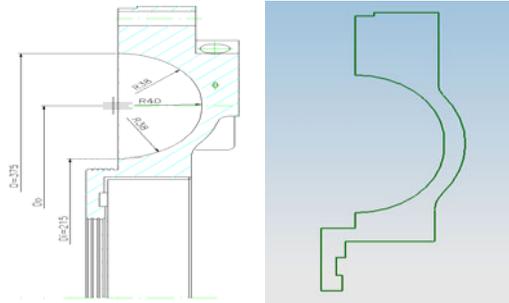
In the program, two three-dimensional arrays had been built respectively to record the point groups of rotor and stator. As the parameters of hydraulic retarder are interrelated, the geometrical relationship between points had also been established in the procedure, so the coordinate of the point could be update automatically by the constraint relationship.



The section file would be generated and initialized firstly, and then the three-dimensional arrays and constraint relationship were endowed in the section drawing program. The section drawing would be built from point to line, from line to face and then form face to body. As Fig 7 showed the CAD and parametric section drawing of rotor and stator, and some detail of profile had been simplified without changing overall structure.



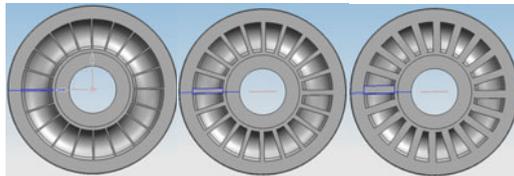
(a) rotor



(b) stator

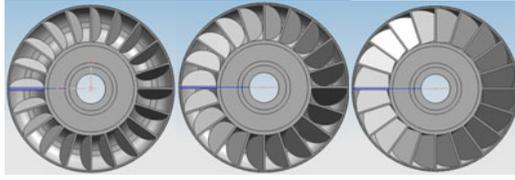
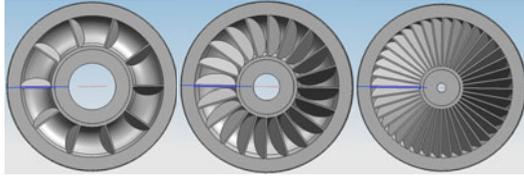
Fig. 7. CAD and parametric section drawing

As Fig 8 showed, by inputting different design parameters, the series hydraulic retarder models could be generated automatically by the parametric design procedure.



(a) stator models with different δ

Fig. 8. Series 3D parametric model

(b) rotor models with different β (c) stator models with different D and n **Fig. 8.** (Continued)

5 Conclusion

In this paper, based on analyzing the design feature of hydraulic retarder cascade, a parametric design method was built. And a software was compiled using the UG/open API as the secondary development tool. The example of series hydraulic retarder cascade automatic modeling showed that the method could complete the modeling work accurately and automatically. It is a foundation of hydraulic retarder automatic design and simulation.

References

1. Guo, L., Du, M.: Configuration optimizing simulation of vehicle hydrodynamic retarder based on CFD technology. *Modern Manufacture Engineering* (1), 104–106, 115 (2009) (in Chinese)
2. Jun, Y., Ren, H., Ming, L.: Numerical simulation of hydraulic retarder with different blade number. *Journal of Jiangsu University (Natural Science Edition)* (1), 27–31 (2009) (in Chinese)
3. Jun, Y., Ren, H.: Performance analysis for hydrodynamic retarder with different vanes. *Transactions of the Chinese Society for Agricultural Machinery* (4), 206–209, 226 (2009) (in Chinese)
4. Li, S.: Research on parametric design and virtual assembly technology of torque converter. School of Mechanical Engineering, Beijing Institute of Technology, Beijing (2006) (in Chinese)
5. Wang, X.: Research on virtual prototype of Hydraulic Retarder. School of Mechanical Engineering, Beijing Institute of Technology, Beijing (2005) (in Chinese)
6. Liu, Y.: Applied manual of hydraulic coupling. Chemical Industry Press, Beijing (2008) (in Chinese)
7. Zhu, J., Wei, C., Zheng, M.: Vehicle hydramatic transmission: upper and lower volume. National Defense Industry Press, Beijing (1982) (in Chinese)

A Hardware Design of CCM3118 Evaluation System

Xu Feng-liang¹ and Shi Binbin²

¹ Information Centre, Suzhou Health College, Suzhou, China

² Dept. of Computer Engineering, Suzhou Vocational University, Suzhou, China

xusongxiaorong@sina.com

Abstract. The basic system of CCM3118 hardware evaluation system and hardware module design approach of extended FLASH, extended SRAM, Ethernet, USB, LCD are given. In order to improve the stability of chip selection under high-speed bus, the usage of CPLD is discussed in detail. Basic method of hardware testing and design framework of hardware driver is briefly described.

Keywords: C*CORE, CCM3118, Hardware Design.

1 Introduction

In 1998, MOTOROLA, INC designed and developed 32 bits central processor unit M*Core. After 2000, several general-purpose microcontrollers (MCU) based on the M*CORE M210 were produced. M*CORE technology transferred to China in September 2000. C*Core Technology (Suzhou) Co., Ltd absorbed this technology, and produced microcontroller (MCU) CCM3118 based on the C*CORE of our own intellectual-property rights in 2004[1]. Features of the CCM3118 include low power, high performance. The application program of C*CORE has better operating efficiency, less code quantity and less memory occupation than other 32bit MCU core's. The overall performance of C*CORE is better than ARM7's. It is applied in information security, consumer electronics, OA system, network communication, industrial control, car electronics and other fields.

Generally, electronic manufacturers will give hardware evaluation system which suit for new produced MCU, in order to accelerate the promotion of new product. At present few companies can provide CCM3118 embedded hardware evaluation system. This CCM3118 Reference design kits was developed by Soochow University and C*Core Technology (Suzhou) Co.,Ltd[2]. It includes hardware evaluation system, integrated development environment, debug and program tools. It can help teaching of embedded system and project development.

The application-specific CCM3118 Evaluation System follows the design rules of embedded hardware evaluation system. It extends common modules of embedded system development. Design methods of hardware evaluation system and minimum hardware circuit are introduced in detail in this paper. There are only 64Kbyte SRAM and on-chip FLASH on CCM3118, so it is difficult to run bigger program and operating system on it. Extending FLASH and SRAM by external bus interface module of CCM3118 to fix this problem. Complex Programmable Logic Device (CPLD) is used to unscramble chip selection and address signals. It can solve the

problem of read-write error caused by signal transferring latency. Hardware evaluation system extends Ethernet, USB, LCD, etc, so the system would have more application value. At last, all formal hardware driver programs are given.

2 Hardware Evaluation System Architecture of CCM3118

CCM3118 is a kind of 32-bit reduced instruction set computer (RISC) architecture MCU, using 180nm HCMOS art manufacture .The operating frequency is up to a maximum of 60 MHz over a temperature range of -40°C to 85°C . Available packages are 208 pin low-profile quad flat pack (LQFP). For applications requiring the full external memory interface support or a large number of general-purpose inputs/outputs (GPIO).CCM3118 hardware evaluation system includes minimum hardware circuit, extended FLASH, extended SRAM, CPLD unscrambling circuit, Ethernet, USB, LCD. Hardware evaluation system structure is shown in figure 1.

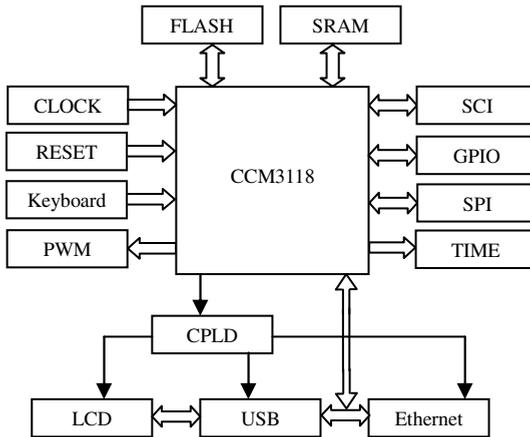


Fig. 1. Hardware Evaluation System Structure

3 Detailed Design of CCM3118 Hardware Evaluation System

3.1 Minimum Hardware Circuit

In hardware evaluation system, Ethernet chip RTL8019AS and LCD operate at voltages 5 volts, USB chip SL811HST, CPLD, extended FLASH, extended SRAM, GPIO of CCM3118 operates at voltage 3.3 volts, the core of CCM3118 operates at voltage 1.8 volts, therefore using AMS1117-3.3 chip convert input voltage DC 5 volts to 3.3 volts and using AMS1117-1.8 convert 3.3 volts to 1.8 volts.

The clock module of CCM3118 can be operated in 1:1 normal PLL mode, normal PLL mode or external clock mode. In normal PLL mode, the PLL is fully programmable. It can synthesize system frequency F_{sys} at PLL reference frequency F_{ref} and has a post divider capable of reducing this synthesized frequency without

disturbing the PLL. The PLL reference can be either a crystal oscillator or an external clock. The frequency of CCM3118 is:

$$F_{sys} = \frac{F_{ref} \times (M/R)}{(1+K) \times 2}$$

M is multiplication factor divider field. R is reduced frequency divider field. K is post-divisor k bit. These values can be set clock module registers. The EXTAL pin of CCM3118 connect to 10MHz active crystal and the values of M, R, K are set to 24, 2, 0, then the system clock can works at 60MHz.

Reset circuit is composed of low-power supervisors chip IMP811S, which is more reliable than resistor-condenser reset circuit. A reset signal is issued if the power supply voltage drops below 2.93 volts and is asserted for at least 140ms after the supply has risen above the reset threshold. It can avoid system working in an unpredictable state.

Extended FLASH and SRAM Circuit.

Two FLASH chip Am29LV800B produced by AMD are connected in parallel in hardware system, which can provide 32bit data width, 2Mbyte memory capacity. It stores program codes, constant table, user's data would not be lost when power-fail. Extended FLASH circuit is shown in figure 2.

SRAM is used as program running space, stack and data block. Data not to be lost when power-fail is not the characteristic of SRAM, but its access speed is faster than FLASH's. Two SRAM chip IS61LV25616AL produced by ISSI are connected in parallel in hardware system, which can provide 32bit data width, 1Mbyte memory capacity. It can meet the requirements of complicated embedded operating system like μC/OS-II or μCLinux. Extended SRAM circuit is shown in figure 3.

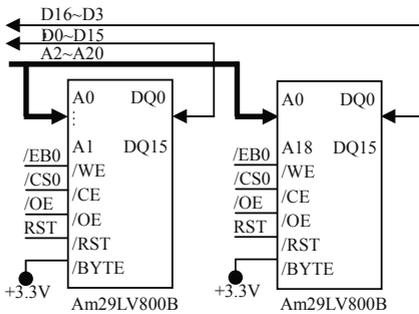


Fig. 2. Extended FLASH circuit

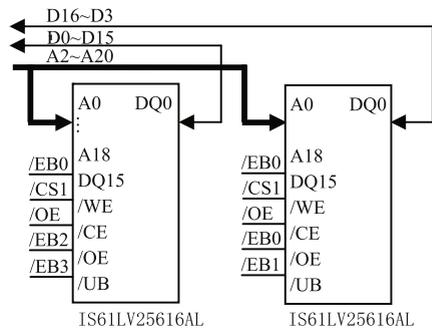


Fig. 3. Extended SRAM circuit

3.2 CPLD Decoding Circuit

CCM3118 chip select module provides four programmable asynchronous active-low chip enable signals CS0-CS3 for external memory and peripheral devices. In hardware evaluation system CS0 and CS1 enable extended FLASH and extended SRAM. CS2, CS3 and address signals are used to address decode, therefore the system can extend more peripheral device modules.



Address decoder using 74HC00 logic NAND gate was tried, but the delay of address decoder caused read-write error of peripheral devices. The switching time of each 74HC00 is more than 20ns, but the clock cycle of CCM3118 is 17 ns. When the write period is over, the write enable signal is not over because of time delay; at last the unpredictable data from data bus is written into the peripheral device. It is same as read period.

To solve this problem, address decoder using XC9572XL produced by Xilinx, which is a kind of high-performance CPLD. XC9572XL has 72 macrocells with 1,600 usable gates. Pin-to-pin logic delays are 5 ns [3]. Because of in-system programmable, users can modify pins' logic functions as required at any time. Using CPLD can shorten wiring distance, reduce the PCB area, and resist the Electro Magnetic Interference (EMI). CPLD connecting to peripheral devices is shown in figure 4.

CPLD decodes address signal A0-A22 and chip enable signal CS2, outputs enable signal for peripheral devices. The logical structure of XC9572XL can be described by Verilog language, and developed by Xilinx ISE 6.2 tools. For example, active-low signal AEN of RTL8019AS is described as follows.

```
assign R8019_AEN =!( (!CS2) && (!A[22]) && (!A[21]) && (!A[20]) &&
(!A[19]) && (!A[18]) && (!A[17]) && (!A[16]) && (!A[7]) && (A[6]) );
```

Address range of chip select signal CS2 is encoding from 0x8000_0000 to 0x807f_ffff. When accessing any address from 0x8100XX40 to 0x8100XX7f, the CPLD outputs low signal and active RTL8019AS. Other peripheral devices are driven by this method. CPLD also processes control signal. For instance, CPLD input low level reset signal and output high level reset signal to reset RTL8019AS chip. Connection of CPLD and main peripheral devices is shown in figure 3.

3.3 Extended Ethernet Circuit

Currently embedded Ethernet technology is more and more widely used. The application value of embedded system would reduce without Ethernet module. Ethernet chip performs physical layer protocol. The RTL8019AS is a highly integrated Ethernet Controller which offers a simple solution to implement a Plug and Play NE2000 compatible adapter with full-duplex and power down features.

In hardware evaluation system the JP pin of RTL8019AS inputs low, RTL8019AS is in jumper mode. CCM3118 is compatible of 5 volts high level signals, so the RTL8019AS is directly connected with CCM3118 through pins. The IOCS16B pin inputs low, so RTL8019AS uses in 8-bit slot not in 16-bit slot. PCB could layout easily, and not hurting performance very much. It is remarkable that the chip select control register 2 must be configured as a 16-bit port size, and only low 8-bit of the data bus is valid for RTL8019AS. Other peripheral devices are all use this 8-bit bandwidth data bus.

3.4 Extended USB Circuit

USB (Universal Serial Bus) interface shows good performance in many aspects, such as the simple architecture, plug-and-play support, good expandability, and low power. The application of USB has been very broad in many fields. The SL811HS produced by

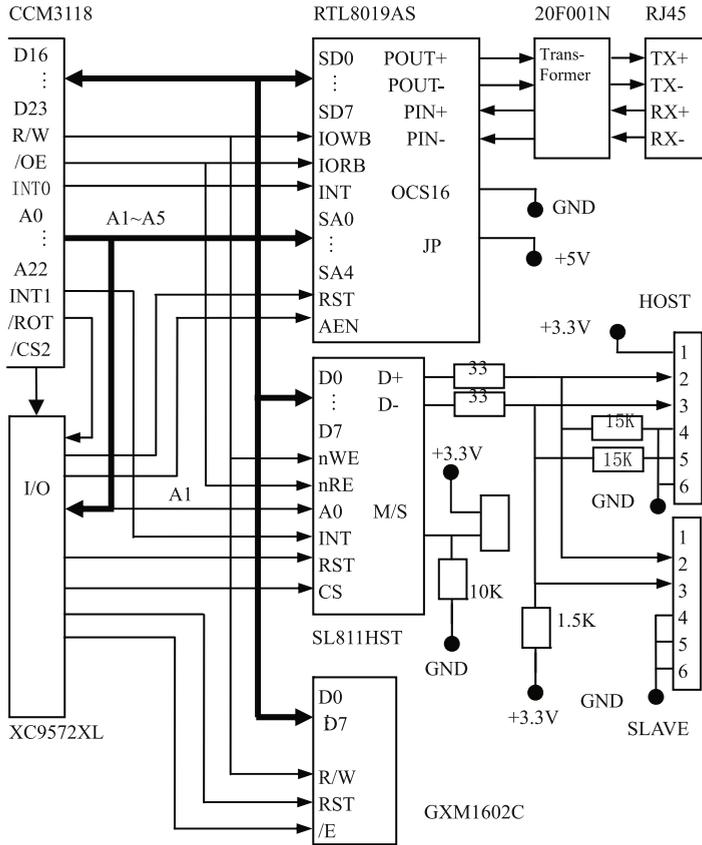


Fig. 4. Connection of CPLD and main peripheral devices

CYPRESS is an Embedded USB Host/Slave Controller capable of communicate with either full-speed or low-speed USB peripherals. The SL811HS USB Host Controller conforms to USB Specification 1.1. The SL811HS data port and microprocessor interface provide an 8-bit data path I/O or DMA bidirectional, with interrupt support to allow easy interface to CCM3118. It contains a 256-byte RAM data buffer which is used for control registers and data buffer.

The I/O pins D0-D7 of SL811HS are used with multiplexed address/data bus interfaces, they are connecting to CCM3118 pins D16-D23. The SL811HS contains 256 bytes of internal buffer memory. The first 16 bytes of memory represent control and status registers for programmed I/O operations. The remaining memory locations are used for data buffering. Access to memory and control register space is a simple two step process, requiring an address Write with A0 set low followed by a register/memory Read or Write cycle with pin A0 set high. SL811HS pin A0 connects to CCM3118 pin A1. The pin M/S input low, SL811HS works in host mode and host interface type A useful. The pin M/S input high, SL811HS works in slave mode and slave interface type A useful.

3.5 Extended LCD Circuit

Human machine interface is used most widely in embedded system. It chooses 16 x 2 character LCD GXM1602C According to the price element and universality. CPLD decodes address signal and chip enable signal to drive LCD. GXM1602C pin RS is connected to CCM3118 pin A1. The pin RS input low, to access instruction registers. The pin RS input high, to access data registers. The R/W input signal indicates the direction of the data transfer on the bus. High level indicates a read from CCM3118 and low level indicates a write to LCD. The CPLD decodes address signal and chip enable signal to drive other peripheral devices like GPIF and keyboard, which circuits are not given in this paper.

4 Drivers of Hardware Modules

Drivers of hardware modules are designed based on modular ideology. The drivers are used in hardware test and practical application. Users do not need or only need to revise a small amount of code, and the drivers can be used in real project. After Testing, the program proved to be right and system proved to be stable. As the space is limited, this paper only lists directory of projects which is shown below in table 1.

Table 1. Directory of Projects

| Project | Function |
|--------------|---|
| GPIO.prj | setup GPIO and control led |
| Sci.prj | Ethernet communication |
| Ethernet.prj | serial communication |
| USBhost.prj | USB works in host mode, Communicate with flash disk |
| USBslave.prj | USB works in slave mode, Communicate with computer |
| LCD.prj | receive characters from SCI and display on the LCD |
| PWM.prj | adjust duty cycle to change brightness of light |
| Time.prj | show time by 8-SEG LED |
| SPI.prj | Read and write EEPROM via SPI interface |
| Keyboard.prj | Receive value form keyboard |

5 Summary

CCM3118 hardware evaluation system has been developed successfully. It has been used in teaching experiments in some universities now, and received good responses. This project passed China Jiangsu province science and technology achievement appraisal. The industrial class MCU based on the C*CORE is applied in handheld devices, consumer electronics, industrial control and other fields. It has good prospects. The design methods of hardware evaluation system, the encountered problems and the

solutions are introduced in detail in this paper. Design framework of hardware driver is also briefly described. It can promote application of C*CORE.

References

1. C*COR Technology Co., Ltd., CCM3118DQ Advance Information Rev. 1.0 (2003), <http://www.china-core.com>
2. Lu, X.-F., Wang, Y.-H., Chen, S.: The Design for CCM3118 Evaluation Development Board and IDE. *Microelectronics & Computer* 22, 82–85 (2005)
3. Xilinx: XC9572XL Data Sheet (2004), <http://www.xilinx.com>

RNP Flight Procedure Analysis and Simulation

Lan Ma and Ying Lin

School of Air Traffic Management, Civil Aviation University of China,
300300 Tianjin City, China
lin_li123@sina.com

Abstract. Required Navigation Performance (RNP) is an important part of the future air traffic industry. It releases the congestion of the air traffic, reduces the distance of airlines, and increases flight effect. Now it has been applied in many airports and airlines. However, the high accuracy type of RNP has several special attributes. Some researches must be done before application. This paper purposes the framework of RNP procedure design through analyzing some related criteria, and then simulates an integrity RNP procedure based on Google Earth, and also considers some error performance.

Keywords: Required navigation performance, Flight procedures, Navigation system error, Google Earth.

1 Introduction

With the development of civil aviation, the air traffic flow is increasing dramatically and the congestion in the en-route and terminal area is accrued frequently. The flight safety is threatened and the operation efficiency is reduced directly. Thereby the International Civil Aviation Organization (ICAO) proposes the conception of “Required Navigation Performance (RNP)”, which applies to navigation performance within airspace and therefore affects both the airspace and the aircraft [1].

In [1] RNP is defined as the navigation performance within certain airspace which the navigation performance accuracy value of the flight is expected to be achieved at least 95 percent of the time by the population of operating aircrafts. It is significant of RNP application in the civil aviation. This paper will show an integrity RNP flight procedure based on Google Earth and analyze some error factors which exist in the real environment.

2 RNP Requirements

Detailed RNP requirements are primarily captured in [2, 3, 4], and [5]. And they can be mainly classified into two categories—flight legs (Track-to-Fix (TF leg) or Radius-to-Fix (RF leg)) and containment tunnels. Requirements related to flight legs, such as maximum bank angle, distance of turn anticipation (DTA), and minimum leg length are used to ensure the flight legs are flyable by the host aircraft. Requirement related to containment tunnels, such as the width ($2 \times \text{RNP}$) of the lateral containment

and required obstacle clearance (ROC) for the vertical containment are used to ensure the safety of the aircraft and enable efficient air traffic management. Fig. 1 shows the profile view of both lateral and vertical containments.

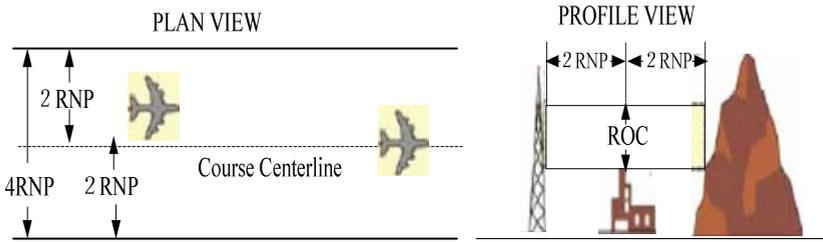


Fig. 1. Profile View of Containment Tunnels

3 RNP Error Analysis

The total system error (TSE) of RNP includes along track error (ATT) and lateral track error (XTT). ATT is made up of navigation system error, RNAV calculation error and display error, and XTT is made up of navigation system error, RNAV calculation error, display error and flight technical error. As Fig. 2 shows.

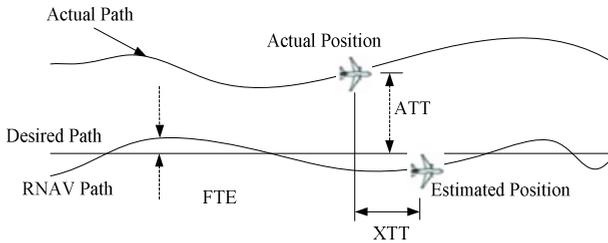


Fig. 2. ATT and XTT

As the detailed description above, the more accurate of the RNP type is, the more requirements for the navigation system. Under the traditional navigation system, such as VOR/DME, DME/DME, the TSE distributes between 1.40nm and 8.6nm, which can realize RNP4, RNP5. For the special type of RNP, such as 0.1-0.3, that is, authorized requirement (AR) has a higher requirement for the navigation performance. As Table 1 show.

Table 1. Error of the Navigation System

| Name | DEM/VOR | DME/DME | GNSS | Special System | Navigation |
|-------|---------|---------|------------|----------------|------------|
| Error | RNP5 | RNP4 | RNP0.1-0.3 | RNP0.1 | |

4 The Approach Procedure of RNP Operation

As the detailed description above, the conception, requirements and some related criterions have been introduced. In this section, the procedure of RNP operation will be shown.

The conduct of RNP procedures requires that the aircraft operator examine its crew information, flight procedures and training record to ensure that they are sufficient to enable operator qualification and operational approval. In this paper, an integrity RNP design frame is proposed, shown in Fig.3.

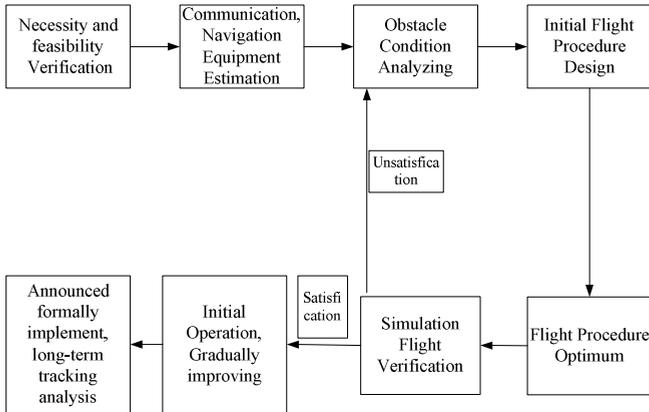


Fig. 3. Procedure of RNP Operation

In the process of design RNP procedure, there are several jobs must be done. Firstly, the necessity and feasibility of the specific airport must be analyzed, including current airport flight flow, the safety level, the future development trend and some related factors. Then make sure whether the navigation facility in this area can meet the accuracy type requirement that the RNP will be operated and ensure the correctness, integrity, continuity of the navigation facility. And in this stage, the condition of obstacle must be considered, especially in the RNP procedure. In the process of analysis, the possible flight path and some unconsidered factors must be considered. At the same time it is necessary for soliciting the opinions, optimizing the procedure and providing several cases to be chosen. Secondly simulate the procedure and verify whether it meet the requirement of the design. Finally, after being announced officially, it is still need to carry on the long-term tracing analysis. Some problems could not immediately appear and need a lot of statistical analysis to obtain the persuasive conclusion.

5 The Design of RNP Approach Procedure

According to the ICAO manual [6], the traditional flight approach procedure contains approach segment, initial approach segment, intermediate approach segment, final

approach segment and missed approach segment. Through the RNP procedure also has the same segment, some special aspects should be considered.

As the detailed description above, there are two main flight legs in RNP procedure. Also in the flight leg construction, the containment should be considered. And the obstacle clearance is an important factor that ensures the aircraft security. Some evaluation about the obstacle area must be made. We also should consider descend gradient in the design process. Descend gradient is calculated by the nominal anchor point locations. The instruments approach procedure should use the standard descend gradient.

6 Simulation and Flight Test Results

Through the analysis about RNP procedure, the procedure and some other conditions which should be paid attention to also have been introduced. This paper conducts an integrity RNP procedure with RNP type is 0.1 based on Google Earth secondary development platform.

Firstly, the Simulation Procedure is introduced. The system Framework is shown as Fig. 4. It has two terminals-Server and Client. The server mainly completes the track display, the position updates, coordinate transformation and flight posture extraction. The client displays the real time position that receives from the server and anticipation position of the plane that reads from the database and displays the error information. At the same time, the client stores the track data in order to analysis after the simulation.

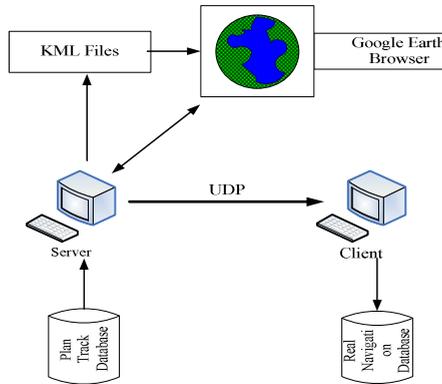


Fig. 4. System Framework

Secondly, the simulation results are shown in Fig. 5.

This paper firstly constructs an integrity flight path based on Google Earth. And then the 3D plane model is loaded on the Google Earth, by reading the position information from the navigation database, the position is updated, and realizes the flight dynamically.

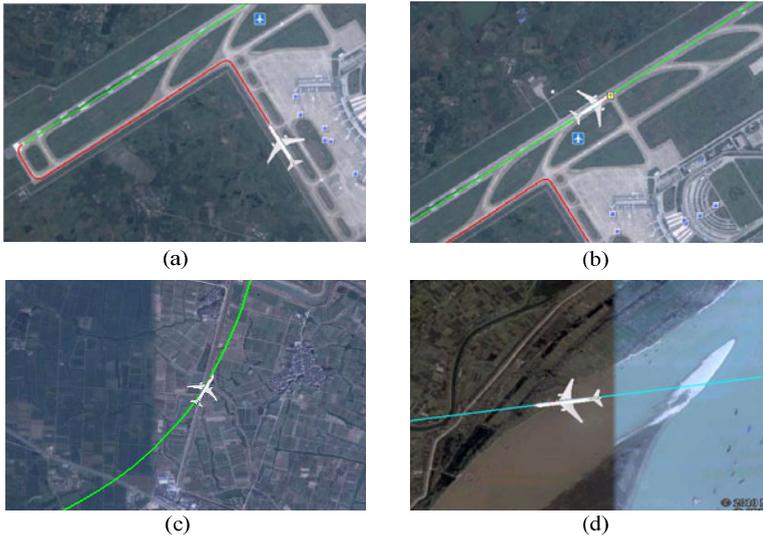


Fig. 5. Flight Simulation

The client displays the real-time and anticipation position information, and calculates the lateral and vertical error based on the flight data. As Fig. 6 shows.



Fig. 6. Client display

The client also has the function of alarm when the position of the plane beyond the RNP containment.

7 Conclusions

According to RNP flight procedure related requirements and standard, the author simulates an integrity RNP flight procedure based on Google Earth. It is significant for our practice application. However, due to lack of navigation data, the simulation is only

under the extremely ideal condition. Considering aircraft in real flight process, there may be more problems, such as pitch climb, turning and some other factors. Thus there are still some problems need to be solved in the future.

References

1. ICAO. Manual on Required Navigation performance. ICAO, Montreal (1999)
2. Do-236B, Minimum Aviation System Performance Standards: Required Navigation Performance for Area Navigation, RTCA, Washington D.C (2003)
3. DO-283A, Minimum Operational Performance Standards for Required Navigation Performance for Area Navigation, RTCA, Washington D.C (2003)
4. FAA ORDERS 8260.51 Required Navigation Performance (RNP). Instrument Approach Procedure Construction, FAA (2002)
5. ICAO. Required Navigation Performance Authorization Required (RNP AR) Procedure Design Manual. ICAO (2008)
6. ICAO. Doc 8168-OPS/611-Volume 1 Procedure For Air Navigation Services—Aircraft Operation (Volume 1) Flight Procedures, 4th edn., Canada (1993)

Infrastructure-Aid Information Forwarding Based on Network Coding in Vehicular Ad Hoc Networks

Gang Wang, Yi Xu, and Xia Dai

School of Electronics and Information Engineering, Beihang University, Beijing, China
wangch318@yahoo.cn

Abstract. In this paper, we present an analytical study for data dissemination in vehicular ad hoc networks (VANET) in the urban areas. Without any routing protocol, random linear network coding (RLNC) is an efficient method in practical scenarios contrasted with flooding. In such scenarios, some vehicles cannot communicate with others due to their mobility and non-connection with the neighbors. To tackle this problem, we propose a mechanism using the roadside infrastructure to assist transmission, which allows the majority of vehicles in a connected domain. Through this mechanism, the packets delivery ratio is improved significantly and other properties are minimally affected. Furthermore, the feasibility of this approach is verified through the simulations.

Keywords: Infrastructure-aid, information forwarding, network coding, vehicular ad hoc networks, urban area.

1 Introduction

Intelligent Transportation System (ITS) is developed all over the world to settle the urban traffic management problems such as traffic jam, emergency treatment and vehicle route planning due to a surge of the vehicles number on the road. ITS is mainly applied in the vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) to provide the traffic information about road obstruction, emergency incident, location and entertainment. Owing to the vehicle mobility, topographic conditions and traffic jam, more efficient wireless technique which can solve those problems to improve communication performance attracts significant interest from research communities[1-2].

In VANET, the vehicle speed is different in the urban scene and in the highway. In this paper, we focus on the urban scene. To ensure the efficient operations of the urban traffic, save driving time and reduce driving risk, the vehicular networks need to enhance the broadcast capacity, assure the transmission quality and share the traffic information immediately and efficiently.

The remainder of paper is organized as follows. Section II provides a review of related works. In Section III, it proposes an infrastructure-aid forwarding mechanism based on network coding and provides the specific steps. In Section IV, we present the simulation results and finally Section V concludes this paper.

2 Related Works

Flooding is the simplest data dissemination protocol in mobile ad hoc networks [3]. The information can almost be transmitted to each node in the network. However, there still exist problems as follows: Firstly, due to the omnidirection of wireless transmission, most parts of the area are within multiple nodes transmission range, which result in broadcast storm for too much redundant information received. Secondly, the use of that mode may suffer from a high number of collisions and channel competition because the transmission time slots between adjacent nodes are highly correlated. Thirdly, abundant hidden and exposed nodes affect the reliability of broadcast information distribution.

VANET has two primary applications: (1) Safe driving application. This application can significantly reduce the number of traffic accidents. According to the research, if the driver has half a second before the collision to be alert, then 60% of the traffic accidents will not happen; (2) User application. User application can give passengers advertising, entertainment and other information in the journey.

For the first application, the alarms spread to each other inter-vehicles. Researchers have proposed some routing protocols such as the improved on-demand routing protocols and the improved geographic-based routing protocols. However, some protocols rely on the priori knowledge such as GPS positioning. In the actual VANET, the popularization of GPS is a process. So the robustness of the protocol should be considered. TIBCRPH [4] is a traffic infrastructure based cluster routing protocol. It also uses GPS to perceive the position. The maintenance of cluster-head will increase the protocol overhead. And some protocols ignore a lot of practical factors. For example, in most experiments all the vehicles are supposed to move at the same and constant speed. What is more, in different parts of the area, the density of vehicles are uneven, which factor is often ignored.

For the second application, both vehicles and infrastructures can push the information to the other vehicles. Many routing protocols model P2P applications as BitTorrent to download multimedia files cooperatively [5]. In this application, most of the protocols only concern the case that only one large data source such as a multimedia file. CodeOn [6] uses Access Point to broadcast two kinds of messages unless the vehicles are outside the AP coverage.

This paper focuses on the safe driving application. The main characteristic of such applications is that some of the data sources will send traffic information simultaneously and the size of the message is small. To avoid hidden terminal problem and increase the reliability in broadcast mode, several schemes such as transmitting RTB/CTB and acknowledgement messages [7] have been put forward. However, since the small safety message size is comparable to the RTB/CTB and acknowledgement message size, it would not be a reasonable choice in this application. Our ultimate goal is to propose a mechanism to improve network performance in the urban scene for safe driving application.

3 Network Coding and Infrastructure-Aid Mechanism

We consider the random network coding schemes of [8] for all-to-all data transmission. The coding arithmetic exploited here is random linear network coding. Each source node that wants to transmit information to all other nodes (broadcasting packets) is independent of any routing protocols and has no data retransmission mechanisms. Nodes exchange coded frames instead of data packets. We define a coded frame c to be a linear combination of packet p_k .

$$c = \sum_{k=1}^n e_k p_k \quad (1)$$

Where e denotes the coefficient of encoding packets in the finite Galois Field $GF(q)$ with $q = 28$. In the header of a coded frame, the encoding vector $e = [e_1 \dots e_n]$ is stored for the purpose of later decoding. When a node receives an innovative encoded packet, it will recombine the packets of its buffer and update the encoding vector $e' = [e'_1 \dots e'_n]$ in the packet header:

$$c' = \sum_{k=1}^n e'_k c_k \quad (2)$$

Then continue to transmit until the coefficient vectors matrix E has a full rank in the buffers of all its neighbor nodes within its receiving range. Then the node starts to decode the packets after the packets coefficient matrix E reaches full rank. Then we can get the original packets through E and C .

$$P = E^{-1}C \quad (3)$$

In the urban area, some vehicles may have no neighbor nodes within the communication range in a certain time. Under that circumstance, we can utilize the roadside infrastructure such as road lamps, signal lamps to assist packets transmission. V2I communications and I2V can be applied to achieve V2V communication among unconnected vehicles, which makes more vehicles be aware of their surrounding road conditions, traffic accidents and other related information in a certain range. Infrastructure is characterized by high energy and wide transmission range. In this paper, infrastructure is employed only to store and forward data without recoding. Similarly, each node including infrastructure and vehicle which just broadcasts or encodes packets does not rely on any routing protocols and has no data retransmission mechanisms. To avoid collisions, each node sends packets with a random delay from 0 to 0.1s.

Our mechanism is presented as the following steps:

1) We set five wireless base-stations uniform distribution. The base-station can receive signals from at least one other base-station. Its sending range is 330m that the vehicles can receive the signals sending by base-stations from 330m away. And the receiving range is 250m the same as the vehicles.

2) Each vehicle node sends to the network a single original packet and wants to collect all the other transmitted packets.

- 3) The base-station forwards the packet it received to assist the node which has no neighbor node in its transmitting range.
- 4) Each original packet is sent only once.
- 5) When a node receives a linear independent packet, it reassembles the packets in the buffer and forwards the combination until the coefficient matrix of each neighbor has a full rank or the timer expires.

4 Simulations

In this section, we discuss the relevant results via ns2 simulations. We adopt the basic MAC protocol IEEE 802.11b that, in the broadcast mode, does not use any acknowledgment mechanism. In case of collision, no retransmission occurs and the packet is lost.

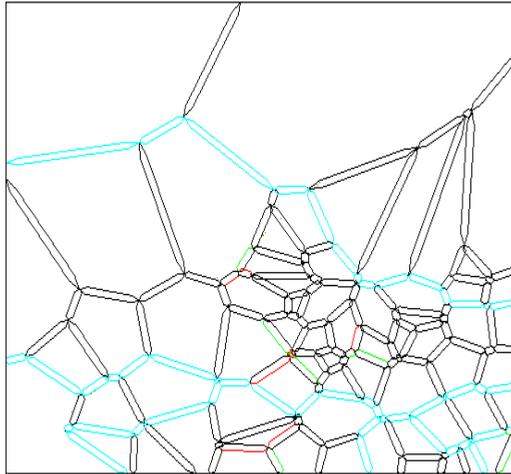


Fig. 1. A street map in the urban scene

Network Topologies: There are some widely used mobility models such as Random Waypoint (RWP) to restrict the areas where nodes can appear. However, those models cannot reflect the real situation of the roads in the city scene. Here, we use the simulator VanetMobiSim to simulate realistic urban scene. In the simulation area: 30% are road-intensive areas; 30% are the residential areas; and the last 40% have sparse vehicles such as inside the hospital, park and playground. In order to be more close to the city scene, it selects at most six of the crossroads to set traffic lights. Vehicles must appear in the road following the traffic laws. Their speed is time varied between 20km/h to 50km/h and the vehicle will stop when the traffic light is red. One of the simulation scenarios is shown in Figure 1, a street map of 1,000m×1,000m in the urban scene.

We test the algorithms varying the number of vehicles, n , from 20 to 60. Their speed is time varied between 20km/h to 50km/h.

NC_I (Network Coding Infrastructure aid mechanism) and flood_I (Flooding Infrastructure aid mechanism) are the infrastructure-aid mechanisms respectively deriving from NC and flood. It sets five infrastructures in each scenario in the infrastructure-aid mechanisms.

Performance Metrics: Packet delivery ratio, average delay, protocol overhead and throughput.

Packets Delivery Ratio

Packets delivery ratio is defined as the ratio between the number of received (or successfully decoded) packets and the number of transmitting packets.

$$PDR = \frac{R_p}{S_p} \quad \text{or} \quad (PDR = \frac{D_{C_p}}{S_p}) \quad (4)$$

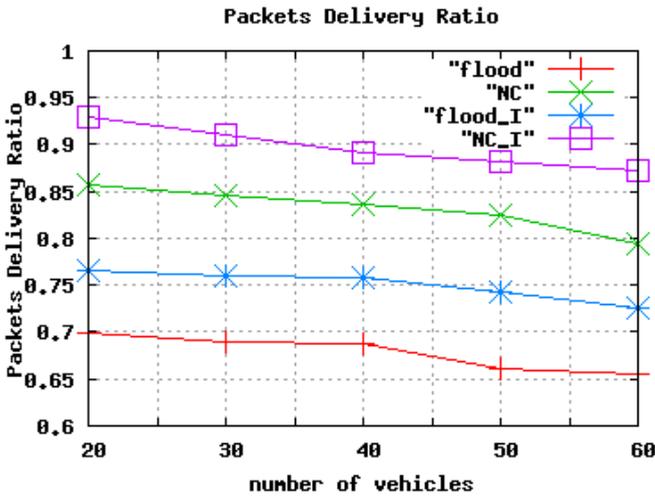


Fig. 2. Packets delivery ratio for the various number of vehicles

SP and RP denote the number of all the transmitted and received packets. Fig.2 is compared to the packets delivery ratio over the number of vehicles. As observed with four protocols, the ratio of network coding is high. The NC_I is the best performance due to the infrastructure-aid. Once the vehicles move outside the sending range of the other neighbors, the infrastructure is working for forwarding packets. It helps vehicles to transmit the packets or coding packets to the farther position. Thus the packets delivery ratio with infrastructure-aid is obviously improved.

Average Delay

In great majority of network application, the end-to-end delay is an important performance that people focus on.

$$D(i) = DT(i) - ST(i), \quad (5)$$

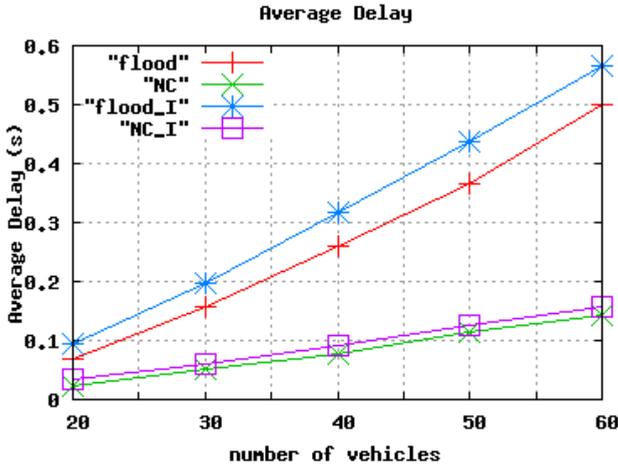


Fig. 3. Average delay for the various number of vehicles

where $D(i)$ denotes the transmitting delay of the Packet i . $RT(i)$ and $ST(i)$ respectively denote the receiving time and sending time of the packet i . In particular, when use the network coding scheme, the decoding time is included in the delay. Thus $D_{NC}(i)$ denotes the time between the first transmission of Packet i and its reception and then successfully decoding at the destination nodes.

$$D_{NC}(i) = D_{NC}T(i) - ST(i) \tag{6}$$

Then, we utilize the average delay \bar{D} as one of performance metrics where N denotes the number of packets.

$$\bar{D} = \frac{1}{N} \sum_{i=1}^N D(i) \tag{7}$$

From Fig. 3, the average delay of NC_I is a little longer than NC. In the infrastructure aid mechanism, the base-station assists the isolated nodes to forward packets. It has more hops than the common mechanism that the delay has a little increased. NC and NC_I are both much better than flood and flood_I yet. Therefore, infrastructure-aid has little negative impact on this performance.

Protocol Overhead

Protocol overhead is defined as providing services to the application to be sent any additional protocol information. Here it is denoted by PO.

$$PO = \frac{S_P}{R_{CBR}} \tag{8}$$

SP is the number of all the transmitted packets in MAC. RCBR is the number of received packet which type is CBR. (Specially, when using network coding, R denotes



the successfully decoded packets.) With the same number of packets sending to the network, NC_I receives more effective coding packets. Therefore, its protocol overhead is lower than NC. The protocol overhead result is described in Fig.4.

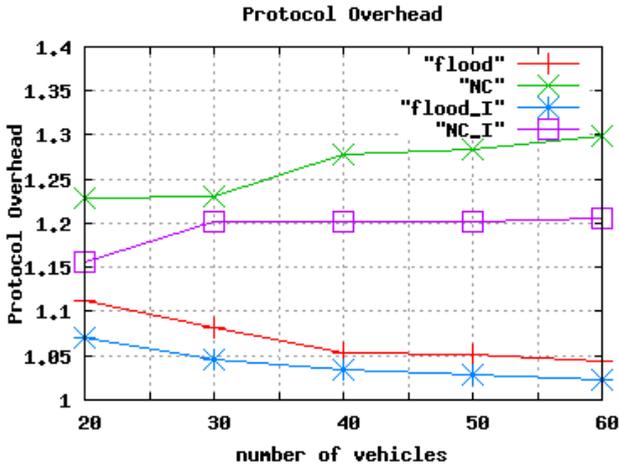


Fig. 4. Protocol overhead for the various number of vehicles

Throughput

The throughput (TH) of network is calculated as the total number of packets received (or decoded), divided by the number of time slots it took until the last packet was decoded.

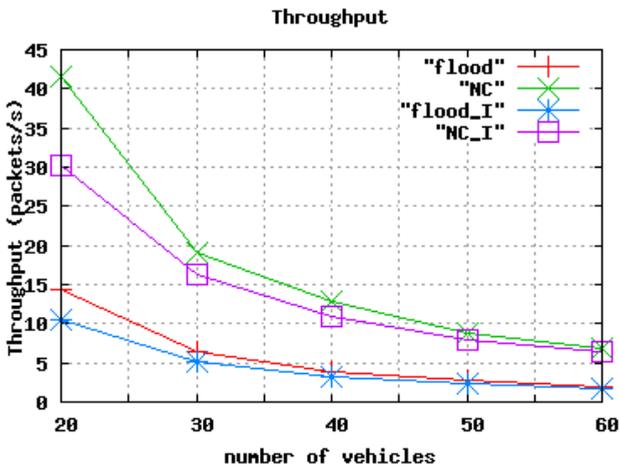


Fig. 5. Throughput for the various number of vehicles

$$TH(i) = \frac{TB(i) - TB(1)}{RT(i) - RT(1)} \quad (9)$$

TB(i) is the number of all the transmitted data when the packet i is successfully received (or decoded) at the destination nodes. RT(i) is the receiving (or decoding) time of the packet i. So TH(i) denotes the average throughput. It is depicted in Fig.5. Since the base-station nodes forwarding information for the vehicles nodes, the connective data sources increase. However, if multi-nodes request to forward messages simultaneously, there must be a collision at the base-station. The throughput is decreased slightly. The results may be due to more neighbor nodes waiting for the matrix full. With the increased number of vehicles in the network, the negative impact on throughput is less and less. By contrast with flood and flood_I, the results of NC and NC_I are both better.

5 Conclusion

In this paper, we concentrated on the infrastructure-aid mechanism for vehicular ad hoc network in urban scene for safe driving application. First we discussed the impact of network coding based protocol. We identified potential problems due to the mobility and non-connectivity of vehicles. To solve this, we proposed the infrastructure-aid information forwarding mechanism. This mechanism improved the packets delivery delay obviously, reduced the protocol overhead and has little negative impact on other performances. The effectiveness of our scheme was demonstrated by simulations.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (Grant No.60972007), the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (Grant No.60921001), and the National Basic Research Program of China (Grant No.2010CB731800).

References

1. Eriksson, J., Balakrishnan, H., Madden, S.: Cabernet: vehicular content delivery using WiFi. In: Proceedings of ACM International Conference on Mobile Computing and Networking (MOBICOM), San Francisco, California, USA, pp. 199–210 (2008)
2. Scheuermann, B., Lochert, C., Rybicki, J., et al.: A fundamental scalability criterion for data aggregation in VANETs. In: MOBICOM, Beijing, China, pp. 285–296 (2009)
3. Karthikeyan, N., Palanisamy, V., Duraiswamy, K.: Performance comparison of broadcasting methods in mobile ad hoc network. International Journal of Future Generation Communication and Networking, 47–58 (2009)
4. Wang, T., Wang, G.: TIBCRPH: Traffic Infrastructure Based Cluster Routing Protocol with Handoff in VANET. In: WOCC (May 2010)
5. Lee, U., Park, J.S., Yeh, J., et al.: CodeTorrent: Content Distribution using Network Coding in VANET. In: MobiShare, Los Angeles, California, USA (September 25, 2006)

6. Li, M., Yang, Z., Lou, W.: CodeOn: Cooperative Popular Content Distribution for Vehicular Networks using Symbol Level Network Coding. *IEEE Journal On Selected Areas In Communications* 29(1), 223–235 (2011)
7. Korkmaz, G., Ekici, E., Ozguner, F., et al.: Urban multihop broadcast protocol for inter-vehicle communication systems. In: *Proceedings of the 1st ACM International Workshop on Vehicular ad Hoc Networks*, pp. 76–85 (2004)
8. Fragouli, C., Widmer, J., Boudec, J.-Y.L.: Efficient broadcasting using network coding. *IEEE/ACM Trans. Networking* 16(2), 450–463 (2008)

Research on Evaluation and Optimization of Agricultural Technology Innovation Resources Allocation in China

Jingcui Sun^{1,2} and Qingran Guo³

¹ Northeast Forestry University, Harbin, China

² Heilongjiang College of Construction, Harbin, China

³ Henan Institute of Science and Technology, Xinxiang, China
qing19328@sina.com

Abstract. In this paper the allocation process of agricultural technology innovation resources is described by a two-stage DEA (data envelopment analysis) model. The allocation efficiency of agricultural technology innovation resources upon the continuous 17 periods in China are evaluated, and the projection analysis of the non-DEA efficient regions was conducted, and a novel optimization scheme designed. And a corresponding improvement strategy was proposed.

Keywords: agricultural technology innovation, resources allocation, efficiency evaluation, DEA model.

1 Introduction

Resources for agricultural technology innovation refer to the required various input elements of agricultural technology innovation, that is, the general term of human resources, financial resources, material resources, information resources and organizational resources of agricultural technology innovation. They are important strategic resources to achieve agricultural technology innovation. The purpose of the allocation of agricultural technology innovation resources is to improve the efficiency of the use and allocation of agricultural technology innovation resources, and obtain maximum outputs of science and technology and economy with minimum inputs of resources in a certain period, within a certain space. Efficiency calculation methods include ratio method, parametric and non-parametric method. In the non-parametric method, DEA method has a unique advantage for evaluating the efficiency of multiple input-multiple output problems. Therefore, this paper selects DEA method to evaluate the allocation efficiency of agricultural technology innovation resources.

2 The Building of an Evaluation Index System

In this paper a comprehensive process of agricultural technology innovation is described by a two-stage model, and the output efficiency of agricultural science and technology (the output efficiency of agricultural technology innovation) and the output

efficiency of agriculture economy (the conversion efficiency of agricultural technology innovation) were respectively evaluated. Therefore, the indicators for allocation evaluation of agricultural technology innovation resources should be designed in two stages (see Fig. 1). Taking into account the availability of the data, during the process of selecting input indicators, human resources and financial resources of agricultural technology innovation are the two types of indicators that are emphatically stressed.

Evaluation indicators of the first stage. ① Input indicators: the two categories of indicators (human resources and financial resources of agricultural technology innovation) are mainly considered as the input indicators for agricultural technology innovation resources. The former selects Personnel Engaged in S&T Activities in Agriculture (I11), Scientists and Engineers among Personnel Engaged in S&T Activities in Agriculture (I12), and Agricultural Professional and Technical Personnel in State-owned Enterprises and Institutions (I13) to reflect the human resources inputs of agricultural technological activities in China from the overall quantity and quality. The latter selects Science and Technology Promotion Funds in Agriculture (I14), and Internal Expenditure on S&T Activities in Agriculture (I15) to reflect the financial resources inputs of agricultural technological in China.

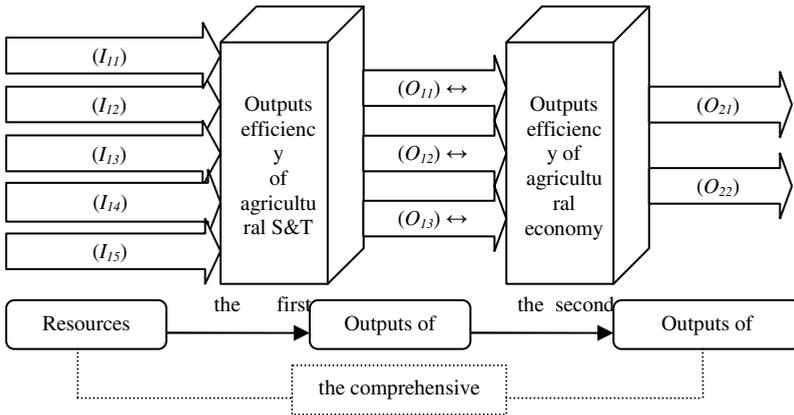


Fig. 1. A two-stage model of efficiency evaluation of agricultural technology innovation resources allocation in China

② Output indicators: Three kinds of indicators (Scientific and Technological Papers in Agriculture (O11), Patents Application Accepted in Agriculture (O12), Major Research Results of Science and Technology in Agriculture (O13)) are mainly considered as the output indicators for agricultural technology innovation to measure the outputs in agricultural science and technology of the first stage.

Evaluation indicators of the second stage. To measure the output efficiency of agriculture economy, the outputs of science and technology of the first stage serve as the inputs of the second stage, namely, the input indicators of the second stage include Scientific and Technological Papers in Agriculture (I21), Patents Application Accepted in Agriculture (I22), Major Research Results of Science and Technology in Agriculture (I23).



At the second stage, Gross Output Value of Agriculture, Forestry, Animal Husbandry and Fishery (O21), and Annual Per Capita Net Income of Rural Households (O22) are selected to measure the agricultural economy outputs of the second stage.

3 Applications of DEA Model of Allocation Evaluation of Agricultural Technology Innovation Resources in China

Data acquisition and processing. According as the two stages of agricultural technology innovation contain a certain time delay, in order to more accurately measure the output efficiency of science and technology in agriculture and economy in agriculture, the two-stage input-output delay time is taken for 1 year.

We will use C2R model selected from DEA model to evaluate the allocation efficiency of agricultural technology innovation resources in China. Each period serves as a decision making unit, and the data during the period 1990-2008 in China will be analyzed. Then 17 decision making units (DMU) in each stage are formed. In order to maintain the comparability of the data, they are taken from the corresponding years' "China Statistical Yearbook", "China Statistical Yearbook of Science and Technology", "China Education Yearbook".

Correlation Analysis of the data. The use of DEA model is under the assumption that the output and input indicators are in the same trend, that is, if the quantity of inputs increases, the quantity of outputs shall be not reduced. To this end, Pearson correlation analysis of the empirical data at the two stages was respectively carried out using SPSS15.0 software (Table omitted). The results show: At the first stage, all the input and output indicators during the period 1990-2008 in China were analyzed. Among them, the three indicators (Personnel Engaged in S&T Activities in Agriculture, Scientists and Engineers among Personnel Engaged in S&T Activities in Agriculture, and Major Research Results of Science and Technology in Agriculture) and other indicators are negative correlated or not significant positive correlated. Therefore, the correlation analysis will be again conducted after excluding the above three indicators. The results show that the new combination indicators comply with directional requirements; At the second stage, all the input indicators and output indicators are obviously positive correlated. Thus, the above indicators can be used in DEA model analysis.

Evaluation results and analysis. In this paper the allocation evaluation of agricultural technology innovation resources in China is analyzed using EMS Version1.3.0 software. At the first stage we use the data during the period 1990-2007, at the second stage we use the data during the period 1991-2008, at the comprehensive stage we use the data during the period 1990-2008. First of all, the values of overall efficiency θ_1 , θ_2 and θ_3 of the three stages are respectively calculated (Table omitted), and their change trend is analyzed. Then according to the calculated amounts of input redundancy and output deficiency of the non-DEA efficient periods, the projection analysis for this period was conducted. Finally, an improvement strategy of the allocation efficiency of agricultural technology innovation resources in China is proposed.

The change trend analysis of overall efficiency value θ of the three stages. The overall efficiency values of the three stages are plotted on the suture removal map and are analyzed (see Fig. 2). At the first stage, these are 8 periods, whose overall efficiency

values are DEA efficient, namely, the first period (the period 1990-1991), the second period, the sixth period, the twelfth period, the fourteenth period, the fifteenth period, the sixteenth period and the seventeenth period. At the second stage, these are 5 periods, whose overall efficiency values are DEA efficient, namely, the fourth period, the fifth period, the sixth period, the seventh period and eighth period. At the comprehensive stage, these are 6 periods, whose overall efficiency values are DEA efficient, namely, the fourth period, the fifth period, the sixth period, the eighth period, the eighth period, the sixteenth period and the seventeenth period. From Figure 2 we can clearly show that only the sixth period (the first stage is in the period 1995-1996, the second stage is in the period 1996-1997, the comprehensive stage is in the period 1995-1997) is DEA efficient at three stages. And its overall efficiency value converges on a point 1. It shows that the allocation of agricultural technology innovation resources in this period is optimal. Overall, from the three lines we revealed that the efficiency value θ_1 of the first stage is relatively high, in addition to the tenth period, the values of overall efficiency were above 0.9. The efficiency value θ_2 of the second stage is relatively low, the values of overall efficiency mostly were at low levels and they are more substantial changes, minimum value of overall efficiency of the second period (the period 1992-1993) reached 0.5659. Comparing with the previous two stages, the efficiency values θ_3 of the comprehensive stage are in the middle level, namely, generally is above 0.8. This means that although the output efficiency of agricultural science and technology in China is higher, but the conversion efficiency from the outputs of agricultural science and technology to the outputs of agricultural economy is lower.

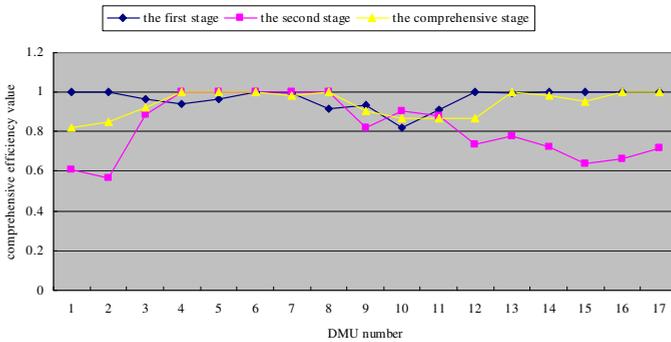


Fig. 2. The change trend analysis of overall efficiency value of the three stages

Projection analysis. When DMU are DEA efficient ($\theta=1$ and all slack variables S^- and S^+ are 0), the allocation of agricultural technology innovation resources in this year is optimal. Speaking of optimal, it means the existing inputs fully play a role and the current outputs are the maximum outputs that can be achieved under the existing inputs. When DMU is non-DEA efficient ($\theta \neq 1$, or $\theta = 1$ and the slack variables S^- and S^+ are not all 0), it can be improved from either inputs or outputs to achieve DEA efficient. According to According to the amounts of input redundancy and output deficiency of the non-DEA efficient years (or known as "adjusted value"), an objective value, which will make the non-efficient period into efficient period (or known as "ideal value"), that is, an optimization scheme of the allocation efficiency, can be calculated using "projection" theory. (see Table 1, Table 2 and Table 3).

Table 1. Adjusted value and ideal value of the inputs at the first stage

| DMU | (I13) (10 000 person) | | (I14) (100 million yuan) | | (I15) (10 000 yuan) | |
|-----------|-----------------------|-------------|--------------------------|-------------|---------------------|-------------|
| | Adjusted value | Ideal value | Adjusted value | Ideal value | Adjusted value | Ideal value |
| | (s_1^-) | | (s_2^-) | | (s_3^-) | |
| 1992-1993 | 1.41 | 84.09 | 0.00 | 2.89 | 43075.13 | 183054.63 |
| 1993-1994 | 7.09 | 83.09 | 0.00 | 2.83 | 74499.79 | 189787.68 |
| 1994-1995 | 0.00 | 92.44 | 0.00 | 2.89 | 33514.38 | 311715.76 |
| 1996-1997 | 0.00 | 99.34 | 0.00 | 4.91 | 0.00 | 396481.14 |
| 1997-1998 | 0.00 | 95.12 | 0.00 | 5.03 | 0.00 | 336320.34 |
| 1998-1999 | 47.7 | 55.03 | 0.00 | 8.54 | 0.00 | 398181.43 |
| 1999-2000 | 0.00 | 90.36 | 0.00 | 7.49 | 0.00 | 398886.13 |
| 2000-2001 | 0.00 | 101.69 | 0.00 | 8.91 | 0.00 | 468244.36 |
| 2002-2003 | 0.00 | 106.77 | 0.00 | 9.84 | 18364.43 | 688944.67 |

Table 2. Adjusted value and ideal value of the outputs at the first stage

| DMU | (O11) (piece) | | (O12) (piece) | |
|-----------|----------------|-------------|----------------|-------------|
| | Adjusted value | ideal value | Adjusted value | ideal value |
| | (s_1^+) | | (s_2^+) | |
| 1992-1993 | 0.00 | 10921 | 202.79 | 1544.79 |
| 1993-1994 | 0.00 | 10588 | 0.00 | 1538 |
| 1994-1995 | 0.00 | 9510 | 0.00 | 1845 |
| 1996-1997 | 1762.46 | 11089.46 | 0.00 | 2685 |
| 1997-1998 | 2323.32 | 11847.32 | 0.00 | 2581 |
| 1998-1999 | 1031.52 | 12898.52 | 0.00 | 3534 |
| 1999-2000 | 2749.97 | 14058.97 | 0.00 | 3420 |
| 2000-2001 | 4185.86 | 16294.86 | 0.00 | 4027 |
| 2002-2003 | 0.00 | 17771 | 0.00 | 4835 |

At the first stage 9 periods, for example, the period 1998-1999, need to be improved. In the current efficiency state ($\theta_1 = 0.9339$), according to the original inputs, namely: the number of Agricultural Professional and Technical Personnel in State-owned Enterprises and Institutions is 1.1 million person, the number of Science and Technology Promotion Funds in Agriculture is 914 million yuan, Internal Expenditure on S&T Activities in Agriculture is 4.263641 billion yuan. The current outputs can be obtained, namely: the number of Scientific and Technological Papers in Agriculture is

11867 pieces, the number of Patents Application Accepted in Agriculture is 3534 pieces. There are the amounts of inputs redundancy and outputs deficiency, namely: $s_{1-} = 0$, $s_{2-} = 0$, $s_{3-} = 47.7$, $s_{1+} = 1031.52$, $s_{2+} = 0$. That will need to be adjusted using projection theorem.

Table 3. Adjusted value and ideal value of inputs and outputs at the second stage

| DMU | (I21) (piece) | | (I22) (piece) | | (O21) (100 million yuan) | | (O22) (yuan) | |
|-----------|--------------------------------|-------------|--------------------------------|-------------|--------------------------------|-------------|--------------------------------|-------------|
| | adjusted value (s_{1-}) | ideal value | adjusted value (s_{2-}) | ideal value | adjusted value (s_{1+}) | ideal value | adjusted value (s_{2+}) | ideal value |
| 1991-1992 | 3180.65 | 3871.44 | 0.00 | 751.04 | 14.18 | 9098.88 | 0.00 | 783.99 |
| 1992-1993 | 1305.94 | 4963.1 | 0.00 | 886.77 | 0.00 | 10995.5 | 0.00 | 921.62 |
| 1993-1994 | 1492.92 | 8198.38 | 0.00 | 1190.90 | 0.00 | 15750.5 | 0.71 | 1221.69 |
| 1999-2000 | 0.00 | 9709.58 | 260.23 | 2631.29 | 81.18 | 24996.98 | 0.00 | 2253.42 |
| 2000-2001 | 0.00 | 10196.19 | 320.32 | 2763.15 | 70.66 | 26250.26 | 0.00 | 2366.4 |
| 2001-2002 | 0.00 | 10666.82 | 656.7 | 2890.68 | 71.14 | 27461.94 | 0.00 | 2475.63 |
| 2002-2003 | 0.00 | 11310.43 | 287 | 3234.94 | 0.00 | 29691.8 | 0.00 | 2622.24 |
| 2003-2004 | 0.00 | 13777.86 | 0.00 | 3748.58 | 0.00 | 36239 | 253.63 | 3190.03 |
| 2004-2005 | 0.00 | 14994.58 | 0.00 | 4232.13 | 0.00 | 39450.9 | 219.51 | 3474.44 |
| 2005-2006 | 0.00 | 15513.24 | 0.02 | 4341.70 | 0.00 | 40810.8 | 6.76 | 3593.8 |
| 2006-2007 | 0.00 | 18585.11 | 0.00 | 5235.74 | 0.00 | 48893 | 165.54 | 4305.9 |
| 2007-2008 | 0.00 | 22044.16 | 0.00 | 6309.63 | 0.00 | 58002.2 | 348.59 | 5109.21 |

We can improve it from the input perspective while maintaining the original outputs, using the formula Eq. 1, "projection" of non-efficient periods is obtained in the production frontier, namely, in order to achieve relatively effective allocation efficiency, the amount of input indicators should be reduced to 0.933 9 times as that of their original inputs, and the value of input indicators continued to decline s_{-} . After calculating the ideal values of the inputs: the number of Agricultural Professional and Technical Personnel in State-owned Enterprises and Institutions is 0.5503 million person, the number of Science and Technology Promotion Funds in Agriculture is 854 million yuan, Internal Expenditure on S&T Activities in Agriculture is 3.9818143 billion yuan. We can improve it from the output perspective while maintaining the original inputs, using the formula Eq. 2, After calculating the ideal values of various outputs: the number of Scientific and Technological Papers in Agriculture is 12898.52 pieces, the number of Patents Application Accepted in Agriculture is 3534 pieces.

$$X'_j = \theta * x_j - s_{-} \tag{1}$$

$$Y'_j = Y_j + s_{+} \tag{2}$$



At the second stage 12 periods, for example, the period 1999-2000, need to be improved. In the current efficiency state ($\theta_2=0.8182$), according to the original inputs, namely: the number of Scientific and Technological Papers in Agriculture is 11867 pieces, the number of Patents Application Accepted in Agriculture is 3534 pieces. The current outputs can be obtained, namely: Gross Output Value of Agriculture, Forestry, Animal Husbandry and Fishery is 249.158 billion yuan, Annual Per Capita Net Income of Rural Households is 2253.42 yuan. There are the amounts of inputs redundancy and outputs deficiency, namely: $s_1^- = 0$, $s_2^- = 260.23$, $s_1^+ = 81.18$, $s_2^+ = 0$. That will need to be adjusted using projection theorem.

We can improve it from the input perspective while maintaining the original outputs. The ideal values of the inputs are obtained: the number of Scientific and Technological Papers in Agriculture is 9709.58 pieces, the number of Patents Application Accepted in Agriculture is 2631.29 pieces. We can improve it from the output perspective while maintaining the original inputs. The ideal values of the outputs are obtained: Gross Output Value of Agriculture, Forestry, Animal Husbandry and Fishery is 249.9698 billion yuan, Annual Per Capita Net Income of Rural Households is 2253.42 yuan.

Improvement strategies. Based on the projection analysis, it is easy to design strategies to improve the allocation efficiency of agricultural technology innovation resources for each period.

That period, which was an example in the above, still is taken for example. As for the case of the above, the following measures can be taken to improve its allocation efficiency of agricultural technology innovation resources: In order to improve the output efficiency of agricultural technology innovation, the number of Agricultural Professional and Technical Personnel in State-owned Enterprises and Institutions should be decreased appropriately. Or it is appropriate to increase the number of Scientific and Technological Books in Agriculture. In order to improve the conversion efficiency of agricultural technology innovation, it should be appropriate to decrease the number of Patents Application Accepted in Agriculture. Or it is appropriate to increase Gross Output Value of Agriculture, Forestry, Animal Husbandry and Fishery.

4 Conclusion

The analysis results show that the efficiency value θ_1 of the first stage is relatively high, the efficiency value θ_2 of the second stage is relatively low, the efficiency values θ_3 of the comprehensive stage are in the middle level. This means that although the output efficiency of agricultural science and technology in China is higher, but the conversion efficiency from the outputs of agricultural science and technology to the outputs of agricultural economy is lower.

In order to further explore approaches and methods to improve allocation efficiency of agricultural technology innovation resources, projection analysis of non-efficient periods was conducted and an optimization scheme was designed, the strategies to improve the allocation efficiency of agricultural technology innovation resources were proposed.

As for the state of relatively efficient outputs and relatively inefficient conversion of agricultural technology innovation, one must continue to strengthen industrial applications of agricultural technology innovation and to achieve efficient interface between technology and economy.

References

1. Guan, J., He, Y.: Evaluation of Regional Innovation System Based on DEA Method. *Studies in Science of Science* 23(4), 265–272 (2005)
2. Hu, M.: *Regional Innovation Systems: Evaluation, Development Pattern and Policy*, 1st edn., pp. 44–56. Hu Nan University Press, Changsha (2008)
3. Duan, Y.: *Data Envelopment Analysis-Theory and Application*, 1st edn., pp. 4–14. Shanghai Popular Science Press, Shanghai (2006)
4. Cui, N., Guo, X.: *Economic Analysis on Soybean Production Technology and Its Application in China*, 1st edn., pp. 101–126. China Agriculture Press, Beijing (2010)

CJSYS-A01 Piezoelectric Fluidic Angular Rate Sensor

Xing Wang, Linhua Piao, and Quangang Yu

Sensor Technology Research Centre
Beijing Information and Technology University
Beijing, China
wang_zhw@qq.com

Abstract. The CJSYS-A01 piezoelectric fluidic angular rate sensor sensitive mechanism was described. The sensor's dynamic characteristics, temperature performance and overload performance were analyzed, then the sensor circuit was designed, and the performance was tested. At last, the sensor environment experiment results were provided. The experiment results show that sensor is provided with the new principle, wide dynamic range, high accuracy and strong impact, high cost performance, which is a practical value new type sensor and good for development prospects.

Keywords: Piezoelectric fluidic angular rate sensor, Sensitive mechanism, Circuit.

1 Introduction

The core component of the angular rate sensor is high-speed rotor or vibration beam. These components are solid, great quality, large shock and strong vibration, the excessive inertia force will damage rotor. The CJSYS-A01 piezoelectric fluidic angular rate sensor is a solid-state inertial device, which uses Coriolis forced airflow circulation beam deflection, and achieve measurement angular parameters. The characteristics are used gas as sensitive quality, no high-speed rotor, and the response time is less than 80ms, long operating life and low cost. It can be withstand the 16000g impact [1]. The sensor can be used in tanks, artillery, missiles, ships and other areas, especially for civil industries.

2 Sensitive Mechanism

Piezoelectric fluidic angular rate sensor is used the Coriolis forced airflow circulating beams deflect its original track to achieve parameters measured. The circulating airflow (jet) is generated by piezoelectric pump at speed V_j signal, which is sensitive by the two parallel hotwires [2]. When the angular rate ω_i inputs, because of Coriolis acceleration effect, jet flow beams deflect the center position, it cools hotwires (r_1, r_2), so the bridge lose balance, and output the proportional to the angular rate electrical signal V_0 , which is shown in the figure 1. In order to further research the signal characteristics, the output signal amplified 100 times, the sensor sensitivity will reach $20\text{mV}/^\circ\text{s}$.

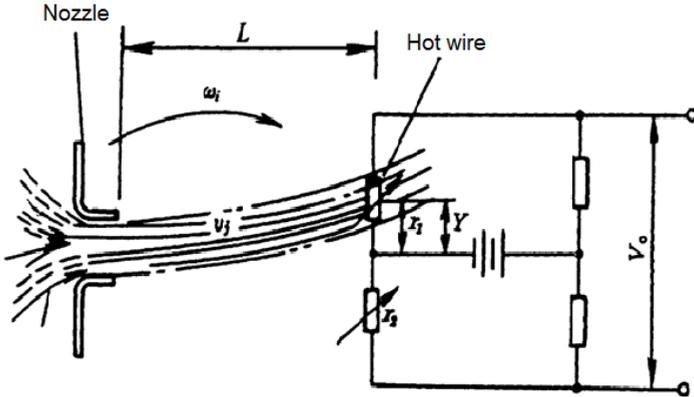


Fig. 1. Working principle of the sensor

3 Characteristics

A. Dynamic Characteristics

In order to research the performance of the sensor and the output of the practical processing, and measure the relationship between amplified output signal and angular rate. The relationship between the bridge output amplified signal and angular rate is measured, which is shown in the figure 2.

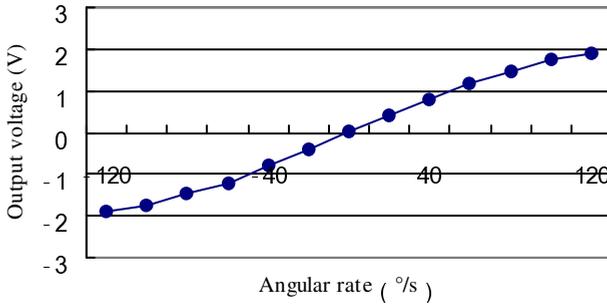


Fig. 2. Relationship between the output of the bridge and the angular rate

The sensor outputs within a small angular to satisfy $V=V_0+Kr$, the output signal can be seen as a linear, ω is the angular rate, the unit is $^\circ/s$, and Kr is a constant. But when the angular rate is more than $45^\circ/s$, the non-linear increase, and the output signal must be compensated. Through compensation, the effective width of sensor dynamic range is $\pm 120^\circ/s$. In this paper, the linear is less than 1% by using the least square method.

Drift exist in the output of the bridge, this mainly due to hotwires and the thermal properties of the surrounding gas, such as the stability of the hotwires, and the gas thermal conductivity. This drift is regular, can be solved with the rational selection and

design; noise existed in the output, too, this is mainly due to the amplifier circuit, thermal noise from hotwire is small. In addition, it is reported that the hot volume air flow exist perturbation. We also measured the resolution of the sensor. When a clear signal outputs, the sensor's sensitivity is better than 0.01°/s.

The hotwires response time constant is:

$$M = \frac{C}{A + B\sqrt{v} - I^2} \quad (1)$$

It can be seen from the formula: the time constant M is determined by the fluidic beam velocity v , the hotwires current I and the heat capacity C . Reducing the heat capacity of the thermal or increasing the fluidic beam velocity will reduce the response time. Therefore, we select the hotwires, when the strength, coefficient of temperature sensitive satisfied; to reduce the response time, we should select with a small heat capacity of the hotwires. The hotwire heat capacity depends on three factors: material, geometry and work status. Made in different material, the response time could be twice time, in which the platinum wire has the smaller response time. Thermal wire certain length, the response time determined by the diameter of the wire. For the hotwire, the smaller the diameter, the shorter the response time. However, if the hotwire certain overheating ratio, the thinner the diameter, the smaller the current; we can see from the above equation, the time constant also has changes. Therefore, for certain overheating ratio, the diameter and the current design optimization problems exist. Time constant is the result of hotwire shape, operating current and fluidic velocity. We made a lot of experiments on the piezoelectric fluidic angular rate sensor, and the step response curve shows in the figure 3. The response time is less than 80ms.

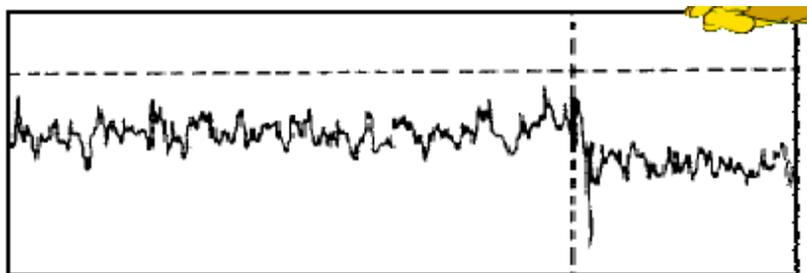


Fig. 3. The sensor step response curve

B. Temperature Performance

The sensor zero output curve with the change of temperature is shown in the figure

4. Bridge output signal varies with temperature, there are three reasons:

(1) Gas thermal conductivity is different under the different environment temperature.

(2) The temperature coefficient is different at two hotwires.

(3) The hotwires are thermal sensitive component, so the bridge output signal should be a function between angular rate and temperature, which is $V=V(\omega,T)$, ω is the angular rate and T is the environment temperature.

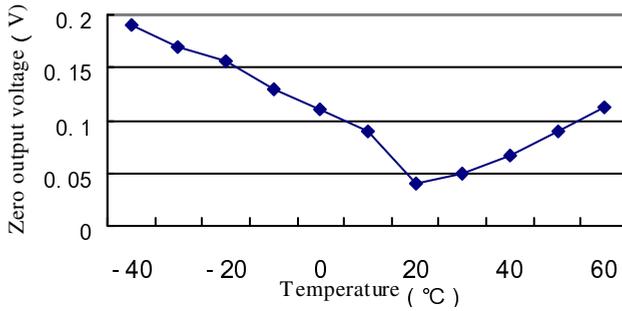


Fig. 4. Sensor zero output voltage with the temperature change curve

C. Environment Performance

1) Vibration Test

Tighten the sensor to the mounting surface in the vibration table, the displacement amplitude transferred to 0.75mm, and then in the 10~55Hz frequency range continuously swept back and forth. In three mutually perpendicular directions of the sensor for test 2h. Logarithmic sweep swept with speed 1oct/min. After the test, the sensor test in the normal atmospheric conditions, detect the sensor output. The value is less than 0.2°/s, the sensor is not short circuit, open circuit and mechanical damage.

2) Impact Test

Pre-reconcile the impact machine; tighten the sensor to the mounting surface in the impact machine, according to the following conditions imposed the impact on the test sample.

- (1) Acceleration peak value: 300m/s²
- (2) Nominal pulse duration: 11ms
- (3) Wave shape: half sinusoid
- (4) Velocity change: 2.07m/s

Applied sensor the three mutually perpendicular axis six directions with each one three times impact. After the experiment, detect the sensor output in the normal atmospheric conditions, and the value is less than 0.2°/s, the sensor is not short circuit, open circuit and mechanical damage.

3) Overload Capacity

Tensile strength of the annealed platinum wires is 15kg/mm². For example, with a diameter 10 μ m, 10mm long platinum wires, the net weight 16.8 μ g, which could bears 1.18g Tension, so it can bears 70kg overload. In all, the sensor overload capacity depends on mechanical structure and hotwires resistance strain coefficient.

4 Circuits

After the bridge output signal amplifies, some problems could exist, such as noise, drift, zero and sensitivity change with temperature. So the signal is to process, the circuit diagram is shown in the figure 5.

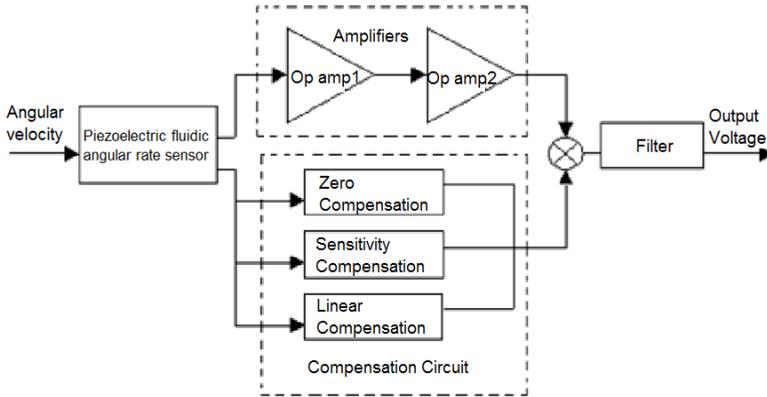


Fig. 5. Circuit diagram

5 Experimental Results

The testing results show that the sensor measurement range is $\pm 120^\circ/\text{s}$, sensor resolution is less than $0.01^\circ/\text{s}$, nonlinearity is less than 1%, cross coupling is less than 1%, response time is less than 80ms and the zero voltage repeatability is $0.2^\circ/\text{s}$.

6 Conclusions

The piezoelectric fluidic angular rate sensor is not high-speed rotor, and use the jet flow deflection sensitized angular rate. Sensor is provided with the new principle, wide dynamic range, high accuracy and strong impact, high cost performance, which is a practical value new type sensor and good for development prospects.

Acknowledgment. National Natural Science Foundation of China (60772012); Supported by program for New Century Excellent Talents in University; Beijing Excellent Training Grant; Supported by program for Beijing Key Laboratory Open Topic; Supported by program for Beijing New Century Talents Training Grant Project; Supported by program for Beijing Education Commission Technical Innovation Platform Foundation (KM201110772020); Supported by program for Modern Detection and Control Technology Ministry of Education Key Laboratory.

References

1. Zhang, F.: Sensor Electronics. National Defence Industry Press, Beijing (1991)
2. Chen, Z., Zhang, F., Li, X.: Gas tilting angle sensor. Journal of Piezoelectric and Acousto-optics 2, 21–22 (1995)

Research on Evaluation of Clients' Service Performance to Bonded Logistic Enterprise Based on Entropy Weight

Li Zhen-fu, He Jian-tong, and Chen Ping

College of Transportation Management, Dalian Maritime University, Dalian, China
lbzh1978@126.com

Abstract. Just for there is no certain theories and methods as proof for clients to select enterprises in bonded logistic zone, which always make blind choice, and cause unnecessary miscarriage. Thus, exact analysis and evaluation of clients' service performance in bonded logistic enterprise can help clients to select bonded logistic enterprise as well as improvement and advancement of clients' service performance to bonded logistic enterprise. On the basis of analysis to clients' service performance to bonded logistic enterprise, this thesis evaluate clients' service performance to bonded logistic enterprise by Entropy Weight Multi-objection Decision-making method for the first time and receive accurate conclusion. Evaluation of clients' service performance to bonded logistic enterprise is highlight of performance evaluation of bonded logistic enterprise which plays an important role in advancing regional bonded logistic benefit and regional economic benefit.

Keywords: Bonded logistics, bonded logistics enterprise, performance evaluation, entropy weight.

1 Introduction

To be a branch of the logistic industry, the bonded logistic has a striking characteristic of "bonded", and is more professional than other style of logistic and more efficient on some aspects. So, it is acting an increasingly important role in our national economic development. With the impacting of globalization, the economy of our country and the global has been combined more and more closely. And many foreign-capital enterprises have set up processing factories or bonded warehouses; they also require the bonded logistics urgently, so it brings a fast developing market for the bonded logistics enterprises in China.

Now, in China, there are several bonded logistic zones like Waigaoqiao in Shanghai, and increasingly numbers of enterprises entered into the bonded zone or used the benefit of it to carry out logistics projects. Because the behavior of bonded logistics enterprises are like "crossing the river by feeling the stones", and there are always appeared some problems for these enterprises, so the enterprises really need to know how to find out and solve these problems, on the other side, these are also belong to what the management in the bonded zone and the customs and other regulatory authorities need to understand and solve by coordination. At the same time, when choosing the bonded logistics enterprise,

without any theory or method, customers are easy to make blindly choose, and cause unnecessary fault. Therefore, analyzing and evaluating the customer service performance of the bonded logistics enterprise correctly can help customers to select the bonded logistics enterprises better, and also benefit for the bonded logistics enterprises to improve the customer service performance.

2 Bonded Logistic Enterprise and Its Related Research

Bonded Logistic and Bonded Logistic Enterprise

Bonded logistic is generally defined as: goods in a state of bonded during the process of import and export, and all the logistic activities about them like transportation, storage, processing and so on are in the supervision of customs. The bonded logistic mentioned in the article is a special kind of logistics, it refers that under the supervision of functional departments like customs, the enterprise do the storing, simple processing, and increasing the goods' value for the export and import bonded goods and the goods which had not fulfill the customs procedures, to meet the need like global purchasing, International distribution and distribution, and international transfer, etc. So it can simplify the logistic links, optimize the logistics process, and reduce the cost of the logistics. In short, the bonded logistics is a form for realizing the bounded goods' flow by using the bonded warehouses and bonded zone.

Bonded logistic fit for the general regulation of logistic scientific, but also has its own typical characteristics which are different from other kinds of logistic:

(1) System border crossing. On geography, the bonded logistic goods are transport in one country, so it belongs to the domestic logistic, but it also has obvious characters of international logistic. For example, bonded zone, bonded logistic center, and district-port joint all has the character of "passing the customs within the boundaries". So they can be regarded as a relay zone between the international logistic and the domestic logistic.

(2) Enlargement of the logistic factors. The factors of logistic usually include transportation, storage, information service, distribution, etc. But beside these basic ones, the bonded logistic also have some key factors like customs supervision, passing the port, bonded, applying to customs, drawback, etc. the basic characters and key ones combined closely and form a complete bounded logistic system.

(3) Whole process management. A basic process of passport for the general trade cargo includes the declaration, examination, levying of taxes, and it is the management of "point" style; but the process of bonded goods is a whole process from entering a country, storing or processing to reexport. The beginning is entering customs, and the finishing point is writing off of goods, so it is a whole process of management.

(4) Bottleneck problems about efficiency. Operating the logistic process under the supervision of customs is the natural difference between bonded logistics and other ones. And in order to achieve an effective supervision, it is necessary for the customs to perform a strict process, complicated procedure, and a high rate of a random inspection. But these are operating to the requirements of operating modern logistics which include convenient, efficient, and low cost. So the relationship between efficiency of logistics and efficacy of customs supervision could be called "Dual Contrary". In short, in the

case of increasing demands of the bonded, the efficiency of customs supervision is a bottleneck problem for the efficiency of the bonded logistics system.

(5) Platform To processing trade enterprises, bonded logistics is the end of supply logistic and the beginning of trade logistic, and even the production logistic (e.g. VMI). Operation efficiency of bonded logistics directly affects the normal production and supply chain function properly of enterprises, and a normal and efficient bonded logistics system is the efficient coordination result of the customs, the government, the logistics enterprises, and ports. So both the perfect policy system and incorporate platforms for integrated logistics service are necessary, for example, the logistic center united by material, financial and information flows will be the main form of bonded logistic.

The bonded logistic enterprise in this article refers to the third-party logistic enterprise whose main business is providing bonded logistic, and its business about bonded logistic accounted for over 60%~80% of its total business volume.

The main difference between bonded logistic and other ones is that, the bonded logistic enterprises are operating the international logistic, and they can achieve the true meaning modern logistics which is the most perfect logistics mode in modern society. When raw material or goods enter into the bonded zone, the bonded logistics enterprises will do a series of works like appropriate storage, processing, cargo agent, warehouse receipt pledge, transportation and financial services (the foreign exchange purchase)etc. And truly realize the interaction among material, financial and information flows. To be the third-party logistics enterprises in the bonded zone, the bonded logistic enterprise should put the idea of this modern logistic throughout the services of the whole supplychain.

The present research situation for the customer service performance evaluation of bonded logistics enterprises

Now, there are not many research results about customer service performance evaluation of bonded logistics enterprises that have been proposed specifically by experts and scholars from home and abroad, most researches are mainly for the general logistics enterprises performance evaluations.

In 1999, the American scholar Mr. Bowersox proposed that measuring the logistic enterprises performance should from the aspects of inside and outside, and the customer service was listed to be one of the indices for measuring the inside performance. But at that time, the content of customer service is very limit. In 2001, SunHongling and QiShijun[1] measured the logistics performance from the aspects of customer service, transportation, storage, logistics cost control, etc. and proposed the evaluation method in which the efficacy factor method is major, with the support of comprehensive analysis method. After that, WangCheng and WangYan[2,3]further extended connotation and extension of logistics customer service, so the logistics customer service would get more emphasis when measuring the logistics enterprises performance.

At the same time, the methods of logistics enterprises performance evaluation was coming in more and fast. ZhouTao[4] and other people proposed using fuzzy comprehensive judge method to evaluate the logistics enterprises performance; LiGuan and HeMingxiang proposed using data enveloping analysis method to evaluate the logistics performance; WeiXinjun proposed using fuzzy cluster analysis method; MaHongyan and other people used the effectiveness theory method; WangYing and

other people used DEA/AHP two-stage model to evaluate the logistic; LiuBinglian and WangPengji build a "3+1" analysis model for logistics enterprises performance base on the System of Balanced Scorecard. JiangJunfeng(2004) combined the Delphi technique and fuzzy comprehensive evaluation method to evaluate the logistics customer service level. ZhouTao(2005)used the analytic hierarchy process (AHP)method to evaluate the logistics enterprises performance. LiuMingfei and LiLeiyu(2005)proposed the principles for establishing evaluation indices and the basic indices, they also pointed out that using the judge matrix to ensure the weight of indices, and how to analyze the evaluation results of logistics customer service.

In summary, it is easy to find that the indices of logistics customer service performance is constantly improving and expanding, and the evaluation methods were bring forth the new through the old continuously, but each method have a style of their own, and there is not a systematic and scientific system on evaluating the logistics enterprises customer service.

During the development process of performance evaluation home and abroad, the performance evaluation has been widely used in the traditional areas of financial, marketing, personnel, and production, etc. and shaped many methods and ideas for evaluation. But, logistic is a new area contrastively, and the traditional evaluation methods have their own shortcomings, so it cannot receive good effects when evaluate the logistic customer service by the traditional evaluation methods directly. The shortcomings of traditional performance evaluation method are shown as follows:

(1) The traditional evaluation method cannot reflect the behavior of enterprises dynamically. For example, the financial reporting in the financial evaluation sometimes annually, so it cannot reflect the recent operating state of enterprises, further more, financial ratios are often outdated.

(2) Traditional evaluation methods sometimes limited to some departments or functions, while the evaluation about the quality and delivery of products are beyond the boundaries of departments and functions, which led to fail of effectively evaluation, and can not give information to support the various decision-making of enterprises.

(3) Traditional evaluation methods always focused on financial ratios but neglect the evaluation of non-financial ones.

(4) Traditional evaluation methods are often lack of feedback function, so the evaluation results cannot feedback to the operation of enterprises timely and effectively.

In order to overcome the shortage of using the traditional performance evaluation method, some new methods appeared in resent years, for example, the omni-directional performance board method, comprehensive balanced scorecard method, benchmarking evaluation method, key performance indexes evaluation method, etc. and many people have done the research about the evaluation method against logistics customer service, and used the balanced scorecard method, analytic hierarchy process method or fuzzy comprehensive evaluation method to evaluate the logistics customer service performance. These methods have made some contribution to the research on logistics customer service performance evaluation, but they also have some shortages, parts of them are shown as follow:

(1) Balanced scorecard method has overcome the limitation of only using the financial technique to evaluate the performance in tradition, and added the performance evaluation from thee aspects of customer, internal business, study, etc. so make the

object of performance assessment is the same as strategic aim of the enterprise, and evaluate the operation of the enterprise scientifically and reasonably. The advantage of it is ensuring the performance evaluation indices system from the aspect of strategy comprehensively and systematically, but the performance evaluation process after it needs other methods and techniques to finish.

(2)analytic hierarchy process(AHP)method makes the thinking process and subjective judgment of human to be mathematicization, gives the quantitative value to every research question according to the relative importance of every factor, uses the judgment matrix to construct the weight of every index, and analyze the result by the number of the weight. The advantage of it is ensuring the weight scientifically, but shortage is that evaluation results do not fit for the thinking way of human.

(3)The advantage of fuzzy comprehensive evaluation method is close to thinking habits and ways of describing things of human, deal the qualitative indices which is hard to quantified traditionally in the way of fuzzification, and get the final result through the indices weight vector of every level and calculation of their relative judgment matrix, but there is no function of ensuring the indices weight and analysis on the evaluation result reasonably and scientifically.

3 Building of the Evaluation Index System

A Selected Principles of Evaluation Indices

In actual operation, for establishing an index system which can evaluate the bonded logistics enterprises customer service performance effectively, and make the results as objective and comprehensive and scientific as possible, the indices selecting should follow the following principles:

(1) The selected indices must be systematic

Generally, one index can only reflect one aspect of the evaluated objectives, but all the selected indices should be able to present the functional characteristics, economic, social and environmental features, and reflect the complete information of planners, users and nonusers, managers, etc. of bonded logistics enterprises.

(2) The selected indices must be consistent with targets

Target is the guiding for actions, and setting targets is the first step of system engineering analyzes methods. The information about the degree of the targets achieving by the action planning is the main one which the directors most concerned about when directing. It is also a main criterion for measuring an action stand or fall. So the selected indices must be able to reflect information of the degree of the targets achieving.

(3) Correlation between among indices should be avoid

Avoid correlation among indices, so it can eliminate the orientation produced by the correlativity among indices.

(4) The selected indices must be operability

The meaning of the evaluation indices should be clear and easy to be understood; the data for quantification of indices should be easy collected; and it could be solved by existing methods and models.

(5) The indices are made scientific

The single index should be comprehensively on the theory, and it can reflect some information about the performance of an enterprise objectively and scientifically.

(6) The indices should have the comparable character

The nature of the indices should be determinately and could be compared. In other words, the indices could be compared between each other, time series of each index could also be compared, so it can reflect the dynamic changes of the data. We should try to use the data of existing statistical system or a data of focus sampling survey, and make the data to be normalization, then appeared as a relatively form.

(7) The indices should focus the key points

The indices should reflect various aspects of the performance comprehensively and objectively, while focus the key points, and avoid the secondary redundant information took the place of the theme. Indices should be representative.

(8) Combined quantitative indices with qualitative ones

In the performance evaluation of logistics enterprise customer service, there are some influencing factors cannot be quantified, for example, the indexes of market and credit are very abstract, and cannot be quantified exactly by models. But if we don't consider about these effects, the evaluation result would lack of comprehensiveness, so we combine quantitative indices with qualitative ones, and use the quantified technology of qualitative analysis like fuzzy quantify and measure quantify to reinforce the shortage of data in simple quantitative evaluation.

B Evaluation Indices

According to the selected principles and problems of evaluation indices, especially the definition and characters of the bonded logistic enterprises, we determine five indices for evaluation, they are: the distribution process, cargo agent, warehouse receipt pledge, transportation and storage, and the financial services(the foreign exchange purchase). These five indices are built for the customers' interests, and they are different from the business performance of the logistic enterprises. By contrast, the business performance of bonded logistics enterprises shall be inclined to indices of the enterprise profit evaluation index, transport efficiency, the rate of utilizing container yard, etc.

Distribution process is the main function of bonded logistic enterprises. The bonded logistic enterprises provide the service like sub-assembly, assembly, and cutting of the products, so the products can be distributed in the market more adaptive. On the other side, this index is fit for the selected principle mentioned above.

Cargo agent is another important function of bonded logistics enterprises. Because formalities and procedures of the bonded logistic are complicated and varied, and different country has different laws and rules, so the cargo agent need to be operated by the professional bonded logistic enterprises. And these are also a major aspect to measure the ability of the bonded logistic enterprise.

Warehouse receipt pledge, refers that borrowers (accredited party) make the confirm receipt which was registered in the trading market to be a pledge, and use it to apply for Short-term circulating capital loans or short-term financing business from banks. When borrowers (accredited party) failto perform repayment obligations on time, the bank has the right to dispose of the receipt of the pledge and the money would be used to repay the loan principal and interest and other related costs. To be the customer of the

bonded logistics corporations, international trade enterprises always need a large number of working capital, and they use the warehouse receipt pledge for short-term financing, therefore, this business become the main content of the customer service for the bonded logistics enterprises.

Transportation and storage are the necessary business capacity of all the logistic enterprises; they are also the main indices for measuring the customer service performance. So the bonded logistic enterprise should make them to be the major indices for evaluating the customer service level.

Financial service is another customer services item which has notable feature of the bonded logistics. That is because there are numbers of financial service items during the bonded logistics process, for example, the foreign exchange purchase, etc. all of these need to be finished by the bonded logistic enterprises for the customers.

Demarcating Methods of Indices

The indices system mentioned above has both the subjective characters and the objective ones of the customer service performance evaluation in bonded logistic enterprises. Some ones could be qualified directly, such as transport and storage area; but some ones are difficult to be quantified and normalized. In view of the different problems when collecting the indices' data, during actual operating process, we use the demarcating methods listed in table 1.

In these indices, objective ones could be obtained from the customer data or using statistical methods directly ; subjective ones mainly adopt the method of marking by the experts, one grading system, the higher the score, the better the index.

Table 1. Demarcating methods of every index

| Evaluate index | Detail | Acquiring method | Feature of index |
|----------------------------|----------------------------------|-------------------|------------------|
| Circulating and processing | Business coverage | experts assess | Subjective value |
| Cargo agent | Business capacity | Customer data | Objective value |
| Warehouse receipt pledge | Financing means and capabilities | experts assess | Subjective value |
| Transportation and Storage | Acreage | Statistical value | Objective value |
| Financial services | Foreign exchange capacity | experts assess | Subjective value |

4 Entropy Weight Multi-objection Decision-Making Methods

As we have mentioned above, most of the existing evaluation methods are determined by which index is more emphasized by the evaluation staff subjectively. And because the weight determining is mainly rely on subjective judgment of the experts, so the credit of evaluation results is low. Another kind of method is determined by using the respected objective information of the indices data, its original data formed by the actual data from various indices when evaluating the objects, and because the indices have been qualified mostly, but on the other side, some indices of bonded logistics



enterprises customer service performance evaluations cannot be qualified, so these methods are not appropriately.

Base on the principle of "practical and realistic operational ability", this article bring in Multi-objection evaluation method according to the practice when evaluate the customer service performance of the bonded logistic enterprises. In addition, in order to avoid the human interference when deciding the weight factors, we bring in the method of Entropy Weight to confirm the weight factors.

The fundamental thought of the entropy weight method is: the weight coefficient should be the changing level of every index in the indices group and measurement of their effects for other indices, the primitive information of determining weight should come from the objective environment, and decide the weight coefficient of the relative indices the number of information value provided by the indices.

A Entropy

The founder of the information theory, Shannon introduced the idea of entropy into the information theory, as a measurement of the system uncertainty.

When the system in several different states, let the probability of each state is $p_i (1, 2, \dots, n)$, and the entropy is:

$$H = - \sum_{i=1}^n p_i \cdot \ln p_i \tag{1}$$

in the formula : $0 \leq p_i \leq 1, \text{ and } \sum_{i=1}^n p_i = 1$

When all p_i are equal (equal probability), H achieve maximum value, it is $H_{\max} = \ln n$. This character is called the extremum of entropy.

Let A is not fuzzy judgment matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \tag{2}$$

Let $S_k = \sum_{j=1}^n a_{kj} (k=1, 2, \dots, n)$ is the sum of the elements in line k ,

$f_{kj} = a_{kj} / S_k$ can approximate replace the appearance probability of the element j in the result k . So the entropy of n result is:

$$H_n = - \sum_{j=1}^n f_{nj} \ln f_{nj} \tag{3}$$

And according to the value of H_n , we can judge the optimal results.

B Entropy-weight

Let evaluation questions which have m objects need to be evaluated and n evaluation indices, use the method of experts evaluating and according to the principle of qualitative combined with quantitative to get the judgment matrix of evaluation indices:

$$R' = (\gamma'_{ij})_{m \times n} = \begin{pmatrix} \gamma'_{11} & \gamma'_{12} & \cdots & \gamma'_{1n} \\ \gamma'_{21} & \gamma'_{22} & \cdots & \gamma'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma'_{m1} & \gamma'_{m2} & \cdots & \gamma'_{mn} \end{pmatrix} \tag{4}$$

Normalize R' , normalize formula is:

$$\gamma_{ij} = \frac{\gamma'_{ij} \min \{ \gamma'_{ij} \}}{\max \{ \gamma'_{ij} \} - \min \{ \gamma'_{ij} \}} \tag{5}$$

After normalization, we can get formula.

$$R = (\gamma_{ij})_{\max} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mn} \end{pmatrix} \tag{6}$$

in the formula : $\gamma_{ij} \in [0,1]$

In the evaluation question (m, n) , we can let the entropy of the j evaluation index is:

$$H_j = -k \sum_{i=1}^m f_j \cdot \ln f_j ; j = 1, 2, \dots, n$$

in the formula : $f_j = \frac{\gamma_{ij}}{\sum_{i=1}^m \gamma_{ij}} ; k = \ln m ;$, and let : when $f_j = 0, f_j \cdot \ln f_j = 0,$

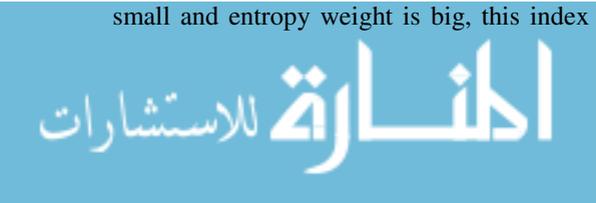
so let the entropy-weight d_j of the j index is:

$$d_j = \frac{1 - H_j}{n - \sum_{j=1}^n H_j} \tag{7}$$

The above formula can draw the following conclusions:

When the index j values of every evaluated object are same, the entropy value is 1, entropy-weight is 0. That means the index didn't provide any useful information to the decision maker, and this index could be canceled.

When the index j values of every evaluated object are different, entropy value is small and entropy weight is big, this index is useful for the decision maker. It also



means that objects have obvious difference on this index, and should be stressful invested.

③ The bigger the entropy of the index is, the smaller the entropy weight of it is , so

the index would be unimportant and fulfill $0 \leq d_j \leq 1$ 和 $\sum_{i=1}^m d_j = 1$

④ As the weight number, the entropy weight has its special meaning. It is not an important factor for an index of the valuation questions in practice, but a factor which represent the comparative competition level of each index when every index value is ensured after given the evaluated objects class.

C Entropy Weight Multi-objection evaluations

According to formula (3- 7) to calculate the entropy-weight d_j of each evaluate index. And ensure the entropy value ω_j of corresponding index by AHP, and calculate the comprehensive entropy-weight by the following formula:

$$\lambda_j = \frac{\omega_j d_j}{\sum_{j=1}^n \omega_j d_j} \tag{8}$$

Then evaluate by entropy-weight.

Mapped the evaluated objects into the distance space, and get

$$L_y(\lambda, i) = \left| \sum_{j=1}^n \lambda_j^p (1 - \gamma_{ij})^p \right|^{\frac{1}{p}}$$

In the evaluation, if only attach importance to the sum of deviation, let $p = 1$, and get:

$$L_1(\lambda, i) = \left| 1 - \sum_{j=1}^n \lambda_j \cdot \gamma_{ij} \right| \tag{9}$$

So we would know that the small deviation is more close to the ideal solution, and we can sort the evaluation of every evaluated object from small to big according to $L_p(\lambda, i)$.

5 Example Analysis

This article choose three bonded logistic enterprises L1、 L2、 L3 of Dayaowan in Dalian, and evaluate their customer service performance. Under the principle of scientific, representative, and operability, we made the practical investigation. And combine the comprehensive evaluation of the advisory group which was formed by the experts in the field of shipping, ports, customs and international trade. Then normalize and dispose the data. So we got the comprehensive scores of each index of the there enterprises, and according to formula (1), we got the data shown in table 2.

Table 2. The estimate of three bonded logistic enterprises

| Index Enterprise | L1 | L2 | L3 |
|----------------------------|------|------|------|
| circulating and processing | 0.79 | 0.86 | 0.76 |
| cargo agent | 0.86 | 0.86 | 0.79 |
| warehouse receipt pledge | 0.87 | 0.89 | 0.81 |
| Transportation and storage | 0.86 | 0.95 | 0.75 |
| financial services | 0.83 | 0.91 | 0.75 |

Thus, the judge matrix of three bonded logistics enterprises L1、L2、L3 was build.

According to the estimate table of three bonded logistics enterprises L1、L2、L3, and formula (5) and (6), we can be get the regular matrix R after standardized:

$$R = \begin{bmatrix} 0.20 & 0.55 & 0.05 \\ 0.55 & 0.55 & 0.20 \\ 0.60 & 0.70 & 0.30 \\ 0.55 & 1 & 0 \end{bmatrix}$$

And according to formula (7) and (8), we got the entropy weight analysis table of three bonded logistics enterprises L1、L2、L3, shown in table 3.

Table 3. Entropy weight analysis table of three bonded logistics enterprises L1、L2、L3

| | L1 | L2 | L3 |
|-------------|------|------|------|
| h_j | 2.48 | 2.53 | 2.01 |
| d_j | 0.37 | 0.38 | 0.25 |
| ω_j | 0.83 | 0.91 | 0.75 |
| λ_j | 0.36 | 0.41 | 0.23 |

do the final evaluation according to formula (9):

$$L_1(\lambda,1) = |(1 - 0.678)| = 0.312 \quad L_2(\lambda,2) = |(1 - 1.154)| = 0.154$$

$$L_3(\lambda,3) = |(1 - 0.158)| = 0.842$$

The calculated results show that:

$$L_2(\lambda,2) > L_1(\lambda,1) > L_3(\lambda,3)$$

So, according to the sorting result of entropy weight multi-objection evaluation method: the customer service performance of bonded logistic enterprise L2 is better than L1, and L1's is better than L3's.



6 Summary

This article use the entropy weight multi-objection evaluation method to evaluate the bonded logistic enterprises customer service performance, and got the cogent result. And also point out that the bonded logistic enterprises customer service performance evaluation is the key point of bonded logistic enterprises performance evaluation, it plays an important role in providing the efficiency of regional bonded logistic and the profit of regional economy.

Considering that in the program evaluation with multi-objection, because the rationality of weight affect the veracity of multi-objection evaluation sorting result, so the research on the weight problem is very important. Base on the practical customer service performance of real bonded logistic enterprises, this article bring in entropy weight multi-objection evaluation model which is base on the entropy idea of information theory, this model could ensure a high reliable optimal alternative under the situation of only judge matrix but no expert weight. Of course, this model could be adjusted the factors like evaluation index, weight, and used for the comprehensive evaluation of other areas, so be a reference for decision making.

References

1. Sun, H., Qi, S.: Modern logistics activity performance analysis. China logistics publishing house, Baijing (2001)
2. Wang, C.: Modern logistics management practice and cases. enterprise Management publishing house, Baijing (2001)
3. Wang, Y.: Logistics service performance standards and evaluation. Logistics Technology (1), 25–26 (2002)
4. Zhou, T., Cheng, J., Qiao, Z.: Performance index system and vague evaluation of logistics enterprises. Modernization of Management (9), 26–28 (2002)

Necessity Analysis of Inside Shoulder Setting in Multi-lane Freeway with Traffic Simulation Method

Zhong Liande¹, Li Xiuwen², and Wu Keman¹

¹ Research Institute of Highway, Ministry of Transport, Beijing 100088, China

² China Highway Engineering Consulting Group Company LTD, Beijing 100098, China
zsy_516@sohu.com

Abstract. “Technical Standard of Highway Engineering” regulated that eight-lane freeway should be provided 2.5-meter width inside shoulder. While several just completed widening to eight-lane freeways in China were not set inside hard shoulder, nor even wider than 0.75-meter curbs. The impacts of the behavior that a vehicle moves from the inside lane to the right shoulder on the operation speed were exploited under different volume and lane-change distances using traffic simulation method. The results indicated the significance and necessity of setting the inside shoulder of multi-lane freeway in terms of safety and efficiency once more.

Keywords: Multi-lane freeway, Inside shoulder, Weaving section, Speed, Traffic simulation.

1 Introduction

Most of the freeways on operation in China are two-way four-lane or six-lane form, in which the inside shoulder setting was not considered except the subgrade separated freeway. While with the increase in the number of freeway lanes, setting the inside shoulder is inevitable, and the inside shoulder has different role from the right shoulder, and the most important role of it is to provide the side clearance and fault-tolerant space for the high-speed vehicles in inside lane, as well as provide temporary parking place in cases of emergency, so the vehicles need not cross multiple lanes to park on the right shoulder. The existence of inside shoulder is not only good for the improvement of the overall traffic operating efficiency, but also avoid severe traffic accidents [1].

Although some researchers and officials have been considering the question of inside shoulder and “Technical Standard of Highway Engineering” also has the regulation that eight-lane freeway should be provided 2.5-meter width inside shoulder [2]. Unfortunately, several just completed widening to eight-lane freeways in China, such as Shanghai-Nanjing Freeway, and Shenyang-Dalian Freeway, and Guangzhou-Foshan Freeway etc were not set inside hard shoulder, nor even wider than 0.75-meter curbs.

The regulation or set criteria about hard shoulder can be found in the road design standards and specifications of other countries. Compared with the United States and Japan, the value of right shoulder width of our country is in the middle. In general the

shoulder width is 2.5 meters with the 1.75-meter minimum for freeway in Japan. In American standards the shoulder width is affected by many factors. Shoulders should be set in both sides, and the right shoulder is at least 10 feet, and when trucks are more than 250 vehicles per hour, then the right shoulder is best to reach 12 feet. In the four-lane freeway, usually the inside hard shoulder width is between 4 feet and 8 feet. The width of inside shoulder for six or more lanes freeways should be at least 10 feet, when trucks are more than 250 vehicles per hour, the shoulder width should be 12 feet [3]. The inside shoulder setting of the highway Specification in our country for renovation and expansion into the eight-lane freeway and newly building eight-lane freeway is non-mandatory provision, and the standards are mainly referred from some foreign practices [3, 4, 5], and the main reason is that there is no related researches to provide theoretical support and also lack of knowledge on the functions of inside shoulder. At present more considerations to be made about the freeway design issues such land conservation and cost reduction when freeway renovation and expansion, which lead to a great bad impact for the operation of freeway in the future.



Fig. 1. Freeway cross section comparison between the USA and China

In this study, the impacts of the behavior that a vehicle moves from the inside lane to the right shoulder on the straight traffics of 8-lane and 10-lane freeway were analyzed, and the operation speed in weaving section and downstream were exploited under different volume and lane-change distances using traffic simulation method. The findings proved the significance and necessity of setting the inside shoulder of multi-lane freeway from the aspects of safety and efficiency.

2 Trial Objective and Project Design

To illustrate necessary of setting the inside shoulder, the road environment was built as shown in Fig.2. The configuration is a 4-lane weaving section which could be equal to the 10-lane freeway without inside shoulder in reality. To reach the right exit ramp the vehicle enter from the left ramp and need traverse 4 straight lanes. This process is the same as that a vehicle moves from the innermost lane to the right shoulder in 10-lane freeway. the purpose that fixing the lane-change distances in the simulation analysis is to analyze the impact of different distances on speed.

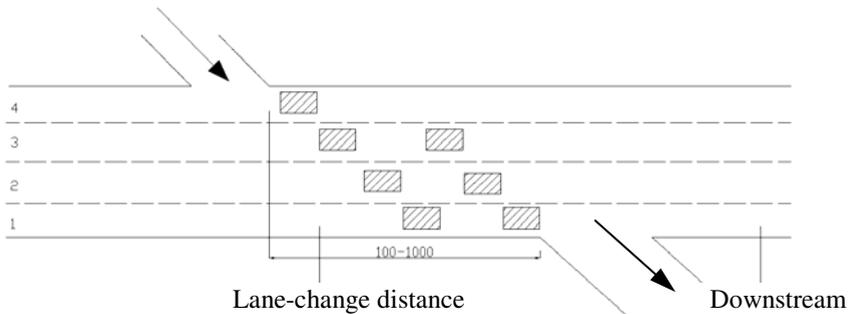


Fig. 2. Highway configuration for simulation

In this study VISSIM was selected as the traffic simulation platform, before the further analysis the simulation system was calibrated using the actual data, so the accuracy could meet the requirements. Simulation inputs primarily consist of three areas:

Lane numbers in weaving section: 3 lanes and 4 lanes represent 8-lane and 10-lane freeway without inside shoulder. (2) Input flow: 4 types: 500, 1000, 1500, 2000 vehicles per hour per lane. For better comparison all the vehicles are assumed as passenger cars in the system. (3) Lane-change distances (the length of weaving section): from 100 meters to 1000 meters with 100 meters intervals, so a total of 10 cases. To summarize, there are total of $2 \times 4 \times 10 = 80$ different cases in the simulation.

Ideal simulation condition is that the process only one vehicle enter from entrance ramp to the basic section then change lanes and exit from the right ramp was studied, which is a little difficult considering the sample and the randomness of simulation. So a substitute process that the vehicles enter the weaving section from the entrance ramp at a 10-second interval in 1 simulation hour was studied. The virtual detectors were placed in the middle of and 50-meter down the weaving section for record the operation speeds. Table 1 is a fragment of analysis data.

Table 1. Fragment of Data Set of Speeds in Weaving Section

| Cases | Lane-change distance (m) | Volume (pcu/h/lane) | No. of lanes | Speed in weaving section (km/h) | Speed downstream (km/h) |
|-------|--------------------------|---------------------|--------------|---------------------------------|-------------------------|
| 1 | 100 | 2000 | 4 | 9 | 70 |
| 2 | 100 | 1500 | 4 | 9 | 70 |
| 3 | 100 | 1000 | 4 | 9 | 69 |
| 4 | 100 | 500 | 4 | 88 | 99 |
| 5 | 100 | 2000 | 3 | 33 | 80 |
| 6 | 100 | 1500 | 3 | 12 | 68.6 |
| 7 | 100 | 1000 | 3 | 15 | 104 |
| 8 | 200 | 2000 | 4 | 8 | 58 |
| 9 | 200 | 1500 | 4 | 9 | 58 |
| 10 | 200 | 1000 | 4 | 12 | 63 |
| ... | ... | ... | ... | ... | ... |

3 Analysis and Discussion

Fig.4 is a screenshot of the process of traffic simulation. The process vehicles change from the inside lane to the right shoulder under heavy volumes and short distance of weaving would cause lots of straight vehicles queuing and stops, and the more lanes, the more evident of this phenomenon. Frequent brakes, and lane changes make an increase in traffic conflicts, which prone to accidents such as rear-end and side-wipe crashes according to Traffic Conflict Theory, so the mandatory lane changing is potential dangerous significantly. At the same time, affect the capacity and level of service of the freeway [6].

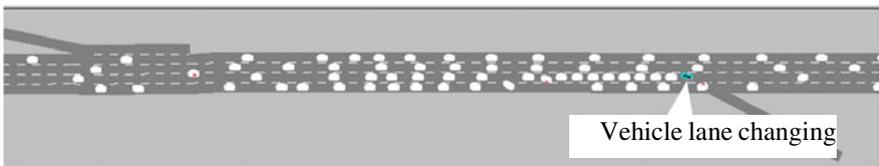
**Fig. 3.** Virtual operation in simulation

Fig.4 provides a more explicit explanation of the result that the speeds of different lane-change distances of 4-lane freeway when the volume input is 1000 pcu/h/lane. The speed in the weaving section and at downstream area are both subject to the lane-change distances, especially the speed in the weaving. the average operation speed is only 9.4 km/h in the weaving section when the distance is 100m, and 11.7 km/h for the 200-meter distance. With the increase of the lane-change distance, the speed in

weaving section also gradually increase, and when the distance is more than 500 m, the speed in the weaving section maintains at the range between 70 km/h and 80 km/h. Relatively the speed downstream is less affected by the lane-change distance. With the distance increasing, there is also a growing trend of the speed. The minimum of speed downstream is 63 km/h, appears in the lane-change distance of 200 meter, and the maximum is 97 km/h under the maximum lane-change distance, namely, 1000 meters.

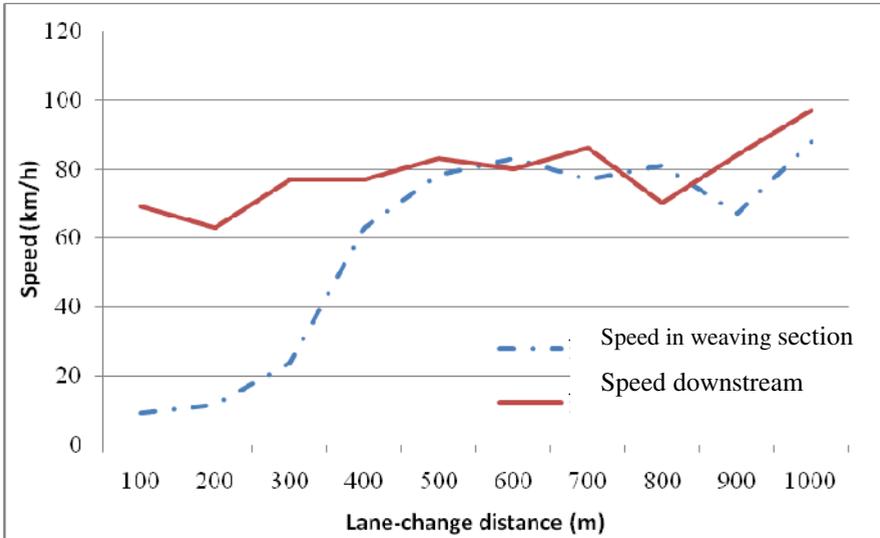


Fig. 4. Speed comparison of different lane-change distances under the condition of 4-lane and 1000 pcu/h/lane

The influences of volume input and number of lanes can be seen in Fig.5.

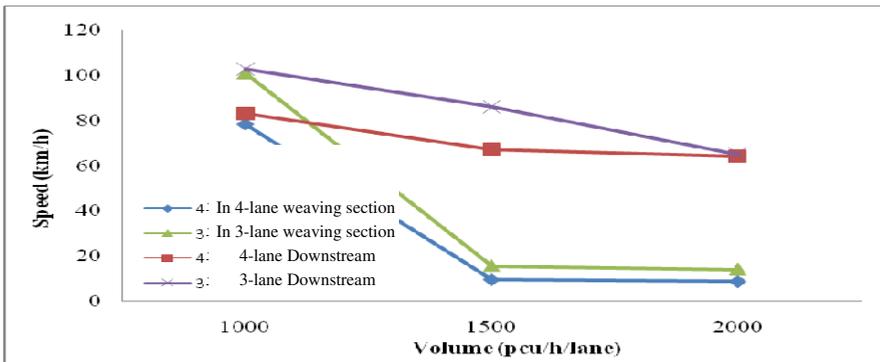


Fig. 5. Speed comparison between in and down of weaving section for 4-lane and 3-lane freeway under 500-meter lane-change distance



The results of simulation indicates that before the volume input less than 500 pcu/h/lane the speeds in the weaving section and downstream are almost the same and unaffected by the number of lanes. When the flow rate of each lane is 1000 pcu/h, the speed difference between in the weaving section and downstream is not particularly large, and the speed downstream is about 20 km/h higher than that in the weaving section. As the flow increased to 1500 pcu/h/lane, until the 2000 pcu/h/lane, the speed in weaving section decreased sharply and the value is as low as 20 km/h. While the speed downstream reduce step by step. When the flow rate is 2000 pcu/h/lane, the speed downstream is 64 km/h for 4-lane and 65 km/h for 3-lane weaving section. It also can be seen from Fig.5 due to one less lane-changing behavior the speed downstream of 3-lane weaving section is bigger than that of 4-lane weaving section. The maximum speed difference appears in the volume input of 1500 pcu/h/lane, and the speed downstream is 86 km/h for 3-lane weaving section as well as 67 km/h for 4-lane weaving section.

4 Conclusions

In this study the impacts of the behavior that a vehicle moves from the inside lane to the right shoulder on the operation speed were exploited under different volume and lane-change distances using traffic simulation method. And the speed of weaving section and downstream were compared under the same volume input and different lane-change distances, as well as the same lane-change distances and different volume inputs and number of lanes. The results indicated the significance and necessity of setting the inside shoulder of multi-lane freeway from the aspects of safety and efficiency

Acknowledgements. This research was supported in part by grants from the National Science Foundation of China (Grant No. 50908106) and National Science and Technology Support Program (Grant No. 2009BAG13A03). The contents of this paper reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the sponsors.

References

1. Zhong, L.D., et al.: Necessity analysis of inside shoulder of Multi-lane expressway. Highway (2), 106–110 (2011)
2. China Ministry of Transport. Technical Standard of Highway Engineering. JTG B01-2003
3. A Policy on Geometric Design of Highways and Streets, American Association of State Highway and Transportation Officials, Washington, D.C (2001)
4. Leisch, J. P.: Freeway and Interchange Geometric Design Handbook, Institute of Transportation Engineers, ITE (2006)
5. Pline, J. L.: Traffic Engineering Handbook, 5th edn, Institute of Transportation Engineers, ITE (1999)
6. Transportation Research Board. Highway Capacity Manual. TRB, National Research Council, Washington DC (2000)

Agricultural Products Quality Cost-Allocating Based on Supply Chain Management

Wang Xin-li and Wei Qiong

College of Economics & Management
Heilongjiang Bayi Agricultural University, Daqing, China
WangXishan@163.com

Abstract. In order to reasonable share the cost of quality security and quality inspection to set up supply chain quality cost-allocating models under the symmetrical information condition, taking quality prevention level and the quality evaluation level as the decision variable. On these bases, we considered a three asymmetric information moral hazard by the principal-gent theory, and then set up supply chain quality cost-return models between the agricultural with manufacturer and processor. The goals were to achieve total optimum profits of supply chain. At last, we solved problems of quality cost-allocating under two-stage supply chain by determining the correlation and quality cost-allocating rates.

Keywords: Agricultural products supply chain, quality cost-allocating, principal-gent relationship, moral hazard.

1 Introduction

Agricultural products supply chain means a moving network that is make up of agricultural products provider, producer, seller and consumer. The basic purpose of management lies in meeting the requirements of consumers and achieving the total optimum profits of supply chain. With the rapid development of community economic and improvement of living standards, the supply of agricultural products took a historic shift from chronic short ages to an overall balance with surpluses during good years. Social demand for agricultural products have become increasing diversified, and people's attention for the quality level of agricultural products have keep rising. However, additional quality cost in the form of human, material resources, information, technology usually occurs in improving quality of products, which covers the advantage of quality improvement. The cost will ascend with the quality improvement after having reached a certain quality level, which have been known as dependency relations between quality level and quality cost. Information is indispensable to supply chain operation. It is important that a perfect information sharing can achieve total optimum profits of supply chain, but we are striving after an ideal impossible of realization. The key to increase the profit of enterprises based on supply chain management under asymmetric information by trying to get rid of unvalued-added quality work and sharing quality cost reasonably.

The earliest related research was contributed by Mazé etc. (2001), they put forward reasonable use the game theory so that comprehensively ascension agricultural product quality firstly, according to the agricultural products supply chain characteristics in European. [1] To validate food chain in the process of vertical integration existence of certain moral hazard problem by using game theory was proposed by Vetter etc. (2001). [2] A plan that the “one to one” of supply chain according to the situation on the quality upgrade to improve the quality of quality costs for both sides undertake have studied up(Rachel Q etc. 2007). [3] Zhang Cui hue etc. (2003) studied the decision of supply chain quality cost effects of moral hazard. The results indicated that, quality prevention cost decreased with decreased quality prevention level, but quality evaluation cost increasing. [4] Zhou Ming etc. (2004) studied the influence of moral hazard between quality prevention decision and evaluation decision. The results supplied equilibrium with hiding for achieving the optimum profits of supply chain. [5] Cao Jian etc. (2006) researched optimal contract design of supply chain, when quality loss cost was shared by both sides. There had internal and external damage coefficient to achieve the optimum profits of supply chain. [6]

In a word, Most recent research of the efficiency problem of quality chain management by using the combinatory game theory focus on increasing products quality level and improving quality cost from macroscopic aspect. Relevant researchs was unconcerned with sharing quality cost reasonably between the agricultural with manufacturer and processor. However, sharing quality cost reasonably depends primarily on whether or not achieving total optimum profits of supply chain. Taking the dependency relations between quality level and quality cost into account, it is hence necessary to achieve total optimum profits of supply chain study on supply chain quality cost-allocating reasonably by the principal-agent theory from quality cost perspective.

2 Problem Descriptions

In the two-stage supply chain environment, the processors usually make the necessary evaluation when the manufacturer sale agricultural producers to the processors. They make decisions on returning merchandise or reducing the price with specific conditions if it is found the quality problems. Therefore, it has occurred quality cost between manufacturer and processors. Including: quality assurance cost, quality risk cost and quality invested cost.

Research Thought

Based on determining quality prevention level and the quality evaluation level under the symmetric information, we study cost-allocating rates of quality assurance cost, quality risk cost and quality invested cost under the asymmetric information by the principal-gent theory. The goal of study lies in achieving total optimum profits of supply chain at the same time meeting requirements of consumers.

Parameter Setting

(1) Quality Assurance Cost, including prevention cost $C(p_m)$ and appraisal cost $C(p_b)$.

Prevention cost can be defined as a cost that the manufacturer pays for guaranteeing quality prevention level (P_m) of agriculture products to prevent quality error when the manufacturer supplies primary agriculture products to processor. Appraisal cost can be defined as a cost that the processor guarantees quality evaluation level (P_b).

(2) Quality risk cost, including internal loss cost (C_I) and external loss cost (C_E).

Internal loss cost can be defined as a cost that manufacturer need to pay for improve quality cost (R) or make a discount to processor (price cut to say) ($\Delta\Pi$), when the manufacturer sell products and processor evaluate the quality error.

External loss cost can be defined as a cost that is shared by agricultural manufacturer and processor, manufacturers the loss of sharing is λE , processor the loss of sharing is $(1 - \lambda)E, \lambda \in (0, 1)$ when the manufacturer sell products and processor don't evaluate the quality error which is found by customers.

(3) Quality Invested Cost.

Quality Invested Cost can be defined as a cost that is spent on improving the quality prevention level and evaluation level of agriculture products. In a two-stage supply chain , firstly, the initial prevention level and evaluation level of manufacturer and processor is (\bar{P}_m, \bar{P}_b) secondly , deciding the prevention level and evaluation level after the processor investing is (\hat{P}_m, \hat{P}_b), thirdly, deciding the prevention level and evaluation level after the follower(manufacturer) is (P_m, P_b), which is the best level.

K1 and K2 indicate quality invested cost rate of the processor and manufacturer to improve quality prevention level;

K3 and K4 indicate quality invested cost rate of the processor and manufacturer to improve the evaluation level. (through negotiation by contract of both parties, as known number in the study).

To realize the role of quality invest to improve the prevention level and evaluation

level, invested cost of the processor is represented as $K_1 \ln \frac{1 - \bar{P}_m}{1 - \hat{P}_m} + K_3 \ln \frac{1 - \bar{P}_b}{1 - \hat{P}_b}$,

which indicates the invested cost increases form (\bar{P}_m, \bar{P}_b) to \hat{P}_m, \hat{P}_b , invested cost of

the manufacturer is represented as $K_2 \ln \frac{1 - \hat{P}_m}{1 - P_m} + K_4 \ln \frac{1 - \hat{P}_b}{1 - P_b}$, which indicates the

invested cost increases form(\hat{P}_m, \hat{P}_b) to (P_m, P_B).

(4) Unit Product Income

Given: Π_m indicates the unit income of manufacturers' basic agriculture products; Π_b indicates the unit income of processors' agriculture products.

Assumption of the Research

To convenient for the study, here gives basic prerequisites for the model:

(1) The ultimate agriculture products with insufficient safety level are all caused by the low safety level of the primary agriculture products.

(2) C (Pm) and C (Pb) both are increasing function.

(3) The external loss of the unit agriculture product is greater than the internal loss,

$$C_E > C_I \text{ [7]}$$

(4) When the two parties sign the purchase contract, the agricultural supply chain is information symmetry, but after they sign the contract, information asymmetry, they both can't know the other side's choice, conform to the moral risk problem of hidden action in information economics. [8]

Model Description

According to parameter and description of prerequisite, give the quality cost-allocating models:

quality cost-allocating function of the whole agricultural supply chain:

$$Z = \Pi_b - C(p_m) - C(p_b) - (1 - P_m)P_b(C_I + R) - (1 - P_m)(1 - P_b)C_E - K_1 \ln \frac{1 - \bar{P}_m}{1 - \hat{P}_m} - K_3 \ln \frac{1 - \bar{P}_b}{1 - \hat{P}_b} - K_2 \ln \frac{1 - \hat{P}_m}{1 - P_m} - K_4 \ln \frac{1 - \hat{P}_b}{1 - P_b} \tag{1}$$

$(1 - P_m)P_b(C_I + R)$ Is used indicating the cost of internal loss about unit products.

$(1 - P_m)(1 - P_b)C_E$ Is used indicating the cost of external loss about unit products.

Unit quality cost-allocating function of the processors:

$$Z_b = \Pi_b - \Pi_m - C(p_b) - (1 - P_m)P_b(C_I - \Delta\Pi) - (1 - P_m)(1 - P_b)(1 - \lambda)C_E - K_1 \ln \frac{1 - \bar{P}_m}{1 - \hat{P}_m} - K_3 \ln \frac{1 - \bar{P}_b}{1 - \hat{P}_b} \tag{2}$$

Unit quality cost-allocating function of the agricultural manufacturer:

$$Z_m = \Pi_m - C(p_m) - (1 - P_m)P_b(\Delta\Pi + R) - (1 - P_m)(1 - P_b)\lambda C_E - K_2 \ln \frac{1 - \hat{P}_m}{1 - P_m} - K_4 \ln \frac{1 - \hat{P}_b}{1 - P_b} \tag{3}$$

3 Decision-Making Optimal Solution of Quality Cost under Information Symmetry

In two-stage Supply Chain, when products quality prevention level and evaluation level both can be detected is the situation of information symmetry, the optimal solution in that situation is the basis to study quality cost-allocating under information asymmetry. The explanation is as follows:

When processors and manufacturers adjust the evaluation level, only think of improve the evaluation level, K1 and K2 has no difference, in other words, the invested cost always rises Pb proportionately whichever side. Of course, we should choose



minimum invested cost to improve evaluation level, set $K_1 = \min(K_1, K_2)$, and $K_3 = \min(K_3, K_4)$.

So, the first order condition of formula (1) is:

$$C'(p_m) = C_E - P_b(C_E - C_I - R) - \frac{K_1}{1 - P_m} \tag{4}$$

$$C'(p_b) = (1 - P_m)(C_E - C_I - R) - \frac{K_3}{1 - P_b} \tag{5}$$

Then, optimal solution under information symmetry (P_m^*, P_b^*):

$$\left\{ P_m^* = 1 - \frac{\min(K_1, K_2)}{E + P_b^*(C_I + R - C_E) - C'(p_m^*)} \right. \tag{6}$$

$$\left. P_b^* = 1 - \frac{\min(K_3, K_4)}{(1 - P_m^*)(C_E - C_I - R) - C'(p_b^*)} \right. \tag{7}$$

4 The Quality Cost-Allocating Model under Asymmetry Information

The principal-agent model is a process of incomplete information games, divide into three stages: the first, the client(who has more information in Principal-Agent Relations) design contract; the second, the agent choose whether to accept the contract; the third, according to rules of the contract, the accepted agent choose reasonable actions. The problem that the model has to solve is how to design a contract which can motivate the agent choose the most advantageous action for the client. The design has two constraints: one is participation constraint (IR), it make that the agent has motivation to participate when accept agents' desired effect is not smaller than cannot accept agents' maximal desired effect (μ); the other is incentive-compatibility constraints(IC), it make that the agent choose actions of the client hoping actively, however, it is that the expected utility that the agent choose actions that the client hoping is greater than choose others.

Supply chain management in reality is always information asymmetry, for the benefit of the corporation, the agent has a possibility to hide its action, or we can say the agent take an act of cheating, that will lead to the moral risk problem. The moral risk problem may be distributed into three classes as follows:

)A Game Analysis on the situation of hiding Pm and detection Pb.

This situation gives Pb, agricultural manufacturer will judge Pm by maximum expected utility of them. Now the manufacturer is the agent who exist moral risk problem, so the processor designs the conditions of mechanism as follows:

$$Max_{(P_m, \Delta II, \lambda, P_b)} Z_b \tag{8}$$

$$s.t. (IR) Z_m \geq \mu \tag{9}$$

$$(IC) \quad P_m = \arg \text{Max}_{P_m} Z_m \tag{10}$$

Suitable for both (8), (9) will get the optimum quality cost-income of the agricultural supply chain:

$$\text{Max}_{\Pi, \Delta \Pi, \lambda, P_b} Z = Z_m + Z_b \tag{11}$$

Then solve one step partial derivative of Pm and make the solution be zero in the incentive-compatibility constraints (10):

$$s.t.(IC) \quad C'(p_m) = P_b(\Delta \Pi + R) + (1 - P_b)\lambda C_E - \frac{K_2}{1 - P_m} \tag{12}$$

In considering the situation above about the processor’s design mechanism, the optimal decision that the manufacturer accepts it will be: if $P_b = P^*_b$, the processor will pay a fixed Π_m to the manufacturer, the manufacturer will choose \hat{P}_m according to (12) to make $\hat{P}_m = P^*_m$. And now:

$$\hat{P}_m = 1 - \frac{K_2}{P^*_b(\Delta \Pi + R) + (1 - P^*_b)\lambda C_E - C'(p^*_m)}$$

When internal loss occurs, the manufacturer will pay a price discount:

$$\Delta \Pi = \begin{cases} C_I + \frac{(1 - P^*_b)C_E}{P^*_b} & ; (K_2 < K_1, \lambda = 0) \\ \frac{K_2}{K_1} \left[C_I + R + \frac{C_E(1 - P^*_b)}{P^*_b} \right] - \frac{C'(p^*_m)(K_2 - K_1)}{P^*_b \cdot K_1} - R & ; (K_2 > K_1, \lambda = 0) \end{cases}$$

When external loss occurs, the manufacturer’s quality risk cost-allocating rate:

$$\lambda = \begin{cases} 1 + \frac{C_I \cdot P^*_b}{(1 - P^*_b)C_E} > 1, \quad (\text{no reasonable key}); & (K_2 < K_1, \Delta \Pi = 0) \\ \frac{(K_2 - K_1)[P^*_b \cdot R - C'(p^*_m)] + K_2[P^*_b \cdot C_I + (1 - P^*_b)C_E]}{K_1(1 - P^*_b)C_E}; & (K_2 > K_1, \Delta \Pi = 0) \end{cases}$$

(2) A Game Analysis on the situation of hiding P_b and detection P_m .

This situation gives Pm, while agricultural manufacturer judges Pb by maximum expected utility of them. Now the processors are the agent who exist moral risk problem, so the manufacturer designs the condition of mechanism as follows:

$$\text{Max}_{(\Pi_m, \Delta \Pi, \lambda, P_m)} Z_m \tag{13}$$

$$s.t.(IR) \quad Z_b \geq 0 \tag{14}$$

$$(IC) \quad P_b = \arg \text{Max}_{P_b} Z_b \tag{15}$$

Suitable for both (13), (14) will get the optimum quality cost-income of the agricultural supply chain:

$$\text{Max}_{\Pi, \Delta \Pi, \lambda, P_m} Z = Z_m + Z_b \tag{16}$$

Then solve one step partial derivative of Pm and make the solution be zero in the incentive-compatibility constraints (15):

$$s.t.(IC) \quad C'(p_b) = (1 - P_m)[(1 - \lambda)C_E + \Delta \Pi - C_I] - \frac{K_3}{1 - P_b} \tag{17}$$

In considering the situation above about the manufacturer’s design mechanism, the optimal decision that the processor accepts be: if $P_m = P_m^*$, the processor will pay a fixed Π_m to the manufacturer, the manufacturer will choose \hat{P}_b according to (17) to make $\hat{P}_b = P_b^*$. And now:

$$\hat{P}_b = 1 - \frac{K_3}{(1 - P_m^*)[(1 - \lambda)C_E + \Delta \Pi - C_I] - C'(p_b^*)}$$

There has no problem of price discount $\Delta \Pi = 0$. Agricultural manufacturer won’t undertake loss, in other words, the improvement cost won’t produce R=0, and external loss-allocating is

$$\lambda = \frac{R}{C_E} = 0.$$

(3) A Game Analysis on the situation of detection Pb and Pm.

In this situation, the agricultural supply chain is a virtual client. Processor and manufacturer are agents with exist a two-way moral risk problem. While the whole supply chain will design the condition of mechanism as follows:

$$\text{Max}_{(\Pi_m, \Delta \Pi, \lambda)} Z \tag{18}$$

$$s.t.(IR) \quad Z_b \geq 0 \quad \text{and} \quad Z_m \geq 0 \tag{19}$$

$$(IC) \quad P_b = \arg \text{Max}_{P_b} Z_b \quad \text{and} \quad P_m = \arg \text{Max}_{P_m} Z_m \tag{20}$$

The first-order condition of quality cost-income in agricultural supply chain is:

$$C'(p_m) = P_b(\Delta \Pi + R) + (1 - P_b)\lambda C_E - \frac{K_2}{1 - P_m} \tag{12}$$



$$C'(p_b) = (1 - P_m) [(1 - \lambda)C_E + \Delta\Pi - C_I] - \frac{K_3}{1 - P_b} \tag{17}$$

The manufacturer chooses \hat{P}_m according to (12), if other side's action is given. The processor chooses \hat{P}_b according to (17), while make $\hat{P}_m = P_m^*$, $\hat{P}_b = P_b^*$, the manufacturer should pay a price discount and the quality risk cost-allocating rate:

$$\textcircled{1} K_2 < K_1, K_3 < K_4$$

$$\Delta\Pi = (1 - P_b^*)(C_E - C_I - R) + C_I$$

$$\lambda = \frac{C_E - P_b^*(C_E - R - C_I)}{C_E}$$

$$\textcircled{2} K_2 < K_1, K_3 > K_4$$

$$\Delta\Pi = \frac{(K_4 - K_3)C'(p_b^*)(1 - P_b^*)}{(1 - P_m^*)K_4} + \frac{K_3}{K_4}(C_E - C_I - R)(1 - P_b^*) + C_I$$

$$\lambda = 1 - \frac{P_b^*}{C_E \cdot K_4} \left[K_3(C_E - C_I - R) - \frac{(K_3 - K_4)C'(p_b^*)}{1 - P_m^*} \right]$$

$$\textcircled{3} K_2 > K_1, K_3 < K_4$$

$$\Delta\Pi = \frac{K_2 \cdot (1 - P_b^*)C_E + K_2 \cdot P_b^* \cdot (C_I + R) + (K_2 - K_1)C'(p_m^*)}{K_1} - R$$

$$\lambda = \frac{K_2 \cdot (1 - P_b^*)C_E + K_2 \cdot P_b^* \cdot (C_I + R) + (K_2 - K_1)C'(p_m^*)}{C_E \cdot K_1}$$

$$\textcircled{4} K_2 > K_1, K_3 > K_4$$

$$\Delta\Pi = \frac{K_1C_I + (K_1 - K_2)C'(p_m^*)}{K_1} - \frac{(K_3 - K_4)C'(p_b^*)(1 - P_b^*)}{(1 - P_m^*)K_4} + \frac{(C_E - C_I - R)(1 - P_b^*)K_3}{K_4}$$

$$\lambda = \frac{P_b^* [(K_3 - K_4)C'(p_b^*) - (C_E - C_I - R)K_3(1 - P_m^*)]}{(1 - P_m^*)K_4 \cdot E} + \frac{K_2}{K_1 \cdot (1 - P_b^*)} - \frac{P_b^*}{1 - P_b^*} + \frac{(K_1 - K_2) \cdot [C'(p_m^*) \cdot (1 - K_1) - P_b^* \cdot (C_I + R - C_E)]}{C_E \cdot K_1 \cdot (1 - P_b^*)}$$

The results indicates that, it's no need to share quality assurance cost and quality invested cost in the running of the two-stage agricultural supply chain, while these costs are taken by the corporation of themselves. The sharing situation of quality risk cost: the sharing scales of quality risk cost both sides with vary quality invested cost rate

(K_1, K_2, K_3, K_4) , When internal loss occurs, processor needn't to take the risk cost which is taken by manufacturer. Then we figure out the price discount $\Delta\Pi$ t of manufacturer supplying. When external loss occurs, both sides will share the external loss cost together. Then we figure out the risk cost-allocating rate λ of the manufacturer and the risk cost-allocating rate $1-\lambda$ of processor. So we can get quality risk cost-allocating by manufacturer is λC_E λC_E , and shared by processor is $(1-\lambda)C_E$.

5 Analysis of Case

In an agricultural supply chain, the cost of manufacturer to guarantee the agriculture products quality prevention level is $C(p_m) = g \frac{P_m^2}{1-P_m}$, the cost of processor to guarantee the evaluation level is $C(p_b) = f \frac{P_b^2}{1-P_b}$, $f = 0.8$; $g = 0.6$ [9], if

$$\Pi_b = 100, \Pi_m = 90, C_l = 15, C_E = 55, C_E = 55, R = 10,$$

$\bar{P}_m = 0.6$, $\bar{P}_b = 0.3, (K_1, K_2, K_3, K_4) = (0.3, 0.2, 0.4, 0.1)$ (given firstly). We can get $(P_m^*, P_b^*) = (0.87, 0.57)$; under information symmetry and quality cost-return $Z = 90.33$ under the whole agricultural supply chain. Based on this, we analysis quality cost-allocating under information asymmetry:

(1) On the situation of hiding P_m and detection P_b , we know already $P_b^* = 0.57$, if $K_2 < K_1$, it get $\lambda = 0$, $\Delta\Pi = 56.49$ under the condition of the internal loss occurs. It is calculated by maple software, $P_m = 0.87, Z_b = 19.81, Z_m = 71.09$. The result is optimum quality cost-return of the whole agricultural supply chain. When external loss occurs $\Delta\Pi = 0$, λ must greater than one to fulfill all conditions. So there has no reasonable key. If $K_2 > K_1$, $\lambda = 0$, $\Delta\Pi = 59.12$ under the condition of the internal loss occurs. It is calculated by maple software $P_m^* = 0.87, Z_b = 9.92, Z_m = 80.69, Z = 90.6$. The result is optimum quality cost-return of the whole agricultural supply chain. When external loss occurs $\Delta\Pi = 0$ and $\lambda = 1.42$ great than one, so there has no reasonable key.

(2) On the situation of hiding P_b and detection P_m , we know already $P_m^* = 0.87$, so manufacturer gives fixed pay to processor. There has no problem of price discount and producers won't allocate the quality risk cost.

(3) On the situation of both sides are hidden:



① If $(K_1, K_2, K_3, K_4) = (0.3, 0.2, 0.1, 0.4)$, $K_2 < K_1$ and $K_3 < K_4$, it calculates that $(\Delta\Pi, \lambda) = (27.9, 0.69)$ in other words, manufacturer make a markdown of 27.9 under the condition of the internal loss occurs; when external loss occurring, manufacturer shares the risk cost 37.95, and processor shares the risk cost 17.05. Then the combined distribution of $(\Delta\Pi, \lambda)$ lets $Z_m = 80.83$, $Z_b = 9.71$, $Z = 90.54$ the result is the optimum quality cost-return of the whole agricultural supply chain.

The process of analysis ②③④ is the same as ①, as follow:

② $(K_1, K_2, K_3, K_4) = (0.3, 0.2, 0.1, 0.4)$, $K_2 < K_1$ and $K_3 > K_4$, $(\Delta\Pi, \lambda) = (31.6, 0.6)$ manufacturer markdown 31.6 under the condition of the internal loss occurring; Manufacturer share 33, processor share 22 under external loss occurring $Z_m = 80.09$, $Z_b = 9.81$, $Z = 90.9$.

③ $(K_1, K_2, K_3, K_4) = (0.2, 0.3, 0.1, 0.4)$, $K_2 > K_1$ and $K_3 < K_4$, $(\Delta\Pi, \lambda) = (29.4, 0.72)$ manufacturer markdown 29.4 under the condition of the internal loss occurring; Manufacturer share 39.6, processor share 15.4 under external loss occurring $Z_m = 80.42$, $Z_b = 9.88$, $Z = 90.3$.

④ $(K_1, K_2, K_3, K_4) = (0.2, 0.3, 0.4, 0.1)$, $K_2 > K_1$ and $K_3 > K_4$, $(\Delta\Pi, \lambda) = (33.26, 0.63)$ manufacturer markdown 33.3 under the condition of the internal loss occurring; Manufacturer share 34.65 processor share 20.35 under external loss occurring $Z_m = 80.67$, $Z_b = 9.94$, $Z = 90.6$.

These results are the optimum quality cost-return of the whole agricultural supply chain.

Thus it can be seen from specific data instance of the two-stage agricultural supply chain that:

1) On the situation the hiding of one side information, manufacturer gives the price discount $\Delta\Pi$ under internal loss cost occurring, and the whole of external loss cost is taken by manufacturer. It guarantees the optimum quality cost-return of the whole agricultural supply chain.

2) On the situation the hiding of both sides, manufacturer gives the price discount $\Delta\Pi$ under internal loss cost occurring. With external loss occurring, quality risk cost be allocated by manufacturer and processor according to λ , the four situations get reasonable $(\Delta\Pi, \lambda)$ combined solutions together. The result is the optimum quality cost-return of the whole agricultural supply chain.

6 Conclusions

In the supply chain of agricultural with a manufacturer and a processors, taking optimum quality prevention level and quality evaluation level as the decision variable under the symmetrical information condition, according to centralized and decentralized quality cost- allocating models, the article analyses the problem of

quality cost-allocating under the asymmetrical information condition. The results indicate that quality assurance cost and invested cost does not need to allocate, which depend on quality prevention level and quality evaluation level. Quality risk cost: when it occurs internal loss, offering relevant discount price by a manufacturer substitutes for quality cost-allocating condition; when it occurs external loss, the two sides can determine quality risk cost-allocating rates amount by means of the invested cost rate. Therefore, in reasonable quality cost-allocating condition, it not only can meet to the requirement of consumer, but also can optimize quality cost coordination management of supply chain and promote the supply chain of agricultural to achieve "win-win" or "common". The conclusions are significant to guide the quality cost coordination management between members in the supply chain, but principal-agent relations of asymmetric information are mostly repetitive game theory of multiple stages in various ways. Therefore, the study of agricultural products quality cost-allocating on repetitive game theory of multiple stages in various ways will be the main direction of further efforts.

References

1. Mazé, A., Polin, S., Raynaud, E., Sauvée, L., Valceschini, E.: Quality Signals and Governance Structures within European Agro food Chains: A New Institutional Economics Approach. Paper Presented at the 78th EAAE Seminar and NJF Seminar 330, pp. 15–16. Economics of Contracts in Agriculture and the Food Supply Chain, Copenhagen (2001)
2. Hennessy, D., Roosen, J., Mirannowski, J.A.: Leadership and the Provision of Safe Food. *American Journal Agricultural Economics* 83(4), 862–874 (2001)
3. Zhu, K., Zhang, R.Q., Tsung, F.: Pushing Quality Improvement Along Supply Chains. *Management Science* 53(3), 421–436 (2007)
4. Zhang, C.-H., Huang, X.-Y.: Effect of Asymmetric Information on Supply Chain Quality Cost Decision. *Journal of Northeastern University (Natural Science)* 24(7), 303–306 (2003)
5. Zhou, M., Zhang, Y., Li, Y., Dan, B.: Optimal Contract Design in the Quality Management of Supply Chain. *Journal of Project Management* (3), 120–122 (2006)
6. Cao, J., Yang, C.-J.: A Game Analysis on Quality Failure of Supply Chain. *China Management Science* (2), 25–29 (2006)
7. Lars, W., Hans, W.: *Contract Economics*. Economics Princeton University press, Beijing (2001)
8. Chopra, S., Meindl, P.: *Supply chain management: strategy planning, and operation*, pp. 36–39. Prentice Hal, New Jersey (2001)
9. Baiman, S., Fischer, P.E., Rajan, M.V.: Information, Contracting, and Quality Costs. *Management Science* 46(6), 776–789 (2000)

Research on Role of Learning Organizational Theory to the P.E. Majors in Teachers College

Yao Bin-bin

Physical Education Department of Anqing Teachers College
yao_summer@sohu.com

Abstract. Physical Education should emancipate our minds, new ideas, from the full course content, improve teaching methods, good sports environment, building a few aspects, in order to develop a comprehensive social development of the compound talents. This article focuses on the P.E. to improve the overall quality of the role of students, to explore the implementation of Physical Education in effective ways and methods. And to study type organization theory the usage carried on a viability assessment at the teaching process and put forward to is the speculation of study of study type organization the study change of classroom.

Keywords: learning organizational theory, P.E. Majors, teaching, research.

1 Introduction

The shared vision means to unite the different people with different personalities through the shared will, value and mission, gradually breaking through the toplimit in organization members' personal ability growing in order to make progress for the shared aim. To improve mental models is to breach the traditional thinking patterns to the disadvantage of the development of organization and to overcome all factors influenced organization activities efficiency. Instead, it aims to improve organization activities efficiency by use of new ideas and ways, methods in favour of its development. The feature of teamwork learning breaks through the individual's being enclosed learning state and organizes the active discussion in members for the purpose of making progressing together. The system thinking requires the notice of the interaction among all factors. It requires one should not take link between the learning organization and other organizations one-sidedly, regarding the links in learning organization activities as the isolated states. It is a kind of beneficial attempt to apply the learning organization to colleges and universities' educational reform especially to the teaching reform in sports major. Currently, educational reform and innovation are the hot topics in colleges and universities. During the process of educational reform, not a few teachers borrow teaching ideas from higher education and educational psychology, getting the good teaching achievement. But, in fact, the experiences and theories on educational reform we can use are not only limited in the fields of education. The learning organization theory in the field of management caters to the improvement of students' passive and objective role in the traditional classroom

teaching and the need of improvement of teaching efficiency. It also supplies an effective way on how to train students' teamwork consciousness and ability, as well as the sustainable development in learning ability[1-5].

2 P.E. Is Important Means to Achieve Quality Education

In addition to physical education students follow the law of education on the comprehensive development of the body, enhance physical fitness, sports knowledge and communication skills but also the students through sports and effective ideological and moral education, intellectual development of students, cultivate sentiments of students, the development of students personality, to promote the socialization of students. So sports is not only an important part of quality education, but also an important means of quality education[6].

(1) The purpose of PE is to make college students have the habit of lifelong physical and capacity. PE is the basis for lifelong physical, life-long sports the most basic and critical and most important part of playing college sports for the past and the role of lifetime sports.

(2) PE is to develop the moral will of the quality of students the most effective way. Helping to train and raise the collective consciousness of students, collaboration awareness, aggressive awareness, sense of innovation, law-abiding consciousness of moral quality[7].

(3) College Sports help promote intellectual development of students and improve learning efficiency. Suitable for physical exercise can improve the brain's balance, flexibility and thinking ability, to ensure abundant energy necessary for learning to enable students to develop physical and mental harmony.

(4) College sports can help improve psychological quality, perfect personality for college students. Sports can develop courage, persistence, self-confidence, perseverance, determination, smart, cheerful, enthusiastic, and other modern personality.

(5) Contribute to the promotion of PE Socialization of College Students. Sports activities are open, social, competitive, practical, because physical activity is always in a social environment, which always occurs with other exchanges and contacts, people in the movement can better overcome the eccentric, to forget troubles and pain, and coordination of relationships, expansion of social interaction and improve social adaptability[8].

(6) PE helps students aesthetic quality. Students in physical activity can directly feel the power and beauty of harmony, physical beauty and spiritual harmony of the United States, while a variety of sports from the body to appreciate the natural beauty of the human body, the development of the students found that the United States, the pursuit of beauty, creating beauty and the ability to appreciate beauty[9].

Accelerate the development of the 21st century is an intense competition, frequent exchanges, international century, personnel quality and training of personnel on a higher requirement. Cultivate and foster a new generation of high-quality 21st century requires us to use "quality education" concept of physical education for rational thinking and bold exploration, the real implementation of the "health first" and "people-oriented" guiding ideology, fully reflect PE Teaching and practice of the dominant position of features to maximize their physical and social benefits, to

encourage students with "good health, the scientific mind, the spiritual transformation of society."

3 Enhancing Mental Health Guidance

According to the Ministry of Education of 126,000 college students nationwide mental health sample survey showed that 20.33% of the students there are different degrees of mental illness, especially with the accelerated pace of life, learning competition, employment continued competitive pressure increased so that today's college students, in varying degrees of psychological disorders and mental disorders the opportunity to increase, and increasing year by year. This shows that strengthen the mental health education has become essential. "Teaching of Physical Education Curriculum Guidelines" will be classified as sports participation physical education, motor skills, physical health, mental health, social adaptation, five fields of study. Active in mental health education, not only to cultivate all-round development of qualified human resources, improve the quality of education is of great significance, but also to strengthen the building of spiritual civilization and promote social stability and unity, to improve the quality of the whole nation of great significance. Particularly in the emphasis on quality education today, in school sports in the implementation of necessary mental health education, to protect the health of college students full physical and mental development.

The human body and is closely related to psychology, mental health education and into the PE in order that students achieve the body and mind through sport. The only way to play sports in college mental health education in the specific function, to better carry out mental health education, in-depth study and to fully exploit the Physical Education on Mental Health of special significance and role of education, school sport to enrich and improve the Mental health education approaches, methods and means.

4 Adjust Teaching Content

Over the years, college physical education by the impact of exam-oriented education, the teaching materials to teach sports knowledge, technology, skills-based, content old, single mode, can not stimulate students learning initiative. Reform efforts should focus on the comprehensive development of the quality of students, combined with the actual school, sports highlights basic theory, quality education and lifelong physical education curriculum. 21st Century Education culture from the past simply "knowledge-based" talent into developing a "creative", "comprehensive" talent, for P.E. to meet the needs of college to meet the social development and change.

In practice, in addition to teaching content, promote our national traditional sports, but also to take into account the knowledge and information age students to modern competitive sports pursuits. Sports has obvious characteristics of the times, a strong sense of competition, unyielding spirit, these are modern people must have the basic quality. The key now is how to make the teaching of modern competitive sports content, simple, interesting, games, living, sports closer to the modern student life, meet the sports needs of more students, so students can enjoy all sports, to improve quality of life of the current urgent problems of P.E.

5 Strengthen Construction of Teachers

Comprehensive quality education, the key is to build a modern educational concepts and implementation capacity of the high level of quality education teachers. P.E. teachers must start from their own, in addition to changing the teaching ideas, teaching methods, teaching contents, the most important thing is the cultivation and improvement of their quality. P.E. teachers in addition to master a professional techniques, but also other specific sports to master the technical capacity, improve and optimize the knowledge structure, to achieve "multi-skill." Strengthen basic theoretical knowledge and related subjects of study, expanding the scope of knowledge to improve the cultural accomplishment.

6 Strengthen Management and Build a Good Sports Environment

Good sports culture on campus students awareness of fitness and improve the overall quality plays a subtle role, so to give full play to the role of schools at all levels of sports organizations, organized and rich in content, diverse, ample and convenient facilities, scientific and effective guidance, thus creating a good sports atmosphere.

Implementation of "integration inside and outside class." Planned and organized extracurricular physical activity is a continuation and extension of classroom teaching. Because the movement techniques and skills to master, are enhanced physical body through repeated practice, the leap from quantitative to qualitative change, classroom instruction alone is difficult to achieve. Therefore, extra-curricular sports activities to enhance the teaching organization and management, teaching the new century the question must be answered.

7 Application of Learning Organization in Teaching of P.E. Major in Universities and College

Cultivating Cooperative Consciousness and Summarizing and Improving Teaching Method Combined with the first practice—personal mastery, the teacher should firstly propagate the advantage of learning organization and mobilize students to join the organization. Competition in class, election of the representatives of subjects, comparing and appraising of groups, setting up progress award and other activities should be held. Its purpose is to place students in dynamic environment and to make the members of class, representatives of subjects; groups and organizations surpass themselves and progress constantly. The second purpose is to make the members of organization process collective honor and the achievement and pride of conquering themselves. The third purpose is to give students inexhaustible motive of study by continuous self-motivation and stimulation of external condition. What's more, by combining the second practice to improve the mental model, physical education teachers' traditional managing modes and teaching ideas should be alter in studying-type class. The theory of learning organization focuses on giving full play to the positivity, initiative and creativity of the members. It also emphasizes humanism,

because human is both the production factor and the factor different from other material form, highlighting that people are a kind of living elements. Teacher as the organizer and guides of the organization is the supervisor of executing the decision and provides the time and space guarantee for organization learning. Teacher should ensure that the information of organization learning is unblocked, try to communicate with students, develop the communication among students and timely praise progressive students. Teacher and students should be together to question the problem in P.E. teaching and investigate and find the solutions.

8 Conclusions

Presently, P. E. Major Teaching in China's institutes of higher learning exist a lot of disadvantages, for example, teaching methods are single and outmoded, course contents break away from practice and teaching arrangements lack the links of practices, etc, the cause of which is that students lack the study enthusiasm and therefore the teaching effect is not that favorable. However, the Learning Organization Theory has a strong objective, which, specifically, can not only enable the teachers to become the organizer, guider, regulator, decider, instructor and participator instead of the role as an authority in an activity so that teacher objectively treats students' individual difference and teaches them according to their aptitude, but also stimulate students' learning enthusiasm and initiative to better excavate students' potential so that students constantly break limitation of ability to improve their learning efficiency. Thus, teaching target will be more reasonable. In addition, it also makes teacher and students to learn communication, cooperation, sharing and creating so as to cultivate the ability of communication and corporation and team spirit. In the united and positive collective atmosphere, teacher and students can gain positive emotional experience and produce innovative thinking in the interaction. The learning organization theory is a relatively new and strange theory system and its application is still at the embryonic stage of exploration. To use learning organization theory in teaching of PE of college and university aims to find a new idea by the theory and to make it lay the foundation for the further teaching practice.

References

1. Senge, P.M.: *The Fifth Discipline—Art and Trend of the Learning Organization*. Shanghai San Lien Book Store (1998)
2. Senge, P.M.: *The Fifth Discipline—Strategies and Methods of Establishing Learning Organization*. Orient Press (2002)
3. Chen, H.: Analysis on Learning Organization's Value of Improving Teaching Quality in Colleges and Universities. *Suzhou S&T University Journal (Social Science Edition)* (4) (2004)
4. Tu, J.: Practice Strategy of Constructing Learning Organization with Analysis of Class Teaching Cases as the Carrier. *Jiangsu Education* (1) (2004)
5. Meng, F.: *Constructing Learning Organization of Modern Schools*. *Research on Comparative Education* (1) (2002)

6. Zhang, S.: How to Create Learning Organization. Beijing University Press, Beijing (2004)
7. Qu, Z.-H.: Development and the idea of. Sports Science (2003)
8. Lai, T.-D.: Quality Education and College Physical reform. School Sports in China (2006)
9. Wang, Y.-P.: The re-construction of college Physical Education Curriculum. Sports Science Research (1994)

Enterprise Logistics Risk Factors Analysis Based on Interpretative Structural Modeling Method

Liu Yongsheng and Li Yanru

Beijing Wuzi University, Beijing 101149, China
lkantadey@yahoo.com

Abstract. How to manage enterprise logistic risk effectively is a main problem that troubles many enterprises now. In this situation, applying Interpretative Structural Modeling(ISM) method, deeply analyzed eight factors that influence enterprise logistics risk. According to analysis results of enterprise logistics risk factors, established the ISM hierarchy model in enterprise logistics risk factors, offered three suggestions to deal with enterprise logistics risk: forecasting demand, developing new products, managing suppliers; strengthening management, communicating information, outsourcing transportation; making emergency strategies and measures.

Keywords: enterprise logistics risk, risk factors, ISM method, risk analysis.

1 Introduction

After enter the 21st century, the process of the economic globalization is accelerating, science and technology innovation is developing rapidly. Modern logistics is developing very rapidly driven by the economic globalization and science and technology. But an inevitable problem is that enterprise logistics risk is increasingly prominent with the progress of modern logistics industry. Strengthen the controlling of enterprise logistics risk has become an important theoretical and practical problem in the academic and industrial fields. Analyze factors that may cause logistics risk in the process of enterprise logistics activities accurately and objectively, study the relationship among risk factors, to make specific strategies for avoiding risk, improving the operation efficiency of enterprise logistics, and increasing corporation's profits. These have become urgent needs to all kinds of enterprises. Meanwhile, domestic and foreign scholars in this research field have got some achievements, and tried to apply various methods to study enterprise logistics risk.

According the existing literatures, the research of enterprise logistics risk mainly revolves logistics risk analysis, identification, evaluation, prevention and controlling, warning etc. The study included the risk management research on a certain link, or a certain function, or a certain activity. Lin proposed measures and suggestions for the small and medium-sized enterprises on how to effectively avoid purchasing transportation risks based on analyzing transportation risks that may happen in purchasing [1]. Wei *et al.* studied the risks in special supply and its causes in the big

non-manufacturing enterprises, and pointed out managers should make some indicators on risk management to remind them of paying attention to such risks adequately in daily production supply management of enterprises, and strategically reserve the unique technology to prevent risks [2]. She *et al.* set up the enterprise procurement early-warning index system from five aspects: delivery risk, market risk, suppliers' credit risk, personnel quality risk and financial risks. Meanwhile, in the domain of index, by distinguishing objective indicators from subjective indicators and discussing respectively, they made the index system truly and objectively reflect the situation of enterprise procurement activities eventually[3]. Jiang gave a kind of quantitative methods to purchase risk management, used the method of CVaR to set up a mathematical model to quantify risks based on the price fluctuations in the procurement process, and finally the empirical analysis shows that the model provides new ideas for enterprise effectively control purchase risks[4]. Li used AHP and the BP neural network technology and vertical combination method to study the emergency logistics risk, and established the assessment and prediction model of the emergency logistics risks[5]. Yue *et al.* assessed and analyzed petroleum products distribution risks by using the method of fuzzy analytical hierarchy process (FAHP), and comprehensively evaluated these risks[6]. Chen *et al.* set up a fuzzy comprehensive evaluation model for assessment of engineering logistics risk through combining the fuzzy evaluation method and risk factor analysis method[7]. Appleguist *et al* studied risks and uncertainty problems of the chemical production logistics, the key is investment risk [8]. Cavinato put forward that enterprise supply chain logistics risk factors include five aspects, namely physical factors, financial factors, information factors, revolution factors and other related factors, and analyzed these factors in detail[9]. Faisal *et al.* classified the supply chain information risks according to their occurred causes, and then established risk indexes to quantify the information risk, they thought there were many variables that could cause supply chain information risk, managers should pay attention to the most important cause variables when manage risks[10].

The modern enterprise logistics is a comprehensive activity, there are many factors may cause enterprise logistics risk, and the relations among risk factors are very complex. Researchers have profoundly realized it, but failed to find an effective way to clear up the complex relation among enterprise logistics risk factors. Thus researching the relationship of enterprise logistics risk factors is the premise and key to adopt effective strategies to control enterprise logistics risk. Interpretative Structural Modeling Method (ISM Method) is an analytic method which is widely used in the modern engineering system, ISM through logic operations of the adjacency matrix to get reachable matrix, then to decompose this reachable matrix, eventually to decompose the complex system factors into multi-stage level form with distinct level. Interpretative Structural Modeling Method has been widely used in formulating business planning, and urban planning and so on, especially for the establishment and analyzing the social system with multi-objective and complicated relationships, the effect of ISM is very significant. In the research of distribution center locating and competitiveness analysis of logistics enterprise, ISM method has been widely used. Therefore, it's a new try to introduce ISM method to the research of enterprise logistics risk, beginning from the research of the relationship among enterprise logistics risk factors, and then research enterprise logistics risk and its controlling strategies.

2 The Application of ISM in Enterprise Logistics Risk Factor Analysis

ISM belongs to a conceptual model, it can make vague ideas or opinions into intuitive model with good relation structure, it is especially applicable to analyze systems with many variables, complex relationship and unclear structure, thus we can use ISM method to find relations among risk factors.

The analysis of enterprise logistics risk relevant factors. Enterprise logistics activity includes numerous links and complicated factors that may cause enterprise logistics risks. Combining causes of enterprise logistics risk and the related experts' opinions and suggestions, we choose demand factors, supply factors, environmental factors, operation process factors, management factors, technological factors, information factors and transportation factors these eight factors which may lead to logistics risk, as shown in table 1.

Setting demand factors, supply factors, environmental factors, operation process factors, management factors, technology factors, information factors and transportation factors that may cause enterprise logistics risk respectively as $S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8$, the collection of all factors set as S_0 . The eight factors are not independent, but mutually correlated and overlapping. But they affect enterprise logistics risk at different degree and different ability which makes analysis the relationship of these factors become very complex. In order to clear up the hierarchical relationship among these factors that affect enterprise logistics risk, and on this basis to determine the key factors laying deep level, we make further analysis of relationship among factors by using Interpretative Structural Modeling Method (ISM).

Establishing Interpretative Structural Model. (1) Finding the interrelationship among factors. According to modeling steps of ISM, the first step application of ISM is to find the interrelationship among influencing factors, so as to determine the digraph of factors. We invited experts in this field and used the result of investigation about relations among enterprise logistics risk factors. It is shown in table 2 specifically. "○" represent line factors have direct or indirect impacts on the column factors, "△" represent column factors have direct or indirect impacts on rank factors, " S_0 " mean the collection of all enterprise logistics risk factors. Such as S_4 said operation process factors of enterprise logistics risk have a direct effect on S_0 , thus the position of rank S_4 or column S_0 should be filled in the "○".

(2) According the directed relationship among enterprise logistics risk factors in table 2, writing adjacency matrix A. The established adjacency matrix A is a 9×9 order phalanx (The phalanx also included the S_0), adjacency matrix A has the same rank elements and column elements, both are sequence of $S_0, S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8$, "1" in adjacency matrix A said the rank factors have influence on column factors or the column factors have influence on rank factors (Including autocorrelation, namely S_i influence S_i), as "0" said no influence between corresponding factors, as shown in figure 1.

Table 1. Enterprise logistics risk effect factors

| Effect factor | The specific performance |
|-------------------|---|
| Demand | Including inaccurate demand forecasting, relying on a few big clients, unpredictable demand fluctuations due to promotion, changing of important demand delivery, changing of important demand contented rate, customer financial deterioration, frequent and rapid new product marketing, low customer loyalty and reserved customer comments etc. |
| Supply | Including non-science procurement process, depending on key suppliers without alternate, some suppliers imperfect finance, long delivery time of suppliers, production quality problems of suppliers, bad fulfillment of plans and ordering, supply market shortage, lack of debt management capacity, low quality suppliers planning and forecasting, etc. |
| Environment | Including the macro-economy, government policy guidance, environment pollution, power supply disruptions, storm, floods, fires, terrorist attacks events, strike, industrial policy adjustment, etc. |
| Operation process | Including yield lower than expectations, production quality standard lower than market acceptable level, failing to deal with the processing quality and accuracy, system failure and backup prevention operation failure, the enterprise failing to respond to customers' demand timely and effectively, failing to meet requirements, etc . |
| Management | Including comprehensiveness of planning, effectiveness of organization, safety stock policies, asset management policies, regulations and procedures, tax and financial system, lack of resist measures and warning plan for uncontrolled factors, human resources management policy, improper management mechanism, etc. |
| Technology | Including logistics technology innovation, logistics technology maturity, the time cost and safety controlling of new technology application, the rate and acceptability of technology achievements conversion, etc. |
| Information | Including integration and sharing of information, transmission and processing of information , the stability of information system, the compatibility and connection degree of information system, timeliness of information, etc. |
| Transportation | Including transportation delay, damage of cargo in the process of transportation, imperfection of transportation system supporting and auxiliary facilities, transportation workers' safety and moral consciousness, etc. |

Table 2. Directed relationship among factors of enterprise logistics risk

| | S ₀ | S ₁ | S ₂ | S ₃ | S ₄ | S ₅ | S ₆ | S ₇ | S ₈ |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| S ₀ | ○ | | | | | | | | |
| S ₁ | ○ | | △/○ | | △ | | | | |
| S ₂ | ○ | ○/△ | | | | | | | |
| S ₃ | ○ | ○ | ○ | | | | | | ○ |
| S ₄ | ○ | | | | | △ | | | |
| S ₅ | ○ | | | | | | | △ | |
| S ₆ | ○ | | | | | | | | ○ |
| S ₇ | ○ | | | | | | △ | | |
| S ₈ | ○ | | ○ | | | | | | |

$$A = \begin{matrix} & S_0 & S_1 & S_2 & S_3 & S_4 & S_5 & S_6 & S_7 & S_8 \\ \begin{matrix} S_0 \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \\ S_7 \\ S_8 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{matrix}$$

Fig. 1. Adjacency matrix of enterprise logistics risk factors

$$M = \begin{matrix} & S_0 & S_1 & S_2 & S_3 & S_4 & S_5 & S_6 & S_7 & S_8 \\ \begin{matrix} S_0 \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \\ S_7 \\ S_8 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \end{matrix}$$

Fig. 2. Reachable matrix of enterprise logistics risk factors

(3) Calculating the reachable matrix M of the adjacency matrix that said on enterprise logistics risk factors. The calculation of reachable matrix M is namely power operation on matrix $(A + I)$ by applying Boolean principles, until $(A + I)^{k+1} = (A + I)^k = M$, by calculating to obtain reachable matrix on enterprise logistics risk factors $M = (A + I)^4$, as shown in figure 2.

(4) Decomposing of reachable matrix of enterprise logistics risk factors, by the theory of ISM method, adjusting column elements of reachable matrix based on the principle of first set to get reduction reachable matrix M' , as shown in figure 3.

(5) Working out the backbone matrix of enterprise logistics risk factors. Backbone matrix maintained accessibility of reachable matrix, while making elements reaching path minimum, namely box marked sub-matrix on the diagonal of matrix R in figure 4.

$$M' = \begin{matrix} & S_0 & S_3 & S_4 & S_1 & S_2 & S_6 & S_7 & S_8 & S_5 \\ \begin{matrix} S_0 \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \\ S_7 \\ S_8 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{matrix}$$

Fig. 3. Reduction reachable matrix of enterprise logistics risk factors

$$R = \begin{matrix} & S_0 & S_3 & S_4 & S_1 & S_2 & S_6 & S_7 & S_8 & S_5 \\ \begin{matrix} S_0 \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \\ S_7 \\ S_8 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{matrix}$$

Fig. 4. Backbone matrix of enterprise logistics risk factors



(6) Decomposing enterprise logistics risk factors at hierarchy, establishing progressive hierarchy model, as shown in backbone matrix of enterprise logistics risk factors, it mainly divides into four levels, the most top said the ultimate goal of the system, downwards each layer respectively represents the reason of up layer, progressive hierarchy diagram, as shown in figure 5.

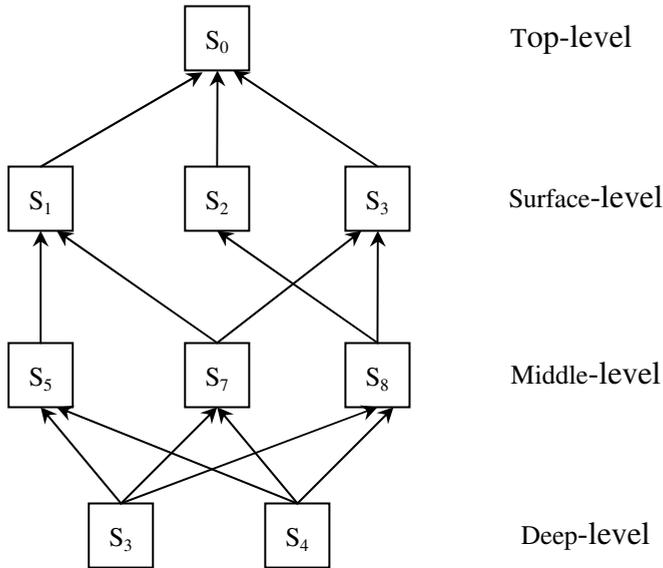


Fig. 5. Progressive hierarchy model of enterprise logistics risk factors

Top-level is the collection of enterprise logistics risk factors, and demand factors, supply factors and technology factors are located in surface-level, it is said they are factors that have the most direct influence on enterprise logistics risk. Factors located in middle-level include management factors, information factors and transportation factors, these three factors will affect surface-level factors indirectly, cause enterprise logistics risk besides directly cause enterprise logistics risk. Environment and operation process factors are in the deep level of the hierarchy model, both of two are mostly influence demand, supply, technology, management, information and transportation, indirectly cause enterprise logistics risk. Therefore, different level factors have the different degree of influence on enterprise logistics risk, when managing logistics risks and formulating the risk management strategies, managers can refer factors' location to give different stress. Factors which directly influence enterprise logistics risk should be predicted and prepared timely response strategies and alternative scheme. Factors which indirectly influence enterprise logistics risk should be prepared some generally and long-term applicable strategies. So not only can timely handling some logistics risks that suddenly happen or have great influence on enterprise, but also greatly reduce time cost on processing some daily logistics risk, save enterprise operation cost mean while increase the income of the enterprises.

3 Conclusions

Through analyzing enterprise logistics risk factors, giving some suggestions that dealing with enterprise logistics risk as follows:

First, forecasting demand, developing new products, managing suppliers. It is found that demand, supply and technology factors would directly cause enterprise logistics risk after the analysis of the eight selected enterprise logistics risk factors by applying ISM method, so managers should realize it as earlier as possible. Enterprises should apply advanced technology to research market deeply and forecast accurately, know about season volatility and regional variation of their own product's demand, keep eyes on the changes of market demand for their products; Enterprises can cooperate with universities or institutes to develop new products, research new production process and improve the scientific and technological content of products, make their products better meet the demand of the market; Meanwhile, the enterprise must strengthen the management of suppliers, develop strategic and cooperative partnership with suppliers, and invite suppliers to participate in the research and development of product to make them know enterprise's requirements for new raw material, thus ensure the supply and quality of raw materials.

Second, strengthening management, communicating information, outsourcing transportation. Management, information and transportation factors locate middle-level in the hierarchy model, their influence on enterprise logistics risk is weaker than demand, supply and technology factors which locate surface-level. Mismanagement would affect product activities, lead to increase of raw material inventory, reduce the satisfaction of market demand, impede information transmission, suppliers cannot understand enterprise's requirements of raw material and the enterprises cannot understand the changes of market demand. The enterprise will become an inefficient system that response slowly. If transportation accidents happened frequently, it would directly influence the whole logistics activity and lead to companies fail to meet the demand, and it also threat personnel safety and damage the social image of enterprise. So the influence of management, information and transportation on enterprise logistics risk is also should not be ignored. Enterprises should strengthen the management of the whole logistics activities, simplify the hierarchy of management and reduce intermediate links when transferring information in order to guarantee information on suppliers-enterprise and enterprise-market transfer smooth and timely; Enterprises should cooperate with the third party logistics enterprise and outsource transportation, which is a non-core activity. Outsourcing transportation to professional transportation companies, not only can get high efficiency but also reduce enterprises logistics cost greatly.

Third, making emergency strategies and measures. Environment and operation process factors are in the bottom of the model, they cause logistics risk to enterprises mainly by affecting other factors, so managers also should give some attention to environment and operation process factors. In China, the macroeconomic environment and government's policy had some certainties, but in recent years, some events of low probability, such as natural disasters, traffic accidents, group events occurred frequently, they made environmental factors be taken seriously more and more. So

managers should make some contingency plans and measures to deal with emergency events meanwhile do the normal management . The smooth operation process will guarantee enterprises logistics activities run normally, so managers should pay more attention to routine check of operation process, solve problems timely and avoid further loss.

It needs to point out that in this article, we applying ISM method to grade enterprise logistics risk factors and distinguish them, offer some references for enterprise managers to strengthen the management of enterprise logistics risk, but in actual operation, enterprise managers should base on the fact of enterprise, use the system approach to manage enterprise logistics risk.

Acknowledgment. This work is supported by the Funding Project for Academic Human Resources Development in Institutions of Higher Learning under the Jurisdiction of Beijing Municipality (PHR (IHLB)) under Grant PHR200907134.

References

1. Lin, R.Z.: Talking About small and medium-sized enterprises how to avoid transportation risk of procurement. *Age Trade* 5, 185–186 (2007)
2. Wei, Q.Y., Tian, Z.L.: Large non manufacturing enterprise special risks of supply chain management. *Chinese Logistics and Procurement* (16), 68–69 (2007)
3. She, L., Yang, Y.Q.: Enterprise purchase management early-warning index system discussion. *Scientific Progress and Strategy* (11), 157–159 (2003)
4. Jiang, W.D.: A new method of purchasing the risk management. *Technical & Economic and Management Research* (2), 74–75 (2007)
5. Li, Z.W.: Emergency logistics risk assessment and prediction model Based on AHP and BP neural network. *Logistics Technology* (9), 75–77 (2008)
6. Yue, W.T., Zhu, M.J.: Product oil distribution logistics project risk identification and evaluation Based on AHP. *Logistics Technology* (3), 85–90 (2009)
7. Chen, H., Wu, Q., Zong, C.Q.: Engineering logistics risk analysis Based on fuzzy comprehensive evaluation method. *Logistics Engineering and Management* (3), 44–46 (2009)
8. Appleguist, G.E., Pekny, J.F., Reklaitis, G.V.: Risk and uncertainty in managing chemical manufacturing supply chains. *Company and Chemical Engineering* (24), 2211–2222 (2000)
9. Cavinato, J.L.: Perspective supply chain logistics risks from the back room to the board room. *International Journal of Physical Distribution & Logistics Management* 34, 383–387 (2004)

A Study and Development of Onboard LED Lighting for Train

Deqiang He, Yixuan Bai, Xiongwen Yang, Xiaochun Su, and Zengqi Yang

Mechanical Engineering Collage of Guangxi University, Nanning, China
hc1258@yeah.net

Abstract. The LED lighting applied to rolling stock adequately based on the operating environment and the characteristic of LEDs is developed. The LED driving circuit, heat sink, and mechanical structure are discussed in the paper, and the simulation to the LED matrix circuit with ANSYS is used to analyze the heat dissipation. The experiment provided the evidence that the train onboard LED lamp designed is adequate for electromagnet and vibration environment of a running train.

Keywords: Train, LED lighting, Thermal simulation.

1 Introduction

Although the railway and urban rail transit is well known as an energy-saving public transit method, their total energy consumption is still tremendous. The traction and illumination system have a large potential in energy saving, which techniques are quite mature now and which makes them a breakthrough point of the energy saving for urban railway transition [1]. Light Emitting Diode (LED) is superior to traditional luminaries for its high lighting efficiency, low power consumption, long lifetime, high reliability, good safety performance, environmentally friendly performance and short response time that make it widely used in every kind of illumination system gradually. The lifetime of onboard LED lamp for train can be as long as 40000 hours with 50% higher lighting efficiency than normal CFL(Compact Fluorescent Lamp). The onboard LED lighting adapt for the railway operation environment is developed, which circuit, heat sink, EMI and vibration designing are mainly described in the paper.

2 Principle of the Onboard LED Lighting Circuit

The circuit of onboard LED lighting is consist with LED driving circuit, LED matrix circuit, PWM dimming circuit, sensors and controlling circuit, as shown in Fig.1.

The average illumination of interior car of train is 150lx [2]. After calculation, the luminous flux needed of every LED lamps should be 1442.8990lm. Considering the transmittance of the transparent board, the actual luminous flux of one LED lamp should be 2404.832lm.

The LL-HP60MW6EB come from Luckylight company is used as the high-power white LED chip. The typical luminous flux under the rated current of 350mA is 100lm. The color of the light is warm white and the rated power is 1W.

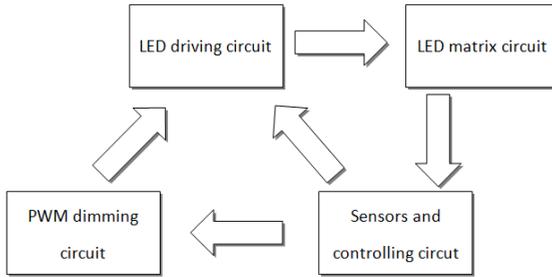


Fig. 1. The circuit structure of onboard LED lamp

The schematic diagram of the LED matrix circuit. LED chips can be connected in series, parallel, series-parallel and cross-tabulation [3]. All in series are used as the LED matrix mode, as shown in Fig. 2.

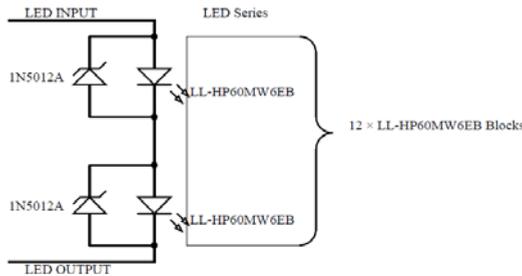


Fig. 2. The LED matrix circuit

To prevent the circuit is shut down by the disconnection of a random LED chip, every LED should be connected in parallel with a Zener diode. When a LED is disconnected with the circuit, its Zener diode will bear the current.

Design for the onboard LED diving circuit. A buck circuit is used as shown in Fig. 3 as the LED driving circuit. The LED INPUT and the LED OUTPUT channel are connected with the anode and cathode of the LED matrix circuit respectively. The HV9910B driving IC is a PWM controller using open-loop peak current mode control in constant frequency or constant off-time mode, of which the efficiency can be above 90% [4].

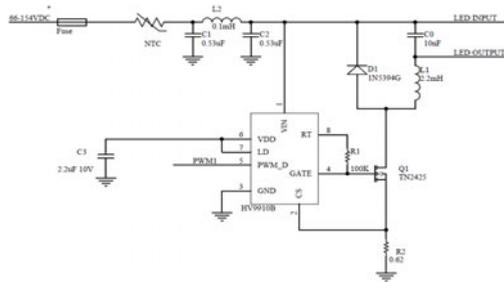


Fig. 3. LED driving circuit



The typical output voltage of train is DC 110V ranging from DC 66V to DC 154V [4]. The voltage of the LED series is ranging from 33.6V to 45V. The rated current of LED (I_0) is 350mA. The efficiency expected is 90%.

Since the maximum duty circle $D_{max}=69\%>50\%$, the HV9910B has to be set in the constant off-time mode, i.e. a timing resistor should be connected between the RT pin and the GATE pin. To set the off-time (T_{off}) to $5\mu S$, the average switching frequency is 110kHz and the timing resistor should be $100k\Omega$.

The filter capacitor (C_0 , about 10nF) is used to lower the ripple current which can bypass any switching peak current, reduce the output EMI effect and increase the lifetime of the LEDs. C_1 is used not only to stable the voltage but also absorb the ripple current in high frequency. Thus, an electrolytic capacitor is inadequate, a metalized polypropylene capacitor or a ceramic capacitor is more appropriate.

$$C_1 = \frac{I_0 \times T_{off}}{0.05 \times V_{min}} = 0.53\mu F \tag{1}$$

V_{min} —the minimum input voltage.

C_1 along with a same capacitor C_2 and a small inductance L_2 form a PI filter which should be near the MOSFET to reduce the area of the high-frequency circulation on PCB. When the MOSFET is switching, C_3 can stable the inner power voltage (V_{dd}) of the HV9910B.

Assuming the ripple current of LED is $\pm 15\%$ (30% in total), L_1 can be calculated as follows [5]:

$$L_1 = \frac{V_{max} \times T_{off}}{0.3 \times I_o} = 2.171mH \tag{2}$$

V_{max} —the maximum input voltage;

The nearest normalized value is 2.2mH, so the ripple current will be lower than 30%.

The effective value of L_1 is 350mA and the rated peak current is 0.4A..

The peak voltage of the MOSFET is equal to the maximum input voltage. Given a margin of 50%: $V_{FET} = 1.5 \times V_{max} = 231V$. The maximum effective current is relate to the maximum duty circle: $I_{FET} \approx I_0 \times D_{max} = 0.24A$. Normally the effective current of MOSFET should be thrice as large as the current above. The TN2425 is appropriate.

Under the worst conditions, $I_{diode} = 0.78 \times I_0 = 0.273A$.

$V_{diode} = V_{FET} = 231V$. So the 1N5394G is chosen as D_2 .

R_2 is calculated using the inner voltage of 0.25V of HV9910B:

$$R_2 = \frac{0.25}{1.15 \times I_0} = 0.62\Omega$$

3 Design for the Heat Sink of Onboard LED Lamp

The basic formulation of heat conduction is:

$$\Phi = -\frac{\lambda \times A \times dt}{dx} \tag{3}$$

Φ —the heat generated or conducted during the conduction;

λ — the conductivity factor;

A —the CSA of the object during the conduction;

dx —the distance of the conduction;

dt —the temperature difference after the conduction;

As aluminum is lighter and has a better capacity of heat-transmission, it is appropriate for train environment and the ideal material for heat sink.

The shape and Size of the heat sink. To conduct more heat during the convection heat transfer, since the conductivity factor and the temperature difference can't be operated, only enlarging the effective area is feasible. To achieve the goal, the fin shape is appropriate for heat sink, as shown in Fig.4.

The heat sink is a half circle. The dimensions of the heat sink are: width $B = 30\text{mm}$, length $L = 1166\text{mm}$, thickness of fins $\delta = 2\text{mm}$, pitch $P = 4$, thickness of fin base $\epsilon = 3\text{mm}$, quantity of fins $n = 8$, the height of the highest fin $l_{\text{max}} = 12\text{mm}$, the height of the lowest fin $l_{\text{min}} = 7\text{mm}$.

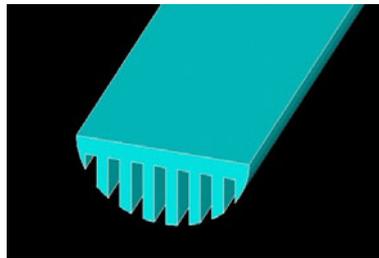


Fig. 4. Structural model of the heat sink

Check the parameters. Assuming that the radiating area of every fin is A_f , the total base area between fins is A_b and the temperature of the base is $t_f = 25^\circ\text{C}$, the total surface area is $A_t = A_b + nA_f$. If the base and the fins share the same conductivity factor, the total heat conducted is the sum of both [8]:

$$\Phi_t = hA_b\theta_b + hA_f\eta_f\theta_b \tag{4}$$



$$\text{Fin efficiency } \eta_f = \frac{th(mL)}{mL}, \theta_b = t_b - t_f, m = \left(\frac{2h}{\lambda\delta}\right)^{\frac{1}{2}}.$$

The conductivity factor of aluminum is 236 W/(m×K). The power of the LEDs is $\Phi_i = 12\text{W}$. The free convection factor of air is 1~10W/(m2×K). As the onboard lamp, 5W/(m2×K) is appropriate, $h = 5\text{ W}/(\text{m}2 \times \text{K})$. To simplify the calculation, we assume that the heat sink is rectangle:

$$A_f = 8 \times 2 \times l_{\min} \times L \tag{5}$$

$$A_b = B \times L - A_f \tag{6}$$

According to (4), $\theta_b = 16.35^\circ\text{C}$, $t_b = 41.35^\circ\text{C}$, lower than the maximum temperature of 80°C. When $l_{\max} = 12\text{mm}$, $\theta_b = 10.001^\circ\text{C}$. So the temperature of the base is $t_b = 35.001^\circ\text{C}$, lower than the maximum temperature of 80°C.

According to the calculation above, when the thickness of a fin is between the maximum and the minimum, the temperature won't pass 80°C. Checking process: Assuming $l = 10\text{mm}$, then $\theta_b = 11.840$. The temperature of the base would be $t_b = 36.840^\circ\text{C}$, lower than the maximum temperature of 80°C.

Finite-element analysis of the heat sink. Assuming the heat-flow density of the upper surface is 274.44W/m2, the result of the FEA of the heat sink is shown as Fig. 5:

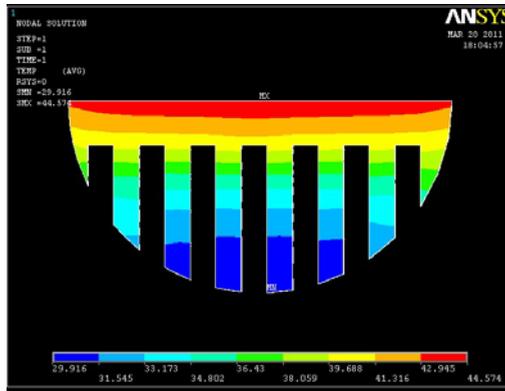


Fig. 5. Temperature distribution of the cross section of the heat sink

The highest temperature of cross section temperature distribution is 44.574°C Zone in red represents that the highest temperature appears on the upper surface on which set the luminaries, i.e. the work temperature of the LEDs is 44.574°C, and it reach the goal that the LEDs lighting work under 80°C. The lowest temperature is 29.916 °C, which exits at the end of the fins like it planned. The temperature is gaining higher from the end of the fins to the upper surface of the base and different colors represent different distribution of temperature. As shown in Fig.5 that the result of the FEA meets the calculations of the theories, suggesting that the structure of the heat sink is appropriate.

4 Mechanical Structure LED Lamp

The structure of LED lamp is as shown in Fig. 6, 1 represents copper needle, 2 represents sleeve, 3 represents heat sink, 4 represents PCB and 5 represents LED chips. As the space available for setting the lamps is limited in the interior car, normally there are a series of LED lamps on the either side of the walking passageway. Along with the lamp-socket, the length of each tube is 1200mm with a distance of 300mm. The wild used T10 lamp-socket is used of which the diameter is 31.75mm. The diameter is set 32mm and the width that touches the lamp cap is 12mm. So the length of the diameter of the LED lamp cap is set 32mm. Except the thickness of the lamp cap, the outer diameter should be 30mm. The total length of the tube is as long as the tube used in trains, about 1200mm.

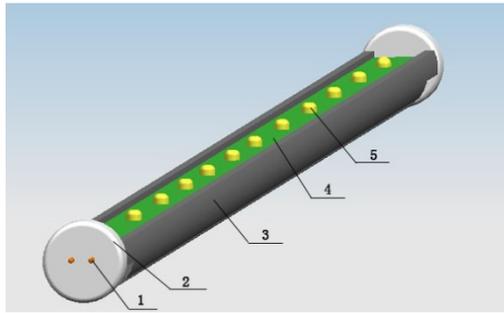


Fig. 6. The structure of the LED tube

The PCB is inserted into the heat sink, and the sleeves at the ends of the tube fasten the heat sink and the PCB together. The copper needle on the lamp cap is used to conduct electricity and the lamp-socket is set in the interior car.

The structure of the heat sink is as shown in Fig. 7. There is a wedge slot to combine the PCB and the heat sink which is as thick as the PCB. During the setting up, a transition fit is used between the slot and the PCB to make sure that the PCB touch the heat sink which avoid the using of fastener and make it stable and easy to repair.

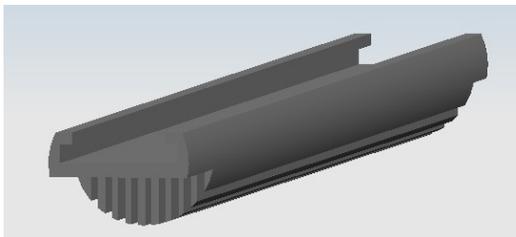


Fig. 7. The structure of the heat sink

5 Design of the Lampshade

The area of the lighting of the chip chosen ranges from -100° to 100° . It is larger than 180° which makes the car covered by lighting. The illumination reaches the peak at 0° and fades as much as it leaves the 0° position. It means that the illumination isn't even and can't meet the request of uniformity of 1: 1.3. So a well-proportion illumination is needed. As shown in Fig. 8, the angle between the incident light L_1 and the emergent light L_4 is $f(\Delta\theta) = \angle 4 + \Delta\theta - \angle 1$, which represents the light change after it pass the lampshade.

$$\frac{df(\Delta\theta)}{\Delta\theta} = \frac{\sin \angle 1}{\sqrt{1 - \sin^2 \angle 4}} \left(\frac{\cos \Delta\theta * \cos \angle 2}{\sin \angle 2} \right) + \sin \angle 1 * \sin \Delta\theta \geq 0$$

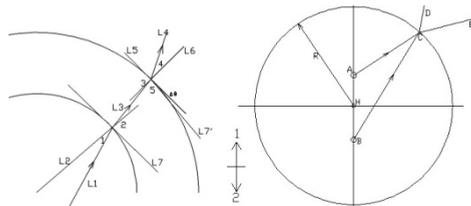


Fig. 8. Process of the reflection

The light deflects after passing the lampshade. When the thickness of the lampshade h is stable, the bigger the curvature appears, the bigger the $\Delta\theta$ is which means the more the light deflects.

As shown in Fig. 8, H represents the center of a lampshade with R as a radius of curvature, A and B represents the different position of a LED. With the conclusion above and Fig. 8, when the LED is above H, the lampshade diffuses the lights while when the LED is lower than H, the lampshade gathers the lights.

To make the illumination even, the space that share more lights needs the lampshade to diffuse the lights and the deflection should be stronger. Here it come up with a conclusion: the closer the lampshade is to the 00 position, the bigger the curvature is and the center of the lampshade should maintain lower than the LED.

6 Summary

The onboard LED lighting which including circuit, heat sink, mechanical structure, etc, based on the special request of the operating environment of train, is designed in the paper. Compared with the lighting used now, the LED lamp not only is less power-wasting, but also can extend the lifetime to 10 times, appropriate for the energy-saving request of railway and urban rail transit.



Acknowledgements. This work was financially supported by the Guangxi Key Technologies R & D Program with Grant No. 0992006-4 and the Guangxi Manufacture System & Advanced Manufacture Technology Key Laboratory Project with Grant No. 0842006_007_Z.

References

1. Gong, W.: Energy Technology 30, 176 (2009)
2. Guo, J.F., Liu, W.P., Hu, X.C.: Rolling Stock 47, 33 (2009)
3. Zhou, Z.M., Ji, A.H.: Design and example of high-power white LED. Posts and Telecom press, China (2009)
4. Steve, W.: Power Supplies for LED Driving. Posts and Telecom press, China (2009)
5. Yan, Y.S.: Electric Drive For Locomotive 42, 54 (2001)
6. Yang, S.M., Tao, W.Q.: Heat Transfer Theory. Higher Education Press, China (2007)
7. Zhao, Z.N.: Heat Transfer Theory. Higher Education Press, China (2005)
8. Tang, X.L., Fan, Q.B., Zhang, C.H.: ANSYS engineering tutorial articles of heat and electromagnetism. China Railway publishing houses, China (2003)

Study on Agricultural Products Logistics Mode in Henan Province of China

Wang Liping

School of Economics and Management, Henan Polytechnic University, Jiaozuo 454000, China
lin_li123@sina.com

Abstract. This paper analyzed the status and problems of agricultural products logistics development in Henan Province of China, put forward the general concept, guiding principle and development goal of modern agriculture logistics development in Henan Province, which will provide theoretical and decision-making references for the relevant government departments and enterprises.

Keywords: Third party logistics, Agricultural products, Empirical analysis.

1 Introduction

Modern third-party logistics has the characteristics of large-scale, fast distribution, wide area, and high efficiency. Therefore, the development of modern logistics has great significance for the development of rural markets, and modern logistics not only provide true, accurate and effective information for business entities, reduce market uncertainty and blindness during the operation, improve the circulation speed of agricultural products, but also reduce the circulation cost of agricultural products, achieve value-added in the process of agricultural products circulation, protect the fundamental interests of farmers, and promote modernization of agriculture, increase the overall efficiency of agricultural production.

2 Introduction to Third-Party Logistics of Agricultural Products

2.1 The Definition of Agricultural Products Logistics and Third Party Logistics

Ministry of Agriculture of China, Rural Economy Research Centre (2005), proposed the definition of agricultural products logistics: agricultural products logistics is a branch of the logistics industry, refers to physical flows of physical entities and related information from producer to consumer that satisfy consumer demand, including agricultural production, acquisition, transportation, storage, loading and unloading, handling, packaging, distribution processing, distribution, and information activities. Development objectives of agricultural products logistics is to increase value-added of agricultural products, save distribution costs, improve circulation efficiency and reduce unnecessary losses, to some extent avoid market risks [1-2].

Third-party logistics of agricultural products is an emerging concept, therefore, the scope of this concept can be expanded appropriately to promote its development.

Currently, the world's major agricultural products logistics mode can be summed up in two: one is European and American mode represented by the U.S. Due to large-scale agricultural producers in Europe and America, the development of large supermarkets selling produce is very fast, so some features of wholesale market have been gradually weakened by integrated production and marketing organization. Another mode is based on the representative of Japan, including Korea, Mainland China and Taiwan and other East Asian countries.

2.2 The Characteristics of Agricultural Third-Party Logistics

The natural attributes of agricultural products determine that agricultural products logistics have different particularity from the general logistics:

A. For fresh agricultural products, its requirements for logistics facilities is particularly high, including logistics equipment with fresh, chilled, and prevention of diseases function.

B. Agricultural products logistics has a strong seasonal and regional characteristic, which is determined by the life cycle of agricultural products.

3 Comparison of Agricultural Products Third-Party Logistics between China and Some Developed Countries

3.1 Status of Agricultural Products Third Party Logistics Development

Cost of agricultural products logistics is high, and value-added ability is weak. In 2003, China's grain logistics costs account for more than 40% in all costs composition, while that of the fresh products account for about 60%, even more, while the logistics costs in developed countries is generally controlled in about 10% [3]. For example, the CIF value of agriculture products in America, Japan, Canada and some EU countries is far lower than the production price of agricultural products in China. According to relevant statistics, China's agricultural products, particularly fresh produce, logistics costs are generally higher than those of other countries, as shown in Table 1 [4-5]:

Table 1. Comparison of national agricultural products logistics fee

| | America | United Kingdom | China |
|--|---------|----------------|------------|
| The share of logistics costs account for the goods price (%) | 10-32 | 8-25 | 20-60 |
| The item has high proportion of logistics cost (%) | food 32 | food 25 | food 30-60 |

3.2 Development of U.S. Agricultural Products Logistics

U.S. agricultural production and trade is a world leader. U.S. has a large, smooth, complex, efficient logistics system of agricultural products. American fruit logistics is more typical, the products after harvest have been in low-temperature environment: the precooling - cold storage - refrigerated truck - cold storage in wholesale station - supermarket refrigerator - consumer refrigerator, which make the loss rate of fruits and vegetables in the physical distribution is only 1% -2% [6-7]. Its specific operation mode of agricultural products is showed below.

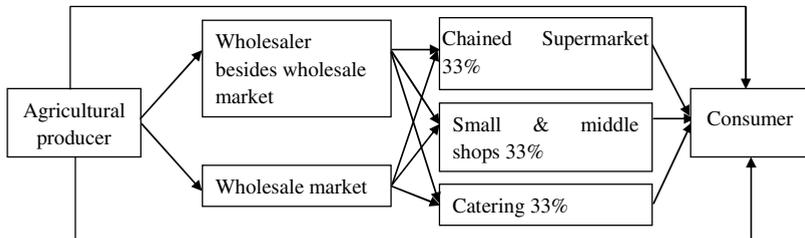


Fig. 1. U.S. agricultural logistics operation mode

3.3 Development of Japanese Agricultural Products Logistics

In Japan, agriculture logistics management system has been formed. This management system not only eliminates the drawbacks of overlapping management from different departments, in line with a unified operating principle of agricultural production and circulation under the conditions of the market economy, but also meets the law of the agricultural product commercial production and operation, in line with the common interests of producers and consumers, reduce administrative costs and improve administrative efficiency.

Ministry of Agriculture of China, Rural Economy Research Centre, thinks that Japan's agricultural cooperative organizations played an active role in the circulation of agricultural products. The agricultural association, as a major supplier groups, has modern logistics advantage, such as processing, packaging, transportation, information networks. According to statistics, 80% -90% of Japan produce reach the hands of consumers through the wholesale markets.

The above experience in several countries tells us that if we want to develop the logistics of agricultural products, we must do three things: (1) the more support (including public facilities, the delivery of supporting funds, the construction of information infrastructure); (2) laws and policies regulating the market; (3) agriculture organization play its due role in coordinating the relationship between farmers and the government.

4 The Development Status of Henan Province Agricultural Products Third-Party Logistics

Overall, the logistics operation mode of industrial enterprises in Henan Province still remain in the self-operation logistics mode, logistics outsourcing is not high, with high operating costs and low efficiency. Many companies are actively seeking suitable logistics agents, if the social logistics services can meet the needs of enterprises, there will be 80% of enterprises outsource logistics business. According to the modern logistics planning in Zhengzhou City, by 2010, third party logistics account for 15% of the total social logistics. This figure is still far behind the proportion of 60-80% in developed countries, showing that the current third-party logistics industry in Henan Province is extremely low [8].

Henan is located in the Central Plain, so the position advantage is very obvious. Henan agricultural logistics infrastructure has been rapid development, including road transport, railway transport, information and communication. Although the logistics needs potential of our province is great, but many companies do not need third-party logistics. Well-known retailers such as Zhengzhou Dennis Store, has established a fairly complete distribution system, they can check the store's merchandise sales at any time through the network. Dennis headquarters say that they will not consider outsourcing, for fear that if they deliver the logistics business to a third party logistics company, one side, distribution system they build own will be wasted; second, Zhengzhou's commercial competition is severe, if letting third-party control their operation, it is likely to result in the disclosure of trade secrets.

The wrong concept and market environment factors currently result in a serious shortage of third party logistics in our province, the development of third-party logistics enterprises is lagging behind. There are very few third-party logistics companies such as Henan Changtong Logistics Company performed active in the market.

5 The Constraints of Henan Province Agricultural Products Third-party Logistics

According to the survey: because the produce's own particularity, poor awareness of the logistics industry and government policies do not keep up with the situation, the logistics development of agricultural products in Henan Province is not optimistic. The restricting factors of Henan agricultural products logistics are [9]:

(1) Poor awareness

Most companies now more interested in self-logistics sector. Although part of the surveyed companies know that the services quality of professional logistics sector is better than the self, but the traditional concept and institution and lack understanding of third-party logistics make the final choice change to self-logistics. According to the author's survey, many watermelons from Anhui, Shandong and so on organize their own cars to ship melons and sell, very few watermelon do this through the professional logistics company. Overall, the province's agricultural logistics system is now outdated, fragmented, and inefficient.

(2) From the main supply subject of agricultural products third-party logistics

The number of supply subject is large, but the intensification is low and level of capacity is weak, such as low service level, scattered distribution of resources, poor coordination and cooperation; the small size of the supply individual, especially poor social and professional level of third-party logistics economies, serious duplicate allocation of supply resources and so on.

(3) From the main demand subject of agricultural products third-party logistics

Because the distribution of rural areas is vast and agricultural production is dispersed, a large part of agricultural products processing enterprises located in the towns, so the main demand subject is dispersed and logistics market environment is poor; because the concept and different organization levels, the development of demand subject of agricultural logistics is not balanced, the difference between supply and demand is large; because of a significant proportion of self-logistics, effective demand for third party logistics is insufficient.

(4) From the infrastructure of third-party logistics

Although the logistics infrastructure of agriculture developed rapidly, but there is still a great gap. The modernization level of most infrastructure is low; integrated support services and facilities should be further constructed; the quality and functionality of most storage facilities is behind; layout is unreasonable, with very clear regional differences, urban developed, rural backward, the tendency of blind construction is obvious, such as the construction of agricultural logistics center lacked overall planning.

(5) From the information system of third-party logistics

The development of agricultural logistics information system lags behind, lack of unified planning and design, lack of regional agricultural logistics exchange platform; development is unbalanced, the gap between enterprises and between urban and rural areas is big; logistics information technology in agriculture has not yet widely used, lack of collection channels and technical standards of agricultural logistics information; logistics software uneven, lack of norms, lack of evaluation and certification system; public platform construction is lagging behind, especially e-government platform related with agriculture logistics have not interoperability.

(6) From the policy safeguard of third-party logistics

Logistics system does not sound and perfect, and there is no regional logistics development strategy and planning, the development of agricultural logistics is in large blindness; agricultural logistics development industrial policy is not issued; agricultural logistics management system is in a fragmented state, lack of a whole management and coordination system; modern agricultural logistics regulatory system is not formed; agricultural logistics standardization is in the initial stage, the technology, services, informational and managerial standards are not yet established; agricultural product and food quality management and inspection system is not perfect; the quality of agricultural logistics employees is generally not high.

(7) Inadequate investment

All along, the community's input in agricultural production is large, but the preservation technology and agricultural products processing technology are relatively low. The huge losses in this area are often overlooked. From the whole industrial chain of agriculture, if we can pay more attention to production, circulation, processing and

one-stop industrial chain management, we can make fresh fruits and vegetables loss rate reduce by half in circulation, which at least will reduce the cost of agricultural products more than 20%.

6 Conclusion

Currently, our agricultural third party logistics research is emerging, especially for the logistics of fresh produce. Many experts proposed development mode and planning of agricultural logistics, but it still in theory stage. For logistics, the practice is ahead of the theory. We must take measures, based on the actual situation of agriculture logistics in our province, from the institution, technical, economic and other angles, to radically accelerate the development of agricultural logistics in Henan Province.

Acknowledgement. This paper is supported by PhD Funds of Henan Polytechnic University (No: B2009-60), here with heartfelt thanks.

References

1. Zhao, L., Xu, Q.: The Exploration to Development of China's Regional Modern Agricultural Logistics System. China Agricultural University (Social Sciences) (3) (2003)
2. Wang, X.: Chinese Rural Logistics Development Mode, vol. (6). Northwest A & F University Press (2003)
3. Ding, J.: The Development of Modern Agriculture Logistics. Economic Daily (4) (2007)
4. Cao, P.: Agricultural Logistics. Logistics Technology and Application (2) (1997)
5. Zhang, S.: China Modern Logistics. China Material Press, Beijing (1998)
6. China Federation of Logistics and Purchasing. China Logistics Development Report (2001). China Material Press, Beijing (2002)
7. Zhang, Z.: Agricultural Supply Chain. Logistics Technology (3) (2003)
8. Wang, Z.: Modern Logistics, 2nd edn. Capital University of Economics and Business Press, Beijing (2008)
9. Hui, C., Zhang, B.: Agricultural Logistics Concept Analysis. Economic Forum (2) (2008)

The Research on Torsional Vibration Signal Processing and Automobile Transmission Shaft

Liu Xiaoqun and Chen Hu

Jiangxi Polytechnic College, Pingxiang, P.R. China
liu_545@126.com

Abstract. Micro-hobbing technology is a good way to process micro gear. The key of the technology is focused on two sides, one is the design of micro gear hob and the other is micro-hobbing technology. It has some advantages including the high efficiency, accuracy in tooth profile, lower cost and widely range of suitable materials. As an extension of traditional technology in small scale, micro-hobbing technology has significance value to research and can be applied in manufacturing the micro gear.

Keywords: steel, concrete, composite shaft lining, mechanical properties, model test.

1 Introduction

Along with the continuous development of the coal resources in mid-eastern China, the thickness of the top soil penetrated by newly constructed drilling shafts has become noticeably thicker. As WANFU mines of JUYE coal field in Shandong and Kouzi West mines of Zhangou coal field in Anhui and Huagou West mine in Huaibei that are about to be developed cross the tertiary and quaternary system alluvium thickness up to 600 ~ 800m. In the future, mines of more deep top soil will be developed. Wells are constructed in so deep alluvium whether by freezing method or drilling method [1,2], they all must adopt high-strength shaft lining structure to resist strong outside loads [3].

The structure form adopted by high-strength shaft lining in our coal mines basically has shaft lining of high strength reinforced concrete, shaft lining of high strength steel fiber and reinforced concrete, composite shaft lining of inner steel plate concrete and composite shaft lining of double steel concrete etc. In foreign forms of high-strength shaft lining structure besides the foregoing, ever use composite shaft lining of cast steel tubing, section steel circle and concrete and so on[1-3].

At present, the majority of the automobiles adopt the reciprocating piston engines widely. Because of the reciprocating force of inertia caused by piston and connecting rod, the inertia centrifugal force aroused by the unbalance mass of the rotating parts and the gas fluctuation pressure in the cylinder, the torsional vibration of the shaft is unavoidable when rotating.

2 Model Design and Test Method

It is not practical to conduct experiments with full-sized prototypes of a high-strength composite shaft lining of double steel and concrete because of the great strength and large size of the structures, hence the model test of reducing scale is used[4].

According to the testing purpose, the model test not only make clear the stress distribution of shaft lining section during loading process, but also need to measure the failure load of shaft lining model. So, the designed shaft lining model not only satisfies the similar conditions of stress and deformation, but also meets the similar conditions of shaft lining strength. On the basis of the basic principle of a similarity theory [5], the available similarity constant was obtained as follows Eq.1 and Eq.2.

$$C_E = C_\sigma = C_P = C_R = 1 \tag{1}$$

$$C_v = C_\mu = C_l = C_\delta = C_\epsilon = 1 \tag{2}$$

Table 1. Model design of shaft lining test

| Specimen numbers | H (mm) | R (MPa) | Thickness of steel plate(mm) | |
|------------------|--------|---------|------------------------------|-------|
| | | | Inner | Outer |
| Z—1 | 97.5 | C65 | 2.0 | 1.4 |
| Z—2 | 97.5 | C70 | 2.8 | 1.7 |
| Z—3 | 97.5 | C75 | 3.5 | 2.0 |
| Z—4 | 102.6 | C65 | 2.8 | 1.7 |
| Z—5 | 102.6 | C70 | 3.5 | 2.0 |
| Z—6 | 104.4 | C75 | 2.0 | 1.4 |
| Z—7 | 107.5 | C65 | 3.5 | 2.0 |
| Z—8 | 107.5 | C70 | 2.0 | 1.4 |
| Z—9 | 107.5 | C75 | 2.8 | 1.7 |

Table 2. Main parameters of the shaft

| Parameters | Value |
|----------------------|-------|
| Module | 0.175 |
| Number of tooth | 12 |
| Variable coefficient | 0.4 |
| Outer diameter | 2.59 |
| Shaft diameter | 0.5 |
| thickness | 0.8 |

Where, C_E is similar constant of elastic modulus, C_σ is similar constant of stress and C_P is similar constant of load. C_ν is similar constant of Poisson ratio, C_R is similar constant of intensity, and C_μ is similar constant of steel ratio. C_δ is similar constant of displacement, C_l is similar constant of geometry and C_ε is similar constant of strain.

In this condition, simply make sure appropriate geometric similar constants. In order to make results comprehensive, There is no test to some specific shaft tube as simulate the object, but consider the range of applications of the ratio of shaft lining thickness h to inner radius (λ , the diameter-thickness ratio) that is a dimensionless quantity and whose similar constant (C_λ) is 1. According to design parameters of shaft lining in the especial deep alluvium, we may take thin, middling and thick three thickness as the simulated prototype. With the test loading devices size, the parameters of model specimen of a shaft lining designed are shown in Table 1. Among them, the diameter and height of model specimen is 925.0 and 562.5 mm respectively, and the inner steel tube and concrete layer are connected with the "U" type anchor card, which is made of 2.5 mm thick plate bar, arranged 3 by per square meter of steel area. The model specimen of a shaft lining is shown in Fig.1. Compressive strength of concrete cube is designed as $R = 65 \sim 75\text{MPa}$.

Pouring model specimen adopts special processing mould. In order to ensure that the boundary conditions of both upper and lower surfaces of a shaft lining model are similar and sealed, the specimen poured need to be conserved for a period of time. From then on, at first put it on the lathe and grinder to process precisely both upper and lower surfaces to obtain high clarity. Then during the test, two rubber seals are installed on upper and lower surfaces of a shaft lining model, and its deformation can eliminate the friction of end faces, so that it ensures that the shaft lining specimen can slide freely and seal in radial direction. The loading test is conducted in a loading device designed specially for high strength shaft lining. Hydraulic oil is used in the tests to simulate the horizontal load of the shaft lining, the vertical load constrained by the cover board and bolt. Because the stiffness of cover and bolt is larger, a shaft lining model is basically a plane strain state during the loading process.

3 Test results and Analysis

Characteristics of the Shaft Lining Deformation. Shaft lining specimen is measured in the loading process on the strain curve of the inner and outer steel and concrete shown in Fig.1 and Fig.2.

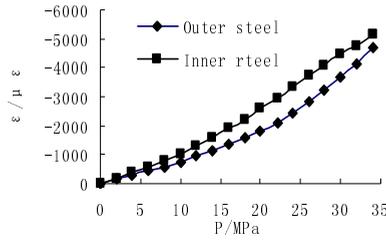


Fig. 1. Curves of load and circular strain of steel plate of Z-7

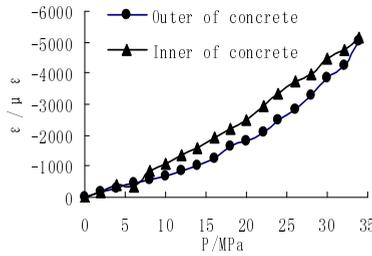


Fig. 2. Curve of load and circular strain of concrete of Z-7

As shown here, the whole stressed process of high-strength composite shaft lining of double steel and concrete can be divided into two stages: during lower load, the relationship of annular strain of steel and concrete and load is approximate linear, which is considered as elastic stage of the shaft lining, when the elastic limit load of the shaft lining is about 40% of the failure load. When the outer load exceeds 40% than the failure load, the relationship of annular strain and load becomes from straight line for curve. With the increasing load, the change of strain accelerates gradually, which is considered as plastic stage of the shaft lining.

The Section Stress of Shaft Lining. Now take the test results of the resistance strain gauge to process, when the relationship of model test loading and the hoop stress of steel and concrete is shown in Figure 3 and 4. The following figure shows that when the load is lesser, the shaft lining stays in the elastic stage. The hoop stress of inner and outer steel and inner and outer edges of concrete is distributed approximately according to the ratio of elastic modulus of steel and concrete. Its $P-\sigma$ relationship which is approximating linear may be calculated by the elastic combination cylinder method. At this time, the hoop stress of shaft lining section presents the distribution of large inner edge and small outer edge; when the load is larger, the material is in the plastic stage and the stress of shaft lining section presents redistribution. At that moment, the hoop stress of inner edge grows slowly; however, the hoop stress of outer edge grows quickly significantly. When the shaft lining is close to destroy, the loop stress of inner and outer edges converge and stress of shaft lining section appears uniform distribution[6-7].

The Maximum Strength of Shaft Lining. The maximum strength values of shaft lining model are listed in table 3 by processing the load test.

As the table 3 shows, high strength composite drilling shaft lining of double steel and concrete in the lateral loads has the very high maximum strength. Therefore, according to the test results, by the limit equilibrium method, this paper deduces the gain factor of high strength composite drilling shaft lining of double steel and concrete, so as to provide a basis for strength calculation of the shaft lining structure [8-10].

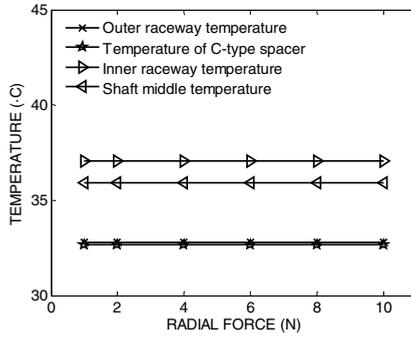


Fig. 3. Effect of radial load upon temperature rise

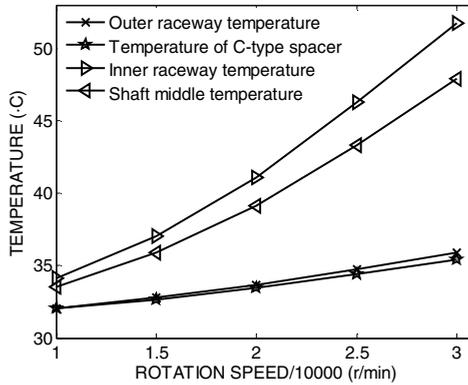


Fig. 4. Effect of rotation speed upon temperature rise

Test results show that shaft lining of double steel and concrete stays in limit state, when inner and outer steel and concrete entrance the plastic stage and section stress distributes in a approximate stage of rectangle. Suppose σ_s is the yield strength of steel, according to limit equilibrium conditions of shaft lining structure of double steel and concrete, Eq.3 was obtained as follows.

$$Pb = \sigma_s h_g + \sigma_{ht} h_h \tag{3}$$

Where, b is outside radius of shaft lining, h_g is total thickness of steel, and σ_{ht} is average hoop stress of concrete in limit state.

It can be calculated for the average hoop stress σ_{ht} of concrete in limit state from Eq.3, thus the gain factor of concrete strength in shaft lining structure was obtained as follows.

$$K = \sigma_{ht} / f_c \quad (4)$$

Where, f_c is the axial compressive strength of concrete, K is the gain factor of concrete strength.

According to the Eq.3 and Eq.4, the gain factor of concrete strength is shown in Table 3.

Table 3. Test results of shaft lining bearing capacity

| specimen numbers | diameter-thickness ratio | concrete strength (MPa) | ultimate strength (MPa) |
|------------------|--------------------------|-------------------------|-------------------------|
| Z—1 | 0.2682 | 67.9 | 28.5 |
| Z—2 | 0.2682 | 70.1 | 32.0 |
| Z—3 | 0.2682 | 76.6 | 36.0 |
| Z—4 | 0.2876 | 66.8 | 34.5 |
| Z—5 | 0.2876 | 73.4 | 39.5 |
| Z—6 | 0.2876 | 75.5 | 36.5 |
| Z—7 | 0.3066 | 64.8 | 35.3 |
| Z—8 | 0.3066 | 72.5 | 34.2 |
| Z—9 | 0.3066 | 77.8 | 41.1 |

4 Conclusions

By the above test study and analysis on high strength composite shaft wall of steel and concrete, the following conclusions are obtained:

Because of the result of the combination of inner and outer steel tube and middle concrete, their own performance was improved. In a shaft lining, the stability of steel tube is enhanced, the brittleness of high strength concrete has been improved also and the destroying shaft lining structure shows good ductility feature. The limit loop strain of steel is about 5000~7000 $\mu\epsilon$ and the limit loop strain of concrete is about 3000~5000 $\mu\epsilon$. Owing to the restraining of double steel, the concrete of shaft lining is completely in the triaxial compressive stress state, and the compressive strength has a greater degree of improvement. The test results show that when the shaft lining is destroyed, the loop stress in concrete can achieve 1.30 ~ 1.90 times of the axial compressive strength. Therefore, during the strength calculation of high strength composite shaft lining of steel and concrete, the gain factor of compressive strength of the shaft lining structure concrete should be considered.

References

1. Cui, Y.L.: A concise handbook of mine construction engineering. China Coal Industry Press, China (2003)
2. Zhang, R.L., He, G.L., Li, D.: Designing guidebook of coal mining. China Coal Industry Press, China (2002)
3. Yao, Z.S., Chang, H., Yang, J.J.: Journal of China Coal Society 29, 167 (2004)
4. Guo, Z.H.: Principle and application of the concrete strength and constitutive relationship. China Architecture and Buildings Press, China (2004)
5. Wei, X.X., Lai, Y.M.: Theory and application of similar method. Lanzhou University Publishing House, China (2001)
6. Tang, Y., Tan, W.K., Fuh, J.Y.H., et al.: Micro-mould fabrication for a micro-gear via vacuum casting. Journal of Materials Processing Technology, 334–339 (2007)
7. Chen, W., Ding, T., Zhang, W.: Research on Fabricate EDM of the Module 0.06 Micro Gear. Micro Fabrication Technology, 49–54 (1998) (in Chinese)
8. Liu, Y., Shi, W., Zhang, Y., et al.: Analysis on qualities of bio-etching processing micro-gear part. In: 2010 International Conference on Mechanic Automation and Control Engineering, MACE 2010, pp. 5792–5794 (2010)
9. Shi, W., Liu, Y.: Fabrication of micro gear based on micro-milling. In: 2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering, CMCE 2010, pp. 608–611 (2010)
10. Shi, W., Wang, X., Liu, Y., et al.: Design of Micro Involute Gear milling Cutter based of AutoCAD. Mechanical Science and Technology for Aerospace Engineering 28(12), 1616–1620 (2009) (in Chinese)

Pattern Analysis of the CRM Application in the China's Telecommunication Industry

Li Wen¹, Wu Meng-yun², and Du Jian-guo¹

¹ School of Business Administration, Jiangsu University, P.R.China, 212013

² School of Finance & Economics, Jiangsu University, P.R.China, 212013
wen_3344@21cn.com

Abstract. As the telecommunications market matures, the demand for telecommunications services to the users is fast-growing, and the telecom operators are rapidly building to change the focus from the network to the business operations, business development strategy from the past to focus on product, now "customer needs first." Against this background, a new management philosophy based on information technologies — customer relationship management (CRM) more and more important. Based on the research and analysis of large amounts of data, combined with the actual situation of China's telecommunications industry, and based on theoretical knowledge of CRM, discussed the application and the problems of CRM in the telecommunications industry, and made a series of measures and proposals, and we will tap the great potential of the telecommunications industry relying on the CRM.

Keywords: CRM, the telecommunication industry, Customer satisfaction, Customer loyalty.

1 Introduction

Since the restructuring of telecommunications in 1999, China's telecommunications service market has gradually developed a new pattern in competition: several major operators led, a number of smaller operators participate and new carriers enter. Meanwhile, some world-class telecommunications giant also began to gradually penetrate into China telecommunication market through various means. In May 2008, after China's telecom industry restructured again, the structure of competition in the telecommunications industry has further optimized. With the development of the China's rapid economic in recent years, both domestic and international situation pose a severe challenge on the public telecom companies' service content, service mode, service quality, management and service awareness in order to meet the international competition of the telecom operation [1]. It request the domestic telecom companies look to the line of foreign advanced telecom companies in the management and operation. In order to promote the exaltation of their own competitiveness, it is combined the traditional marketing concept and innovated CRM knowledge system, referenced the theory and experience of the domestic and foreign customer relationship management, analyzed the status and problems which was implemented by telecom

CRM, proposed the targeted suggestions and measures, provided the practice reference for the telecom to establish the customer relationship management in the paper.

2 CRM Overview of Relevant Theories

2.1 The Definition of the CRM and Its Core

CRM is the customer relationship management. CRM is a system for managing the activities of bilateral contacts which established between enterprise and customer, through managing customer information resources effectively, analysis the customers' demanding characteristics, continue to find the value of customers, to provide customers with products and services, from every place where can contact with customers, establish a long-term, stable, mutual trust relationship between enterprise and customers, in order to stabilize old customers and attract new customers, obtain the excess profits and improve the competitiveness of enterprises through achieving the customers' effectiveness.

Customer relationship management is the core of value creation. Even the best the product manufactured by the enterprise, if it is not the customers need, it has no exist value. So, every enterprise need offer products and services to the potential customers and their requirements, improve customer satisfaction and customer loyalty, achieve a double win situation between enterprise and customer [2].

3 The Relationship among Relationship Marketing, Customer Marketing, Customer Management

Relationship Marketing (referred to as RM) is the relationship of strategic management, with all relevant stakeholders, to achieve long-term shareholder value. The key task are consist of determining the form of the relationship between the different stakeholders and market segments and sub-groups and optimize the management of the interaction in the network of stakeholders.

Customer relationship management is a cross-functional strategic to improve shareholder value. By establishing an appropriate relationship, client and customers are the key, it usually includes the determination of appropriate business and customer strategy, acquisition and dissemination of customer knowledge, determine the appropriate intensity, manage the customer value together create, channel development strategies and integrated data and intelligent use of technology solutions, to create superior customer experiences.

CRM (Customer Management) is concerned that the implementation of tactical CRM[3], this involves customer interaction, including the use of tools such as event management, sales force automation, Web-based personalization and call center management. The relationship of the three parts above is shown in Fig.1.

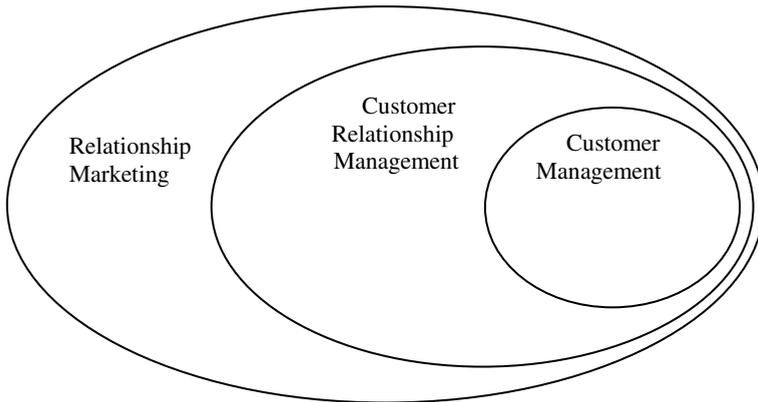


Fig. 1. Relationship among Relationship Marketing, Customer Relationship Management and Customer Management

4 The Application Status of CRM in China's Telecommunications Industry

4.1 Industrial Background

In terms of the whole telecom industry, from 2006 to March 2009, the national telecom business volume has been the raise basically. Main business income of the telecommunications industry has maintained steady and rapid growth, but from 2009, its growth rate gradually becomes slow.

In 2008, China's telecommunications industry was restructured, formed a telecommunications business enterprise which wholly based on China Telecom, China Mobile, China Unicom. According to the 2010 annual statistical bulletin of the National Telecom industry, it is estimated the market of the three telecom operators. The results are shown in Table.1.

Table 1. Estimation of the Three Market Operators in 2008

| Telecom Operators | China Mobile | China Telecom | China Unicom |
|-------------------------------|--------------|---------------|----------------|
| Content Statistics | | | |
| The numer of Mobile users | 482 million | 36.52 million | 153 million |
| Mobile market share (%) | 73.7 | 5.53 | 20.77 |
| Fixed number of users | 18.79 millin | 224 million | 112.5 million |
| Mobile market share (%) | 5.21 | 62.7 | 32.09 |
| The number of broadband users | 6.19 million | 48.28 million | 31.393 million |
| Broadband market share (%) | 6.59 | 56.6 | 36.81 |

Restructuring will ultimately form the three companies to operate a full service, China Mobile will face stronger competition than in the past two ones. By the suppression of asymmetric regulation, dominant position has been weakened. China Telecom and China Unicom mobile's fixed service and net one come into union. It will change the sluggish growth of fixed & net trajectory, and get more space for development[4]. However, the continued growth of China Mobile in mobile voice services is difficult to shake, the high-quality network, high-end users in the cluster and the surplus cash flow are the strong guarantee of the company performance.

Currently, the workflow of domestic telecom companies are implemented organizationally which based on production-oriented system, such systems are production-oriented, enterprise-oriented, and which is implemented in a seller's market circumstances. In the new situation, such systems are not suited to a customer-centric business model and the new competitive market environment (the buyer's market). Customer-centric business model is the symbol of the new generation of business. In today's increasing competitive telecommunications environment, in order to improve the competitiveness of telecom companies, increase profits by change production centers center into customers centers, implementing CRM is an inevitable trend [5].

5 The Application Status Analysis of the Telecommunication Industry

After the telecom reorganization, in telecommunications market, China Mobile, China Telecom, China Unicom include all service of the telecom. China Mobile will undoubtedly walk in the forefront of three major operators. Three operators have transformed on the business operation support system and decision support systems, receive and handle customer inquiries and complaints, market actively, collect the customer data, analyze the data and other services. The construction of these systems is no doubt the further implementation of the telecom operators to lay a good foundation for CRM, major telecom operators as on construction of CRM system as shown in Table.2.

The telecom operators in China have put in use the years of construction, its information system has begun to take shape, including network resource management, boss, internal management and decision support, etc. It has established their own systems accordingly. The implementation of these systems has laid a good foundation

Table 2. CRM System, Major Telecom Operators Current Situation [Year]

| Stage | Infrastructure | Check and adjustment | Application time | Fusion Interactive |
|---|----------------|----------------------|------------------|--------------------|
| China Mobile Communications Corporation | 2001-2003 | 2003-2004 | 2004-2006 | 2005-2006 |
| China Telecommunications Corporation | 2004-2005 | 2004-2006 | 2006-2007 | 2006-2008 |
| China Unicom | 2004-2005 | 2005-2006 | 2006-2007 | 2007-2008 |

for introducing CRM, which part of the system itself is an integral part of CRM. At the same time, telecommunications company in some provinces and cities, with a data warehouse and data mining based on the quasi-functional CRM system module has quietly emerged, such as: the call center in ShangHai is not only the phone center to provide customer complaints and consult services, but also the covers marketing, sales, logistics, customer service, internal management and many other links.

6 CRM Implementation Problems and Its Solutions Problems

For starting lately, domestic telecom operators associated with the CRM philosophy, organizational structure, customer strategy, business processes, information technology planning, performance and other construction applications are still immature. Therefore, telecom operators in the implementation process of CRM have encountered many problems. They are shown as follows:

1. Serious loss of customers. The loss of customers means customer is not going to repeat purchase, or the termination of the original use of the service. In the mobile communications industry, loss has been divided into passive loss and active loss. According to statistics, domestic carriers annual churn rate of 30% or more each year.

2. Customer information scattered distribution. A variety of business-related data in information silos state, which is lack of effective use. First, the enterprises don't have effective means to extraction, processing and integration these data scientific even control a lot of customer information. They can't dig out useful information for business management decision-makers, can not to provide strong support for managers decision-making; Second, part of the information is very important, but it's difficult to collect, the information related to customer privacy and to reflect the tendency of future consumption of users. The information in these areas can help for telecom companies to tap the potential users effective.

3. The potential value of Customers didn't fully be dug out. To our current situation in the telecommunications industry, enterprises can't assessment in an effective way of future income, future costs and cross-selling of customers. In addition, in the telecommunications industry, develop a new customer is costs 5 times than to retain an old customer. Lose of loss a valuable customer need secure 10 new clients to restore. Therefore, to the telecommunications operators, how to improve customer loyalty is still a problem.

7 Solving Strategies

1. Attract potential customers to network, increase the existing customer satisfaction. Reduce the risk of loss of customers, improve customer consumption levels, adequate market share, is a top priority of business decision-making. The method of analysis the loss of customers could be introduced, using SPSS multivariate analysis and data mining analysis to analysis the calls, customer service complaints or payment information over a period of time of the losing customer, establish the model of losing customers, extracted the behavior characteristics from the customers who has been lost or in trend of loss, improve customer retention and customer value.

2. Establish a customer relationship management system with a data warehouse as the core, provide customer information and analysis for marketing. On the one hand, the data warehouse obtain the useful customer information to relevant market from the business, billing, channel, call center and other production management systems, formed a unified date platform operate on market within the enterprise, and provides the information to the people who are doing data analysis and the marketing decision-makers; On the other hand, the data warehouse system based on the needs of customer interaction targeted marketing, service strategy to the corresponding customer after analysis or mining the information what produce by customer abnormal consumption, loss of customers early warning, and other marketing activities feedback to the customer contact system.

3. Add data mining capabilities, using decision tree or clustering method to segment customers. Make the customers into different categories, and then, use different strategies in order. Fig.2 is the customer portfolio analysis chart of the telecommunications industry.

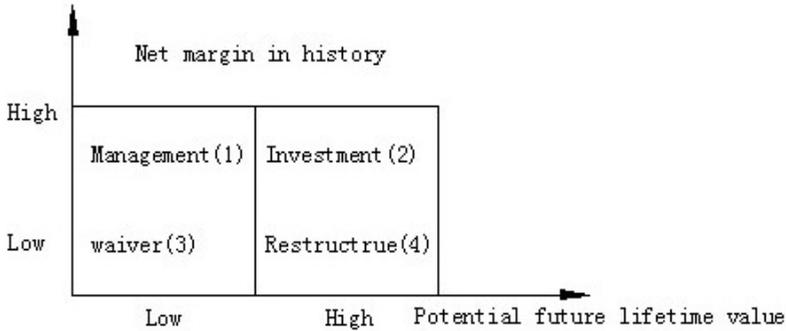


Fig. 2. Customer Portfolio Analysis Chart

From the above Fig.2, you can see for the first class of customers. Although they are profitable, but there is very little earnings in future. This requires developing the relationship with them, increasing their loyalty, exploiting its potential and value of joint consumption, establishing a long-term relationship with them, so as to increasing the potential customer lifetime value constantly; the second type of customer is most concerned by the enterprise, in general, it can be used these measures such as a number of telecommunications services, providing flexible billing date, giving preferential payment terms (such as rebate calls); The third type of customer will not have a obvious impact on enterprise now and future. Therefore, enterprise should concentrate its resources to those customers who can coordinate with the enterprise and have a common development; For the last class of customers, telecom operators should restructure existing client relationships, find cut costs, increase the value methods[6].



8 Summary

Customer relationship management is the management philosophy that relied on "customer satisfaction as the center". Bing a new management philosophy based on information technology, CRM will continue to be a powerful magic of the competition in the telecom operators and moreover. Its importance will become more and more remarkable. With the decrease of technical resistance and the transition of management, it is bound to create significant value for telecom companies continuously.

Acknowledgement. Humanities & Social sciences Project of Ministry of Education of P.R.China(10YJC630281), Zhenjiang Soft Science Funds of Jiangsu Province of P.R.China (RK2010009), Philosophy & Social Science Project of Jiangsu Province of P.R.China (09SJB630014).

References

1. Qian, H.: Telecom Market Space Research on the Customer Value. Beijing University of Post and Telecommunications (10), 27–30 (2007)
2. Ding, X.: Big customer saling strategy and project management, pp. 56–70. China Machine Press, Beijing (2005)
3. Zhuang, C.: The countermeasure and the measure of implementing CRM in the telecommunication enterprise. Enterprise Economy (06), 19–21 (2007)
4. Pennie, F., Adrian, P.: Customer Relationship Management: A Strategic Perspective, vol. (01), pp. 9–10. Gabler Verlag (2009)
5. Min, M., Wu, J., Yang, Y.: The Application Analysis for Client Relationship Management of Telecom Industry Home and Abroad. Designing Techniques of Postes and Telecommunication (12), 79–80 (2006)
6. Zhou, J.: Customer Relationship Management and Value Creation, pp. 8–10. Shanghai Jiaotong University Press, Shanghai (2006)

Service-Oriented Data Integration Using GIS: An Anchorage Management Tool for Wenzhou Government, China

Li Yubao^{*,**}, Li Longchang, and Liao Fei

Wenzhou University
Wenzhou 325035, Zhejiang Province, China
leesilee@163.com

Abstract. GIS technologies can help officers from local governments in developing countries with spatial information. This project, taking Wenzhou Anchorage Management Information System as an example, presents a framework of heterogeneous data integration to serve the officers who are used to get information from paper maps. We employed VB to integrate spatial and nonspatial data into a software tool that allows user to browse and output information of interest. ArcReader, a GIS data viewer of ArcGIS, ESRI, was introduced in the information system for the users to deal with feature layers. It received good evaluation from the users in Wenzhou and promises wider application to more users.

Keywords: Data integration, Service Oriented Architecture, GIS, Management information system.

1 Introduction

Many officers from local functional departments in China use maps frequently. It is common to find maps on the wall of an office opposite to the back of an officer who is in the front of a desktop. Administrations need more spatial information that is hard to get from paper maps. Environmentalists want the exact position, pollutant chemicals and their concentrations of an air pollution source. Transportation officers need the length, width and vehicle density of a road. Departments of agriculture and land resources demand the shape and area of a piece of land and the water, vegetation and other landuse around.

Geographical information system (GIS) can better satisfy the need of officers for spatial information. With GIS, users easily integrate heterogenous information of spatial and attribute features. Spatial features of various GIS data layer provide geographical positions, shapes and topologies of the objects of interest on the earth. GIS tables and annotations contain descriptive, statistical and other non-spatial information. GIS could be a useful tool for officers, especially from local governments in developing countries like China to improve work efficiency.

* Foundation item: Wenzhou anchorage management information system supported by Wenzhou Government.

** Corresponding author.

However, fewer officers from administrative departments expert in GIS, even less can successfully operate professional GIS tools. The information used by different governmental divisions differs greatly. So there are many demands for heterogeneous data integration from various data sources. Although the spatial objects of people’s interest on the earth surface are different, the technologies used for data integration are similar.

The present project is to establish a frame to serve officers with information, elaborating the procedures of a GIS-based heterogeneous data integration with a example of an anchorage management tool for the functional department of Wenzhou, China. Wenzhou is a coastal city where gale, storm and typhoon are main natural hazards. Along the coast line of 355 kilometers there are 42 anchorages harboring 6913 ships. We focus on the needs of officers who have been frustrated with file-folder working style and deals with varied types of data including text file, database, map and so on.

2 Design of Data Integration

GIS is powerful to produce useful information by: (1) data acquisition; (2) data management; (3) spatial analysis; (4) statistical analysis. User’s models can also be incorporated with GIS database. GIS is also powerful in executing data mining and information integration.

Users can easily query information from spatial features and attribute tables. But officers from local government may need more information than what GIS data can provide. And users themselves have accumulated great volume of data such as Word and Excel files over years of working in their positions. So data integration projects should mine new useful information before being assembled with existing data. For the purpose of the users to use spatial and non-spatial data integrated, one computer language (such as C, VB) is necessarily employed to develop a computer information system or a software tool based on Service Oriented Architecture (SOA)(Fig. 1).

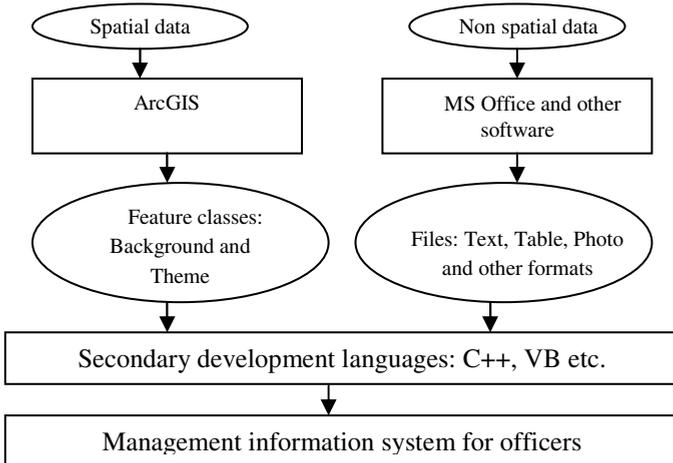


Fig. 1. A diagram showing the SOA framework of data integration



3 An Anchorage Management Tool for Flood Controllers

The users and their needs. The users of the project are from Wenzhou, China, a coastal region where surging storms and typhoons cause heavy losses almost every year. The officers from flood control departments who are in charge of settling fishing ships during the times when storms and typhoons affect the nearby sea. They frequently use the data of each anchorage. For directing a ship to a safe harbor within the shortest time, the officers need to know the distance between the site of the ship and the nearest safe anchorage. Specifically they want this project is able to: (1) save various types of data including text and table, digital photo and map; (2) query, retrieve, print and output the information of interest; (3) map each anchorage with detailed roads, residential areas and waters; (4) develop a computer-based tool for leading and managing ships to the nearest anchorage as sea storms are approaching.

The data needed for the project are heterogeneous. Non-spatial data including files of Microsoft Word and Excel were basically prepared by the Wenzhou Department of Flood Control. This project looked at the preparation of spatial data involved, integration of data in various formats and development of computer software named Wenzhou anchorage management information system (WAMIS).

Data Acquisition and Manipulation. Sources of spatial data used for the project are heterogeneous. Data of residences, transportation lines, waters and other geographical background features were digital linear graphics in Shapefile. Satellite images of the SPOT of 2008 with the resolution of 2.5 meters and thematic maps of Wenzhou Sea were used to extract anchorages, water depths and other coastal and marine information.

All spatial data were manipulated by using ArcGIS 9.3. The spatial area is from to N and from to E as the user required. The raster images and thematic maps were rectified to a local coordinate system as the same with the shapefile layers. The digitalization and edit technologies of ArcGIS help complete feature class layers needed. The feature classes fall into 2 categories: background and theme. Background information includes local residences, transportation, rivers etc.. The thematic layers are Anchorage and Directing Ships that were completed using GIS information mining technology.

Information Mining for Directing Ships. The capability of anchorages to protect ships from typhoon was classified 3 levels, storm (Beaufort Scale number 10), Violent storm (Beaufort Scale number 11) and hurricane (Beaufort Scale number 12) & above as avoiding-wind levels. For each anchorage, area, capacity to anchor ships and anti-wind level were included in the attribute table, and a thematic map in the format of Jpeg and an ArcReader PMF were prepared and outputted for the latter data integration.

We produced a polygon layer to present the information for directing a ship to the nearest anchorage in case the risk of storm. The sea within the area of the project was divided into graticules with the cellsize of 0.1 degrees both in longitude and latitude. For each graticule the distance to the nearest anchorage of 3 avoiding-wind levels each were calculated and stored in the attribute table (Table 1).

Table 1. The Thematic Feature Classes

| Name of layer | Fields of attribute tables |
|-----------------|---|
| Anchorage | Name, Manager’s name, Area, Capacity, Anti-wind level, County name, Township name |
| Directing ships | Nearest anchorage, Nearest anchorage for storm, Nearest anchorage for violent storm, Nearest anchorage for hurricane, |

Data integration. We first examined the total volume of all the data and proposed possible choices of development plan. The users finally decided a single-computer software. Visual Basic computer language was employed to accomplish the integration of heterogeneous data including image, feature class and other formats such as text and numerical table. Feature classes consist of layers of background and thematic information (Table 2), which were integrated and outputted as PMFs for the users using ArcReader to get information. The software developed with VB allows the user to query, browse, export and print the texts, tables and maps.

Table 2. Data Integration Using Arcreader PMF and Feature Classes

| Feature class (layer) | Arcreader pmf: anchorage | Arcreader pmf: directing ships | Feature class (layer) | Arcreader pmf: anchorage | Arcreader pmf: directing ships |
|------------------------------|--------------------------|--------------------------------|-----------------------|--------------------------|--------------------------------|
| Residential points | √ | | Dike lines | √ | |
| Transportation lines | √ | | Sea-depth polygons | | √ |
| Water polygons | √ | √ | Island polygons | √ | √ |
| County and township polygons | √ | √ | Anchor-age polygons | √ | √ |
| Coast lines | √ | √ | Dike lines | √ | |

4 Wenzhou Anchorage Management Information System (WAMIS)

Structure. WAMIS provides a user interface on which seven buttons appear (Figure 2). On the left side is button Regions that actually is the directory of



anchorages the counties they belong. For each anchorage the data were stored separately according to the data type. Text files contain the descriptions of the anchorage. Data tables present statistical and quantitative features. Each anchorage has a thematic map (PMF) showing geographic and geometric information. Photos give details of reality.

The button of Emergency Management links the PMF of Directing Ship that consists of the information indispensable for directing ships to avoid storms and typhoons.

Functions. WAMIS established a directory linking anchorages with the counties where they locate. While selecting a specific anchorage, users can easily read, export and print texts, tables and other non-spatial data of each anchorage and its vicinities. Clicking the button of Maps opens a PMF file with ArcReader that allows users to browse thematic maps of the anchorages. Zoom in/out, pan, distance measurement and other functions are available.

Clicking the button of Emergency Management on the WAMIS interface opens the PMF of Directing Ships. Once having the location (degrees of latitude and longitude) of a reported ship, the officers can immediately find the best shelter anchorage according to the wind force (Figure 3).



Fig. 2. The user surface of the anchorage management system

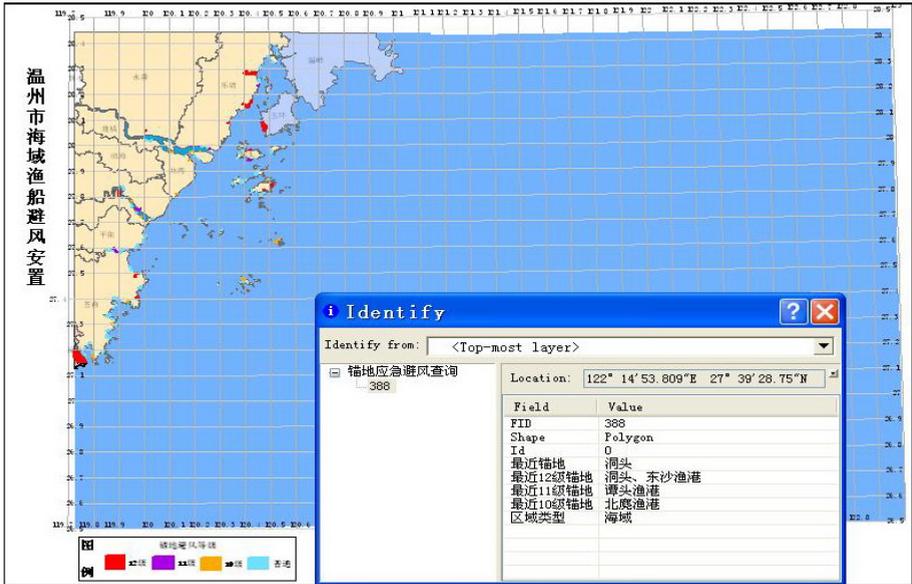


Fig. 3. Directing ships to the nearest anchorage

References

1. Li, D., Wang, S., Li, D.: Spatial Data Mining and Application. Science Press, Beijing (2006) (in Chinese)
2. Gao, L.: Research on Key Technologies of the Present Marine Utilization Data Integration. Nanjing Normal University (2008)
3. Wang, X.G., Yan, Z.L., Shen, W.W.: Research on constructing the Geographic Information System of flood prevention decision support. Science of Surveying and Mapping 4, 41–43 (2010) (in Chinese)
4. Chi, W., Wu, X., Er, D., Wang, M., Wang, Z.: Research on the geographic information data integration technology based on service-oriented architecture. Science of Surveying and Mapping 1, 203–205 (2009) (in Chinese)
5. Breunig, M.: An approach to the integration of spatial data and systems for a 3D geo-information system. Computers & Geosciences 1, 39–48 (1999)

An Index Assignment Algorithm over Noisy Channel

Wang Yue

Institute of Electronic Information, Zhejiang Gongshang University,
Hangzhou, 300018, P.R. China
ywk1298@163.com

Abstract. Enlightened by the defence ability of immune system, an immune colon algorithm (MCIAA) based on the genetic algorithm (GA) is presented. The new algorithm has a good effort to adjust the dynamic balance of colony's convergence and individual's diversity, effectively overcome the problem of traditional GA's local optimal resolution. In this paper, we apply this algorithm on the vector quantization (VQ) index assignment over noisy channel. Simulation result conforms that the algorithm yields a significant reduction in average distortion and converges in less short time compared to other algorithms.

Keywords: vector quantization (VQ), noisy channel, genetic algorithm (GA).

1 Introduction

Vector Quantization(VQ) is an important and successful source coding technique in many digital communication applications. Vector quantization(VQ) operates by mapping a large set of input vectors into a finite set of representative codevectors. The transmitter transmits the index of the nearest codevector to the receiver over noisy channels. However transmitting VQ data over noisy channels changes the encoded information and consequently leads to severe distortions in the reconstructed output. One approach is to optimize the vector quantization(VQ) index assignment to minimize the distortion. Finding the best index assignment requires search among all possible codebook permutations for that which yields the minimum distortion under noisy channels conditions. As there are $N!$ ways to order N numbers, exhaustive search for the optimal ordering is not feasible, many research efforts in the past has been focused on the index assignment problem. R.Zelinski[3] has been shown the effects of transmission error on the PCM signal. Rydbeck and Sundbery[4] proved the importance of good index assignment for scalar quantizers. Zeger and Gersho[1] published the Binary Switch Algorithm(BSA) which operated by continually swapping codewords until there was convergence to a local minimum. In [2] the parallel genetic algorithm was applied to design the codevector indices. Knagenhjem and Agrell [7] utilized the Hadamard transform as a tool for index assignment In [8] an algorithm(EAIAA) was proposed based on the evolution algorithm to assign index assignment over noisy channels. Hamidreza[9] proposed a COVQ for symbol-by-symbol maximum a posteriori(MAP) hard decision demodulated channels.

In this paper, an immune colon algorithm based on the genetic algorithm is proposed, which has a good effort to adjust the dynamic balance of colony's

convergence and individual's diversity, effectively overcomes the problem of traditional GA's local optimal resolution, The feasibility and efficiency of our algorithm is confirmed by experimental results.

2 Vector Quantization over Noisy Channel

Vector quantization encodes each vector from a sequence of source vectors with a channel symbol—a binary word chosen from a finite set[1]. A typical VQ system contains a finite predetermined collection of codevectors (a codebook), and a vector distortion measure, which, when given two vectors, yields a distance (or distortion) between them. A sequence of input vectors is coded by the VQ system by associating each input vector with the binary index of a codevector whose distance from the input vector is minimized. This coding processing can be well done by properly assigning the indices for the codevectors, which is the index mapping π . Then this index is subsequently transmitted to a receiver which decodes the codevector associated with the index (by a lookup operation) and uses the codevector as an approximation to the original input vector to the system. A block diagram depicting a noisy channel vector quantizer is shown in Fig.1.

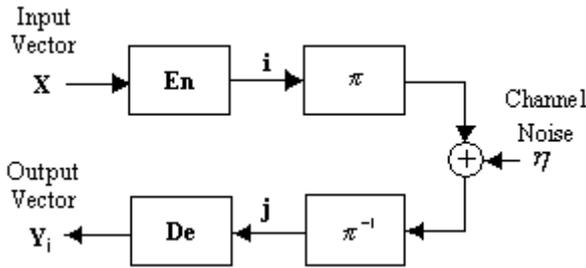


Fig. 1. Block diagram of VQ on discrete memoryless channel

In Fig.1 En is quantizer encoder, De is quantizer decoder, π is index mapping.

A vector quantizer Q is a mapping of input vector X which is k-dimensional Euclidean space R^k to its finite subset Y of R^k given by $Q : X \rightarrow Y$, where $Y = \{y_1, y_2, \dots, y_N | y_i \in R^k\}$. The ordered set Y is known as a codebook and the N elements of Y are called codevectors. The codevector of the codebook Y, $y_i^* = (y_{i1}^*, y_{i2}^*, \dots, y_{ik}^*)$ satisfies:

$$d(X, y_i^*) = \min_{1 \leq j \leq N} (d(X, y_j)) \tag{1}$$

where $d(X, y_i)$ is the distortion between vector X and codevector y_i .

N codevectors $y_i, i = 1, 2, \dots, N$, are assigned codevector indices with a m bit

string $b(y_i)$, where $N = 2^m$. The effects of channel errors on indexes can result in significant distortion in decoded vectors, if the index in the set $\{0,1\}^b$ are transmitted across a noisy channel, their values in general will be received as different index in the same set $\{0,1\}^b$. Therefore, a memoryless noisy channel can be represented by a mapping $\tau : \{0,1\}^b \rightarrow \{0,1\}^b : \tau(i) = i \oplus \eta(i \in \{0,1\}^b)$, η is channel noisy.

For each index $i \in \{0,1\}^b$, π uniquely maps i to another index of $\{0,1\}^b$, namely $\pi(i)$. Hence, a noisy channel vector quantizer Q_π is a mapping from R^k to the set of codevectors Y given by

$$Q_\pi = D e_{\pi^{-1} \circ \tau \circ \pi} E n \tag{2}$$

For a fixed vector quantizer Q and a given channel, one seeks to minimize the average distortion given by

$$e_\pi = E[d(X, Q_\pi(X))] = E[d(X, Y_i)] \tag{3}$$

Therefore, it means to minimize the average distortion over all permutation functions. This procedure needs to jointly consider the statistical characteristic of source and channel. We consider metric distortion functions and r th-power of a norm distortion functions, the latter being of the form

$$d(X, Y) = \|X - Y\|^r \tag{4}$$

we then have

$$e_\pi = E[\|X - Y_j\|^r] = E[\|(X - Y_i) + (Y_i - Y_j)\|^r] \leq U_\pi$$

$$U_\pi = \{ (E[\|X - Y_i\|^r])^{1/r} + (E[\|Y_i - Y_j\|^r])^{1/r} \}^r \tag{5}$$

From above, we can see e_π contains two components, the first part is independent of the permutation π , the second part varies depending on the choice of π , to minimize the upper bounds, it suffices to minimize the quantity [1]

$$D_\pi = E[d(Y_i, Y_j)] \tag{6}$$

In memoryless binary symmetric channel (BSC), we could have

$$D_\pi = \sum_{k=0}^{N-1} \sum_{m=0}^b f_p(y_k) q_m \sum_{w \in N^m(\pi(k))} d(y_k, y_{\pi^{-1}(w)}) \tag{7}$$

where N is codebook size, $N = 2^b$, $f_p(y_k), (0 \leq k \leq N - 1)$ is codevector probability of the input codebook Y , which gives the probability that a particular codevector is selected by the encoder to represent an analog input vector; $q_m = \mathcal{E}^m (1 - \mathcal{E})^{b-m}$, where \mathcal{E} is channel error probability; $d(\cdot, \cdot)$ is defined as a distortion function, which is the Euclidean Distance in our case.



3 Immune System and Relational Algorithms

(1) Immune system

Enlightened by the biological system, many algorithms were presented using for engineering calculation. Immune system is our basic defense system against bacteria, viruses and other disease-causing organisms. It has powerful and complex mechanisms that recombine the gene to cope with the invading antigens, producing the antibodies and exclude the antigens.

(2) The colon selection principle

When an animal is exposed to an antigen, B cell of the Lymphocytes responds by producing antibodies [5]. Each cell secretes only one kind of antibody, which is relatively specific for the antigen. By binding to these antibodies, and with a second signal from T cell, the antigen stimulates the B cell to proliferate and mature into terminal antibody secreting cells called plasma cells. in addition to proliferating and differentiating into plasma cells, it also can differentiate into long-lived B memory cells. Memory cells circulate through the blood, lymph and tissues, and when exposed to a second antigenic stimulus commence to differentiate into large lymphocytes capable of producing high affinity antibodies. The idea of colon selection is that only those cells that recognize the antigen are selected to proliferate.

Based on the colon selection principle, De Castro[5] proposed colon selection algorithm, which can perform machine learning, pattern recognition and optimization problem.

(3) The immune algorithm

In traditional genetic algorithm (GA), the big problem is its early convergence and local convergence. Many algorithms were presented to solve this problem. Jang_Sung Chun[6] enlightened by the variation of the biological immune system, proposed immune algorithm to encourage the variation of the colony, let some individuals can overcome local optimal resolution, and get global optimal resolution. The algorithm used the concept of information entropy to guarantee the diversity.

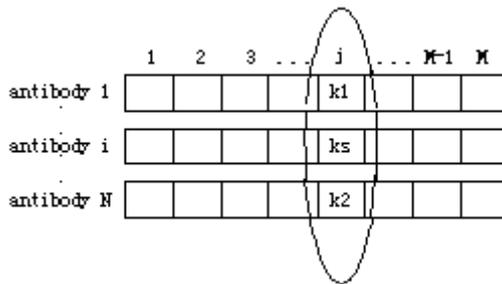


Fig. 2. Information entropy of genes



As in Fig 2, let the immune system be composed of N antibodies having M genes, the information entropy $H_j(N)$ of j th gene is $H_j(N) = -\sum_{i=1}^N P_{ij} \log P_{ij}$ where P_{ij} is the probability that i th allele comes out at the j th gene. For example, if all alleles at j th genes are the same, $H_j(N)$ is equal to zero. Therefore, the average information entropy is

$$H(N) = \frac{1}{M} \sum_{j=1}^M H_j(N) \tag{8}$$

We can use information entropy to present the affinity between two antibodies:

$$ay_{vw} = \frac{1}{1 + H(2)} \tag{9}$$

Therefore we can use (9) to exclude individuals whose affinity is high to guarantee the variation of the colony.

4 Colon Immune Algorithm for Index Assignment over Noisy Channel

One of the goals of the GA is the convergence of the colony, but to get the global optimal resolution, it must keep individuals' variety as well; this needs to adjust balance between convergence and variety. The colon selection algorithm can accelerate the speed of algorithm' convergence, but could not ensure the variety of individuals, so we introduce the concept of information entropy to keep the variety of individuals, which corresponds to the global optimum.

The colon immune algorithm is based on the genetic algorithm. We adopt the VQ codebook designed by LBG algorithm as the first generation of our algorithm. The algorithm eventually halts and gives as output an assignment of binary indexes to the vectors of the codebook. The algorithm thus receives a codebook as input, and outputs the same codebook but with better index assignment.

(1) Creation of Populations

In our algorithm, it is the codevector index permutation that is to be optimized. The genes of an individual are the indices of codevectors, so each individual of colony is a long ordered chain of indices corresponding to codevectors[8]. According to an individual's representation, all the genes of the individual should be different from each other and cover all the indexes completely.

(2) Mutation Probability P_m and Mutation Mechanism

The probability of mutation and exchange rate determine the occurrence frequency of mutation. if probability of mutation is too high, algorithm will search in a large-scale solution space in favor of searching global resolution, but because of always in a big change of individual structure, it leads to the difficulty of convergence. Otherwise if probability of mutation is too low, it will slower the searching speed, usually lead to local optimal resolution. With this consideration, we design higher probability in the

early stage of iteration to search for good individual, then with the increase of the number of iteration, probability begin to decrease, then at the anaphase of the iteration, probability can decrease according Gaussian function to make a fine tuning.

$$P_m = \begin{cases} 0.07 & 0 \leq t \leq T/3 \\ 0.05 & T/3 < t \leq T/2 \\ 0.04 & T/2 < t \leq 2T/3 \\ 0.035 & 2T/3 < t \leq 5T/6 \\ 0.8 * \frac{1}{\sqrt{2\pi} \text{sharp_improved}} e^{\frac{-((t+150)/(T+120))^2}{2(\text{sharp_improved})^2}} & 5T/6 < t \leq T \end{cases}$$

$$\text{sharp_improved} = 0.5 * \frac{1}{\sqrt{2\pi} * 0.5} e^{\frac{-(1/3)^2}{2 * 0.5^2}} \tag{10}$$

In conventional GA, the mutation is usually to add a random number to each gene to be mutated to produce a new individual, but in our case, it's likely to result in an invalid individual with different genes having the same index by such mutation method, therefore, we change the mutation mechanism by first generate a random number in range (0,1) for each gene of an individual to be processed, if the random number is less than the probability of mutation calculated by equation (10), mutation is conducted on this gene by exchanging with the gene at location s ($0 \leq s \leq N - 1$), which is a uniform random number generated randomly. It is shown in Fig 3.

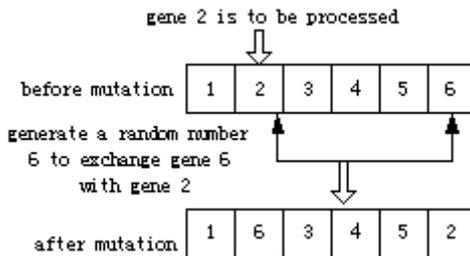


Fig. 3. Mutation mechanism

(3) Algorithm Description

Step 1: Initialization. According to the index assignment of the input codebook, an initial individual $G_0(0)$ is produced.

Step 2: Generation of initial colony. The first generation,

$G_1(0), G_2(0), \dots, G_N(0)$, is produced by mutating $G_0(0)$ with our proposed mutation mechanism. Calculate the fitness value of each individual according equation

(7), $D_0(t), D_1(t), \dots, D_N(t)$, Sort $D_i(t)$ in increasing order, then select and reserve n individuals with smaller $D_i(t)$ to memory pool.

Step 3: Clone individual. Select n individuals from memory pool to be the new colony and clone them respectively, the smaller the individual's fitness value, the more clonal individuals.

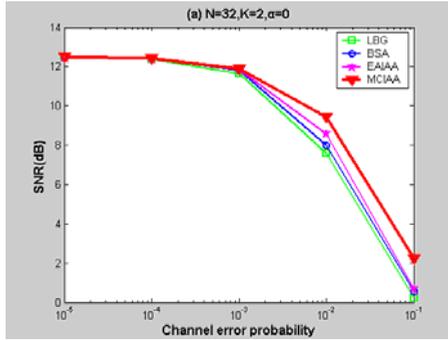
Step 4: Mutation and Sorting. Apply this mutation operation all clonal individuals. Calculate the fitness value of the colony and all clonal individual and sort them in increasing order.

Step 5: Selection. If the generation number is the multiples of 10, according to equation (9), calculate each individual's affinities with other individuals, which can be the standard of similarity. Set a threshold Tr , compare each individual's affinity with other individuals, keep one individual with lowest fitness among the individuals whose affinities with this individual are higher than threshold, and exclude other individuals from this colony since their affinities with this individual is higher than the threshold. Then sort the colony in decreasing order, and put the n individuals with smaller $D_i(t)$ to memory pool. if the generation number is not the multiples of 10, put the n individuals with smaller $D_i(t)$ to memory pool directly.

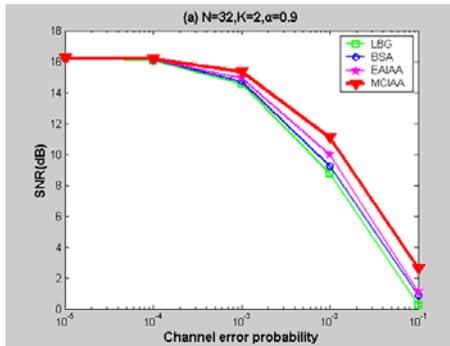
Step 6: Iteration. Set $t + 1 \rightarrow t$, and repeat Step 3 through Step 5 for all the training vectors, until a terminating condition is satisfied, i.e., the total number of sweeps T reaches a predetermined value. When the process terminates, assign indices according to the individual with the smallest D to produce a new codebook with optimal index assignment.

5 Experiment Result

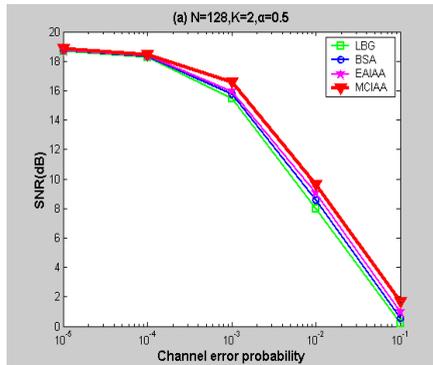
In our algorithm, The input signal tested is 1st-order Gauss-Markov processes $x_n = \alpha x_{n-1} + w_n$ where the correlation coefficient α has magnitude less than one, and w_n is a zero mean, unit variance, i.i.d. Gaussian source, the values for α in our experiment are 0, 0.5 and 0.9. Each of these training sequences is tested with quantizers having the following codebook size $M=16, 32, 64, 256$ and vector dimension $k = 2$ and 4 . The total number of sweeps T is 180. The threshold Tr is 0.95. the number of memory pool is 5, the clonal number of individuals is 5, 4, 3, 2, 1 respectively with the increasing of the fitness of individuals in memory pool. The SNRs are calculated for VQ systems using LBG nature index permutation, BSC iteration index assignment, EAIAA and our algorithm (MCIAA). a comparison is made between them. Ten runs are conducted for each data set and the average SNR value is calculated. The various SNRs with $(M, k) = (32, 2), (128, 2)$ and $(256, 4)$ are plotted in Fig.4. It is easy to find that our MCIAA in SNR is higher than other algorithms and is more prominent for higher channel error probabilities. Furthermore, the other merit of our MCIAA is the speed of convergence, the iteration number of EAIAA is about 300, our algorithm only need 100 to nearly obtain global optimal resolution.



(a) $M = 32, k = 2, \alpha = 0$



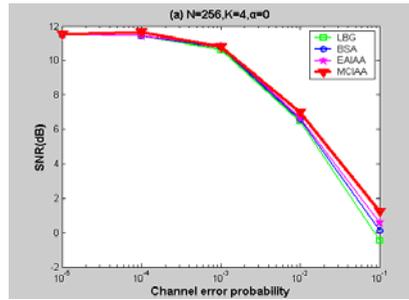
(b) $M = 32, k = 2, \alpha = 0.9$



(c) $M = 128, k = 2, \alpha = 0.5$

Fig. 4. Various SNRs with (32,2),(128,2)and (256,4)



(d) $M = 256, k = 4, \alpha = 0$ **Fig. 4.** (Continued)

6 Summary

According to the biological immune system, we propose an immune colon algorithm and apply it to the index assignment over noisy channel; it provides efficient means of obtaining globally optimal index permutation, leads to increased performance in the noisy channel. Simulation result conforms that the algorithm yields a significant reduction in average distortion and converges in less short time compared with other algorithms. This algorithm also can be applied to pattern recognition, combinational optimization problems.

References

1. Zeger, K., Gersho, A.: Pseudo Gray coding. *IEEE Trans. Commun.* 38(12), 2147–2158 (1990)
2. Pan, J.S., McInnes, F.R., Jack, M.A.: Application of parallel genetic algorithm and property of multiple global optimal to VQ codevector index assignment for noisy channels. *Electronics Letters* 32(4), 296–297 (1996)
3. Zelinski, R.: Effects of transmission error on the mean-square error performance of transform coding system. *IEEE Trans. Acoust, Speech, Signal Processing (ASSP-27)*, 531–537 (1976)
4. Rydbeck, N., Sundberg, C.W.: Analysis of digital errors in nonlinear PCM. *IEEE Trans. Commun. (COM-24)*, 59–65 (1976)
5. de Castro, L.N., Von Zuben, F.J.: Learning and optimization using the clonal selection principle. *IEEE Trans. on Evolutionary Computation* 6(3), 239–251 (2002)
6. Chun, J.-S., Kin, M.-K., Jung, H.-K.: Shape optimization of electromagnetic devices using immune algorithm. *IEEE Trans. on Magnetics* 32(2), 1876–1879 (1997)
7. Knagenhjekm, P., Agrell, E.: The Hadamard transform—A tool for index assignment. *IEEE Trans. Inform. Theory* 42, 1139–1151 (1996)
8. Litianhao, Yusongyu, Zhanggao: An vector quantizer index assignment. *The Journal of Electronical* 30(6), 876–879 (2002)
9. Saffar, H.E., Alajaji, F.: COVQ for MAP Hard-decision Demodulated Channels. *IEEE Communication Letters* 13(1) (January 2009)

Application of Artificial Neural Network (ANN) for Prediction of Power Load

Fei Xia¹ and Li Fan²

¹ School of Information Engineering, Nanchang Institute of Technology,
Nanchang, Jiangxi, China

² School of Mechanical and Electrical Engineering, Nanchang Institute of Technology,
Nanchang, Jiangxi, China
Feauser@yahoo.cn

Abstract. This paper focuses on the development of an artificial neural network based on power load in power engineering. The front section represents areas in power systems that artificial intelligence has been applied to. And then overviews the artificial intelligence techniques which have been used and makes suggestions for the improvement of existing artificial intelligence tools. Following this, the paper concentrates on neural networks and their applications to power systems. The multi-layer feed forward network is introduced and the problems in establishing neural network approaches based on this network for power system applications are discussed. Further subjects for further research in artificial intelligence and neural network applications in power systems are presented. This paper focuses on the development of an artificial neural network based on power load.

Keywords: Artificial Neural Network (ANN), Power Load Prediction, Make Model.

1 Introduction

Modern power supply systems are needed to generate and supply high quality electric energy to customers. To reach this requirement, computers are used in power system planning, monitoring and control. Power system application programs for analyzing system results are stored in computers. [1]In the designing stage of a power system, system analysis programs are executed repeatedly. Engineers adjust and modify the input data to these programs according to their experience and professional knowledge about the system until satisfaction plans are determined. However, the programs so developed for power system analysis and planning are based on mathematical models and are implemented using languages which are suitable for numerical computation only. For Complicated approaches to system planning, development of methods and techniques to incorporate practical knowledge of planning engineers into programs which also include the numerical analysis programs are needed.

In the field of power system monitoring and control, computer based Energy Management Systems are now widely used in energy control areas. Power system analysis programs and other application programs are used in Energy Management

Systems for the purposes of investigating and predicting the appearance of power systems under steady-state operations. Though these programs are powerful tools, their ability to help operation engineers to make efficient decisions is very limited when unforeseen or unexpected models of system operation occur. The abnormal pattern of system operation may be caused by network faults, active and reactive power imbalances, or frequency deviations. An unpredicted Operation may lead to a failure or a complete system blackout. Under these emergency situations, power systems are restored back to the normal state according to decisions made by experienced operation engineers. For efficient diagnosis of network faults, determination of operational strategies for network restoration, and balancing active and reactive powers, there is clearly a need to develop new computer techniques and methods to build programs where the precious knowledge of experienced operation engineers can be accounted for in addition to the conventional power system application programs. There is need to build a rapid and effective method to predict the abnormal state power system .

Artificial intelligence (AI) provides people with coding and reasoning technology and descriptive knowledge. The incoming neural networks (NNS), in addition, provides neural network modules which can be executed in an online environment.[2] These new technologies make up the traditional calculation method and technique in power systems used to solve design, prediction, operation and control problems.

The traditional artificial intelligence and neural networks are usually considered suitable for solving the different types of problems. These two methods on the surface looks very different, but more and more current research on how to dedicatively everything can be incorporated into one and tectonic model, including their common ground. Artificial intelligence technology and neural network according to certain principles steps the integration of internal rules and things correlation, potential assumptions and Suggestions related reconciliation and baseless integration of artificial intelligence and neural networks. From leading researchers in the field of accomplishment have, this comprehensive text provides a very detailed introduces basic symbol processing, coupling network and their integration. Plenty of examples of integration of artificial intelligence and neural network and various specific applications for insight into this evolving area provides a very good method and help.

This paper expounds the artificially intelligent neural network in the power system engineering application. Former one illustrates the artificially intelligent neural network in the power system engineering have been tested in application. And then summarizes the application of artificial intelligence technology is widely, and puts forward Suggestions of improving existing artificial intelligence tools. Then, the paper focuses on the paper neural network power system and its application. Introduces multilevel feed-forward networks and problems existing in the application of established on the basis of the neural network method for power system application of the network are discussed. For further development of future theme in artificial intelligence and artificial neural network is applied to power system was put forward.

2 Feed-Forward Neural Networks Model

A lot of hyperbole has been devoted to neural networks, both in their first wave around 1960 and in their renaissance from about 1985, but the ideas of biological relevance seem to us to have detracted from the essence of what is being discussed, and are certainly not relevant to practical applications in pattern recognition. A formal definition of a feed-forward network is given in glossary. Informally, they have units which have one-way connections to other units, and the units can be labeled from inputs to outputs so that each unit is only connected to units with higher numbers. The units can always be arranged in layers so that connections go from one layer to a later layer. [3]

The network model such as Figure 1.

$$y_k = f_k \left(\alpha_k + \sum_{j \rightarrow k} w_{jk} f_j \left(\alpha_j + \sum_{i \rightarrow j} w_{ij} \right) \right) \quad (1)$$

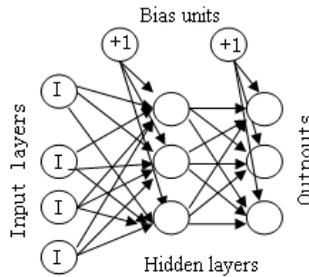


Fig. 1. ANN Model

Each units sums its inputs and adds a constant to form a total input χ_j and applies a function f_j to χ_j to give output y_j . The links have weights w_{ij} which multiply the signals travelling along them by that factor. The input units are there just to distribute the inputs, so have $f \equiv 1$. [4]

The function f_j are almost invariably taken to be linear, logistic (with $f(x) = \ell(x) = e^x / (1 + e^x)$) or threshold functions (with $f(x) = I(x \geq 0)$).

The general definition allows more than one hidden layers, and it also allows skip-layer connections from input to output. If all units in a layers have the same function f_h or f_o , we have

$$y_k = f_o \left(\alpha_k + \sum_{j \rightarrow k} w_{jk} x_j + \sum_{i \rightarrow j} w_{ij} f_h \left(\alpha_j + \sum_{i \rightarrow j} w_{ij} x_i \right) \right) \quad (2)$$

The bias terms can be eliminated by introducing a new unit 0 (the bias unit) which is permanently at +1 and connected to all other units. We set $w_{oj} = \alpha_j$. Thus we have the function as following

$$y_k = f_o \left(\sum_{i \rightarrow k} w_{ik} x_i + \sum_{j \rightarrow k} w_{jk} f_h \left(\alpha_j + \sum_{i \rightarrow j} w_{ij} x_i \right) \right) \quad (3)$$

We will briefly consider how such functions came to be suggested, and the theory which shows that they form large and flexible classes of functions. However, in practice the main factor are how the parameters should be chosen, and how the architecture (the number of layers and the number of units in each, as well as which connections to include) is selected.

3 Example

We give a example of electronic system load of an factory for a week. The data as following:

| | | | |
|--------|--------|--------|--------|
| 0.2085 | 0.6577 | 0.8389 | |
| | 0.1928 | 0.7877 | 0.8431 |
| | 0.1228 | 0.5468 | 0.8048 |
| | 0.1133 | 0.4636 | 0.8149 |
| | 0.1718 | 0.6011 | 0.7540 |
| | 0.3237 | 0.4426 | 0.8037 |
| 0.3721 | 0.6453 | 0.7627 | |

We take use six days data as training samples; take every two days data as input vector data, make the third day data as goal vector, take the seventh data as test samples, checking whether the Artificial Neural Network can prediction the electronic system load accurate or not.

4 Conclusion

The performance of the figure was examined in predicting electronic load of a week. Several ANN architectures were designed for this specific data devaluated for modeling load changes, and the mean measurement error was used for the evaluation accurate. The ANN model converged to the desired mean measurement error (MSE) for the training data, and the MSE was within the acceptable range for the test data. The performance of ANN can be significantly improved by omitting correlated data. The type of transfer function also has a key role in obtaining a converged ANN model. The translate function gives satisfactory results for this type of data. The results of this study show that ANN can be used as an effective tool in predicting electronic load.

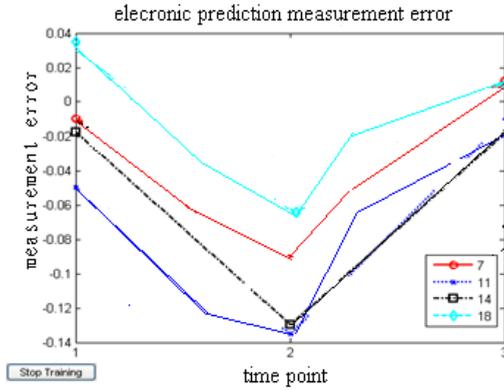


Fig. 2. Electronic measurement error

References

1. Kjeller, L., Jones, K.C., Ton, A.E.J., Rappe, C.: Evidence for a decline in atmospheric emissions of PCDD/Fs in the U.K. *Environ. Sci. Technol.*, 1398–1403 (2008)
2. Ridluan, A., Manic, M., Tokuhira, A.: *Nucl. Eng. Des.*, pp. 239–309 (2009)
3. *Pattern Recognition and Neural Networks* 231,266 (2010)
4. Information on, <http://www.itechate.com/en/products.jsp?sortid=001>

The Research on Tracking Concept Drift Based on Genetic Algorithm

Zhong Zhishui

Department of Mathematics and Computer Science
Tongling University
Tongling, China
zyj752@yeah.net

Abstract. Concept drift is a crucial problem in machine learning. The machine learning model, such as classifier, should adapt to the concept change rapidly to keep the precision. At present, most methods for solving this problem is based on the mechanism, which is referred to as Time –Weight. However, current existing methods can not quickly adapt to the concepts, which reappear. Therefore, a method with the ability to adapt concept changing all the time is necessary, especially for the situation of time precious, such as data stream classification. In this paper, we proposed a method of tracking concept drift based on genetic algorithm (TCDGA). We experimented with TCDGA on public dataset and made comparisons. Our results show that TCDGA can yield better performance when concepts reappear.

Keywords: machine learning, concept drift, genetic algorithm.

1 Introduction

In predictive analytics and machine learning, the concept drift means that the statistical properties of the target variable, which the model is trying to predict, change over time in unforeseen ways. This causes problems because the predictions become less accurate as the times passes. The term concept refers to the quantity you are looking to predict. More generally, it can also refer to other phenomena of interest besides the target concept, such as an input, but, in the context of concept drift, the term commonly refers to the target variable [1, 2].

The problem of changes or discontinuities that occur during learning or between learning episodes has attracted the attention of machine learning researchers for at least ten years now [3]. At present, most existing methods are based on different time – weight mechanisms [5, 6]. However, these methods have bad performance when adapting to concept reappears. Concept reappear is also called recurring contexts, the problem may be due to cyclic phenomena, such as seasons of the year or may be associated with irregular phenomena, such as inflation rates or market mood. In such domains, in order to adapt more quickly to concept drift, concept descriptions may be saved so that they could be reexamined and reused later. Not many learners are able to deal with recurring contexts [7, 8].

To avoid limitations of traditional method, we proposed a method of tracking concept drift based on genetic algorithm. The basic steps of our method are as follows:

(1) analyze the context and regard the context as a combination of basic context elements; (2) after step one, all the contexts in a specific problem can be represented by a set of context elements. In another word, each context can be coded as formalization; (3) when the specific machine learning model detects the performance is much worse that a predefined threshold, our method will launch the process for replace current context with a new one. The process is based on the genetic algorithm and the contexts which emerged before will be import as chromosomes.

Our contribution is proposing a method that is used for tracking concept drift. Our method can be combined by a specific machine learning model to improve its performance when handing the problem of recurring contexts.

We experimented with TCDGA on public dataset and made comparisons. Our results show that TCDGA can yield better performance when concepts reappear. The rest of this paper is organized as follows: Section 2 examines current tracking context drift technologies. In chapter 3, we describe the problem in formulation. In chapter 4, we present TCDGA for improving performance of tracking recurring contexts of machine learning model. In chapter 5, we experimented with TCDGA, made a comparison between TCDGA and traditional methods. Finally, we give a conclusion.

2 Related Work

Much research and effort has focused on concept drift. In this section, we will only review some notable work due to space limitation.

Most notable method for solving concept drift is based on time-weight mechanism [8]. An intelligent agent called NewsDude that is able to adapt to changing user's interests learns two separate user models: one represents the user's short-term interests and the other one represents the user's long-term interests. The short-term model is learned from the most recent observations only. It represents user models that can adjust more rapidly to the user's changing interests. If the short-term model cannot classify the story at all, it is passed on to the long-term model. The purpose of the long-term user model is to model the user's general preferences for news stories and compute predictions for stories that could not be classified by the short-term model.

An offline meta-learning algorithm for identification of hidden context is presented in [9]. The approach assumes that concepts are likely to be stable for some period of time. It uses batch learning and contextual clustering to detect stable concepts and to extract hidden context.

In [10] and [11] a method for selecting training examples for a partial memory learning system is described. The forgetting mechanisms of the method select extreme examples that lie at the boundaries of concept descriptions and remove examples from partial memory that are irrelevant or outdated for the learning task. The method uses a time-based forgetting function to remove examples from partial memory, which are older than a certain age.

All the above mentioned methods do not consider how to handle the recurring contexts in a more efficient way. Hence, we are motivated to proposed TCDGA for avoiding this limitation.

3 Preliminary

We shall first define some notations that we will be using throughout this paper, and then give the problem description.

A difficult problem with learning in many real-world domains is that the concept of interest may depend on some hidden context, not given explicitly in the form of predictive features. A typical example is weather prediction rules that may vary radically with the season. Another example is the patterns of customers' buying preferences that may change with time, depending on the current day of the week, availability of alternatives, inflation rate, etc. Often the cause of change is hidden, not known a priori, making the learning task more complicated. Changes in the hidden context can induce more or less radical changes in the target concept, which is generally known as concept drift. An effective learner should be able to track such changes and to quickly adapt to them. In many domains, hidden contexts may be expected to recur. Recurring contexts may be due to cyclic phenomena, such as seasons of the year or may be associated with irregular phenomena, such as inflation rates or market mood. In such domains, in order to adapt more quickly to concept drift, concept descriptions may be saved so that they could be reexamined and reused later. Not many learners are able to deal with recurring contexts.

Since there are two many machine learning technologies and every technology have different representation of hidden context. Thus, in this paper, we take classification as an example to describe our method TCDGA.

The hidden context in the problem of classification is the mapping relationship between attribute and category.

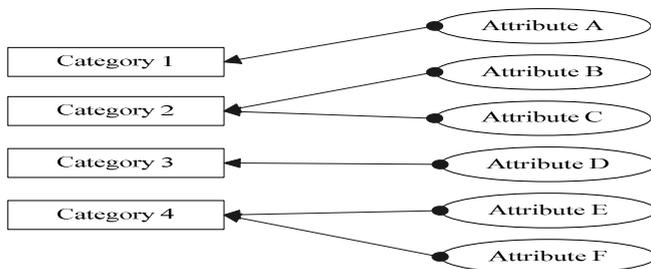


Fig. 1. Hidden Context

Then we can use a vector to represent the hidden context. Suppose there are attributes $\{a_1, a_2, \dots, a_m\}$ and categories $\{c_1, c_2, \dots, c_n\}$. if $i_1, i_2, \dots, i_m \in \{1, 2, \dots, n\}$, then a hidden context can be described as follows:

$$\text{Context} = \langle c_{i_1}, c_{i_2}, \dots, c_{i_m} \rangle \quad (1)$$

Of course, the mapping relationship between the categories and attributes is the simplest situation. In most cases, the mapping relationship might be much more complex. However, the context can also be described by a sequence of category.

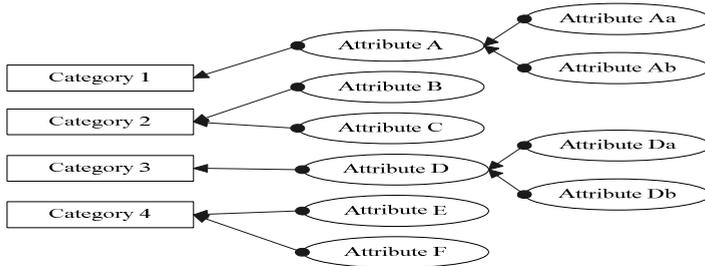


Fig. 2. Hierarchical Hidden Context

In figure 2, we present a hierarchical mapping relationship. In this situation, we can still use the formulation method in formula 1.

The two cases mentioned above can be used, when the value of attribute is discrete and Boolean value. When the value of attribute is continuous, we can use another formulation method. Suppose there are two attribute A and B. The value range of A is [a, b], the value range of B is [c, d], we can separate the value range of A and B to small pieces, such as $A=[a_1, a_2, a_3, \dots, a_m]$ and $B=[b_1, b_2, b_3, \dots, b_n]$ and we can represent the context as a vector, which length is $m+n$. if $i_1, i_2, \dots, i_{m+n} \in \{1, 2, \dots, n\}$, then a hidden context can be described as follows:

$$\text{Context} = \langle c_{i_1}, c_{i_2}, \dots, c_{i_{m+n}} \rangle \tag{2}$$

4 Method

In this section, we first present the basic steps of TCDGA. We then present necessary details of some step respectively. The basic steps of TCDGA are as follows:

Step 1: analyze the context and regard the context as a combination of basic context elements

Step 2: after step one, all the contexts in a specific problem can be represented by a set of context elements. In another word, each context can be coded as formalization

Step 3: when the specific machine learning model detects the performance is much worse that a predefined threshold, our method will launch the process for replace current context with a new one. The process is based on the genetic algorithm and the contexts which emerged before will be import as chromosomes.

We design a genetic algorithm. Evolutionary algorithms are popular approaches to improve the performance of machine learning [12, 13, 14]. The detail of our algorithm is as follows:

(1) Chromosome: Context is used as the chromosomes for our method;

(2) Crossover: two chromosomes, called parents, are combined together to form new chromosomes, called offspring. We use single point crossover, the position for crossover is chosen at random; in this paper, the probability of chromosomes for crossover is 0.85.

(3) Mutation: the mutation operator introduces random changes into characteristics of chromosomes. Each position of chromosome will change between different categories with a certain probability; in this paper, the probability of chromosomes for mutation is 0.1.

(4) Fitness function: in our method, the fitness function is the performance of classification when we use the hidden context, which is represented as chromosome.

(5) Selection: reproduction involves selection of chromosomes for the next generation. In the most general case, the fitness of an individual determines the probability of its survival for the next generation. We use proportional selection for selection procedures.

(6) Initialization: the most important of the genetic algorithm is the step initialization. Each time, when we found a nice context, we will code the context as a chromosome and store it in the database. We the context is not proper for current data; the genetic algorithm will be launched. Before the launching, we will import all context as the chromosomes in the generation 1, and the number of each context' chromosomes is depend on the fitness of the context.

We did necessary experiments in next section.

5 Experimental Results

Experiment Description

The operation system of server is Windows XP Professional Edition. The CPU of server is Inter(R) Core2 Duo CPU 2.8GHZ. The RAM is 1.0G.

We experiment with TCDGA on the dataset, which we retrieve from a text classification dataset and separate them in to 10 different groups according to their topic. The hidden context is the mapping between keywords and categories.

Evaluation Criterion

We use there evaluation criterions to measure the effect of TCDGA.

An ideal concept drift handling method should be able to: (1) quickly adapt to concept drift; (2) be robust to noise and distinguish it from concept drift; and (3) recognize and treat recurring contexts.

In this paper, we focus on the third criterion to compare TCDGA with traditional method.

Experiment Result

We experimented according to the simulated data exactly and make a comparison between traditional method and TCDGA:

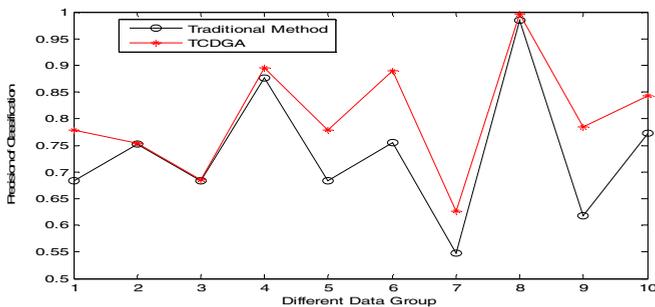


Fig. 3. The presicion of classification

The result shows that TCDGA can do good performance on classification. When the hidden contexts changes frequently, TCDGA has a better performance than traditional methods.

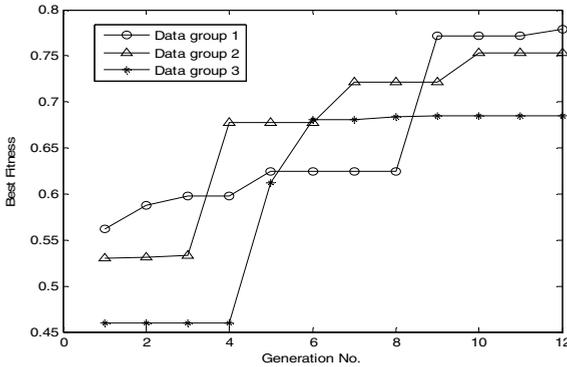


Fig. 4. Convergence analysis

The convergence analysis results show TCDGA is effective for finding a hidden context, which fit the current data.

6 Conclusion and Future Work

Concept drift is a crucial problem in machine learning. The machine learning model, such as classifier, should adapt to the concept change rapidly to keep the precision. At present, most methods for solving this problem is based on the mechanism, which is referred to as Time –Weight. However, current existing methods can not quickly adapt to the concepts, which reappear. Therefore, a method with the ability to adapt concept changing all the time is necessary, especially for the situation of time precious, such as data stream classification. In this paper, we proposed a method of tracking concept drift based on genetic algorithm (TCDGA). We experimented with TCDGA on public dataset and made comparisons. Our results show that TCDGA can yield better performance when concepts reappear.

References

1. Kolter, J.Z., Maloof, M.A.: Dynamic Weighted Majority: An ensemble method for drifting concepts. *Journal of Machine Learning Research* 8, 2755–2790 (2007)
2. Scholz, M., Klinkenberg, R.: Boosting Classifiers for Drifting Concepts. *Intelligent Data Analysis (IDA), Special Issue on Knowledge Discovery from Data Streams* 11(1), 3–28 (2007)
3. Widmer, G.: Learning in Dynamically Changing Domains: Recent Contributions of Machine Learning. In: *Proceedings of the MLNet Workshop on Learning in Dynamically Changing Domains: Theory Revision and Context Dependence Issues*, Czech Republic, Prague (1997)

4. Learning drifting concepts: Example selection vs. example weighting
5. Klinkenberg, R., Renz, I.: Learning for Text Categorization, Adaptive information filtering: Learning in the presence of concept drifts (1998)
6. Maloof, M.A., Michalski, R.S.: Selecting examples for partial memory learning. *Machine Learning* 41(11), 27–52 (2000)
7. Mitchell, T., Caruana, R., Freitag, D., McDermott, J., Zabowski, D.: Experience with a Learning Personal Assistant. *Communications of the ACM* 37(7), 81–91 (1994)
8. Gradual forgetting for adaptation to concept drift
9. Schlimmer, J., Granger, R.: Beyond Incremental Processing: Tracking Concept Drift. *AAAI* (1986)
10. Wang, H., Fan, W., Yu, P.S., Han, J.: Mining concept-drifting data streams using ensemble classifiers. In: *KDD* (2003)
11. Widmer, G., Kubat, M.: Learning in the presence of concept drift and hidden contexts. *Machine Learning* 23, 69–101 (1996)
12. An introduction to genetic algorithms
13. Learning with Genetic Algorithms: An Overview
14. Development of Genetic-based Machine Learning for Network Intrusion Detection

Prototype Design of Hierarchical Cluster Multiprocessor on Platform FPGA

RuiHui Mu¹, PeiZhang Zhao¹, and Dong Li²

¹ Xinxiang University, Xinxiang, P.R. China

² Bureau of Land and Resources, Xinxiang, P.R. China
mym0822@yahoo.cn

Abstract. The MPSoC(Multi-Processor System on Chip) is a promising solution for the coming complex computer and embedded systems. And the the NoC(Network on Chip) architecture has been proposed as an effective interconnection on chip and been used in many designs of high performance embedded systems. But the NoCs and MPSoC bring more challenges on parallel programming and synthesis for applications design. This paper proposes a new hierarchical cluster multiprocessor architecture, which adopts NoC interconnection and hierarchical clusters architecture. And this design has been implemented as a prototype on FPGA. The performances of the prototype is evaluated by a parallel JPEG decoder, and the approach can improve by 143.2%.

Keywords: Multi-Processor System on Chip, Network on Chip, Embedded System Architecture.

1 Introduction

With the rapid improvement of microprocessor technology, MPSoC system, featured as high in performance, extensible, small in size and low in power consumption is able to meet requirements of complex embedded systems for computing performance, power consumption and cost, thus having become one of the mainstream development trends of embedded systems [1]. Currently, MPSoC is widely used in computer communication, digital media, auto electronics, medical electronics, aviation and aerospace, industrial control and many other fields, and the relevance of research topics on MPSoC has engaged many enterprises and institutes in research on MPSoC, such as programming model, task scheduling, system development, debugging environment and system safety[2]. Recently, there has been a growing interest in design and modeling methods targeted towards the co-design of complex MPSoC system [3], [4].

The use of stable formal methods to design and analysis the MPSoC system is a difficult problem. The best way to tackle such a problem is to design the system by communication driven method [5]. So formal verification research becomes more and more popular and using Petri net is a distinct feature of common design methodology, an example of this overall methodology based on formal specification of SoC system design is shown in [6]. Petri net is used as a formal executable method for design and analysis of concurrent, discrete-event dynamic systems [7]. Our design methodology is based on recent advances on asynchronous design with Petri net.

Recent years Model Checking has been extensively used in the verification of hardware and software of the embedded systems. Predicate Abstraction [8] is modified by many scientists and used in the hardware and software verification of the embedded systems. The PISH co-design system [9], partitions the hardware and software of the embedded systems with Petri Net. PRES [10] uses Petri Net to describe the hardware and software and then verifies the description in hybrid automata model [11]. But these methods mentioned above have few considerations about multiprocessor architecture.

In this paper, we design and analyze the parallel processing of a MPSoC (multiprocessor system on chip) system architecture, and establish a Event Count Model to capture the timing characters of system, for generally evaluation on the system performances of CPU utilization and processing speedup. We illustrate the function of high level colored time Petri net and CPN Tools in the system design. Petri nets show more advantages to make the complex design and analysis more simple and clear about the process of the video system active figure. We also can analyze the Petri net graph using CPN Tools which will give great help to the designers.

2 MPSoC System Design Method

A high performance FPGA was used to implement an h.264 encoder, and the primary MPSoC video system is built on the XUPV5-LX110T Development Board from Xilinx. In figure 1, a simplified block diagram of the FPGA board architecture is shown. Our experiments are built on the FPGA, with setting the MPSoC system on FPGA, the multi-core processor architecture has been designed for the special use of mass data computation. The MPSoC has five MBs (MacroBlaze), which is a software IP core of 32bit general CPU. These MBs are connected by the on chip high speed bus directly. The Main Processor is in charge of the system controller and initialization.

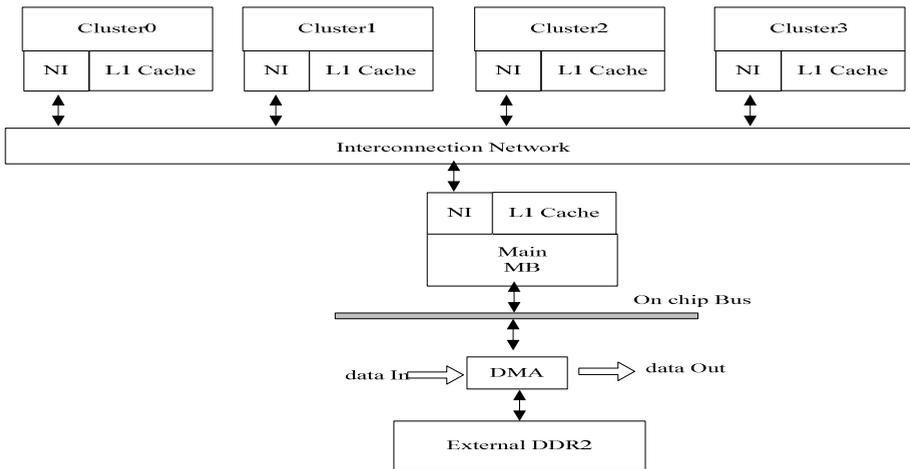


Fig. 1. Architecture of MPSoC system



And in our MPSoC design, there are a main MB, which is in charge of the system controller and initialization, and other Synergistic Processors are organized into clusters. Synergistic Processors execute parallel data computation separately. Each processor has a local L1 cache to load and store data used by executing unit. So there is potential competition to the occupation of bus data transmission. There should be some methods to be evaluated for scheduling and communication of system.

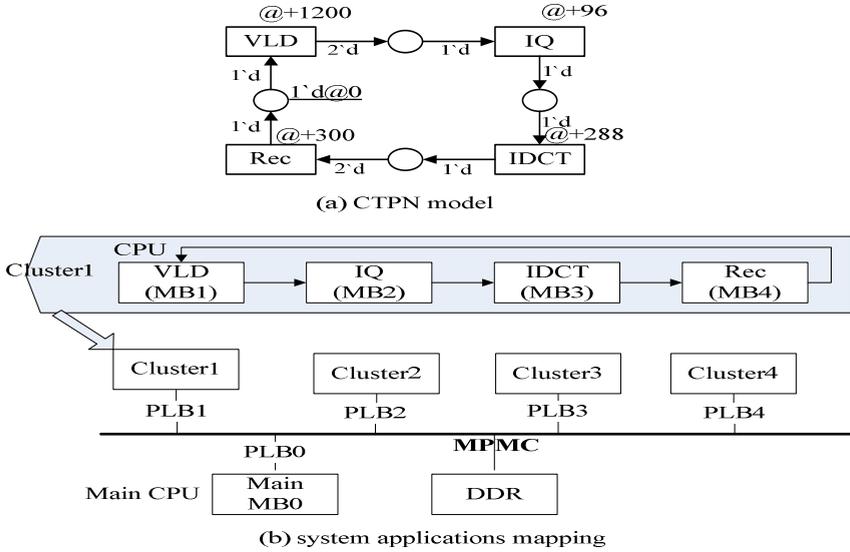


Fig. 2. Architecture of the cluster and application mapping

The cluster multiprocessor is based on symmetric multiprocessor architecture, as shown in figure 2. For obtaining better performances, we adopt hierarchy communication architecture. And in the design of application The system application model can be represented in a ten-tuple CTPN (Colored Time Petri Net), as

$$CTPN_s = \{\Sigma, P, T, W, Co, C+, C-, G, E, I_0\}$$

- (1) Σ is the color set of the system net model, determines the data types, operators and functions can be used in the net.
- (2) P is a finite set of places, represents the input and output units of the cores.
- (3) T is a finite set of transitions, represents logic functions of the cores.
- (4) W is a multi-set of arcs, defines the relations of places and transitions, $W : P \times T \cup T \times P$.
- (5) Co is the color function defined from $P \cup T$ to Σ , Co maps each place p to a set of token colors $Co(p)$ and each transition t to a set of occurrence colors $Co(t)$.



- (6) $C+$ and $C-$ are the post-incident and the pre-incident, labelled time for set of colors $Co(p)$ and $Co(t)$, so the flow matrix or incidence matrix can be defined as $C = (C+) - (C-)$.
- (7) G is the expressions of guard function for transitions, which controls the fire condition of transitions.
- (8) E is the arc expression function, describing the input and output tokens transferred between places and transitions.
- (9) I_0 is the initialization function, which specifies the initial states of the net.

And Figure 3.(a) shows an application dataflow model of CTPN, which is a JPEG decoder. Each transition represents the data processing, which is located on the processor; and place is the communication channel between processors, number of tokens means the capacity for the data transfer. In Figure 3.(b) these application tasks are assigned onto CPUs in a cluster, and the CPUs are connected with FSL fast links directly. So the MPSoC system can be designed in a fast exploration.

3 Case Study

A high performance FPGA was used to implement an MPSoC of JPEG decoder, as shown in Figure 3 and the multi-core processor architecture has been designed for the special use of mass data computation. The MPSoC has five MBs (MicroBlaze), which is a software IP core of 32bit general CPU. These MBs are connected by the on chip high speed interconnection network directly. Each processor has a local L1 cache to load and store data used by executing unit. So there is potential competition to the occupation of data transmission. The Multi-Fuzzy Scheduling is running on the Main MB to be evaluated for scheduling and communication of system.

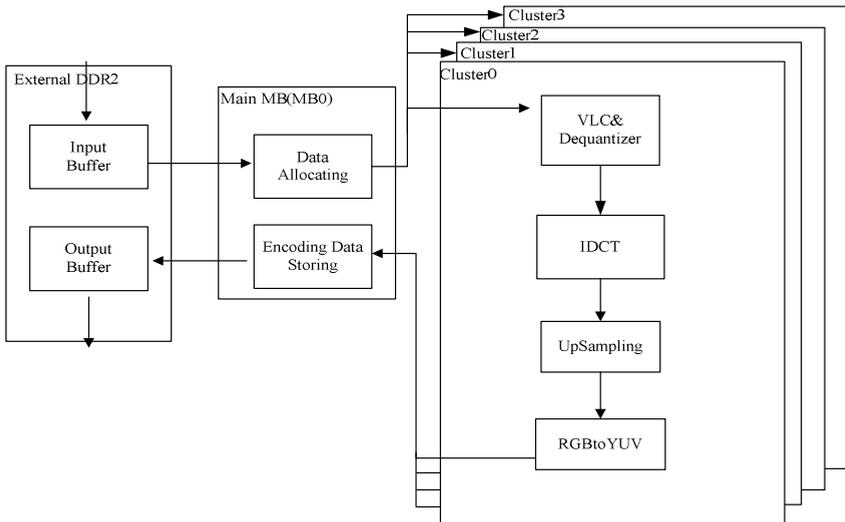


Fig. 3. The Data Flow and Task Allocation of JPEG decoder

In the JPEG decoding processing, a picture is divided into many blocks. Hence it becomes the basic processing object for each core. The shared local memory (L1 cache) and some buffers between cores (such as FIFOs) are using for cache data, by which the one-way data stream can be divided into Multi-stage pipeline to greatly improve the system parallel performances.

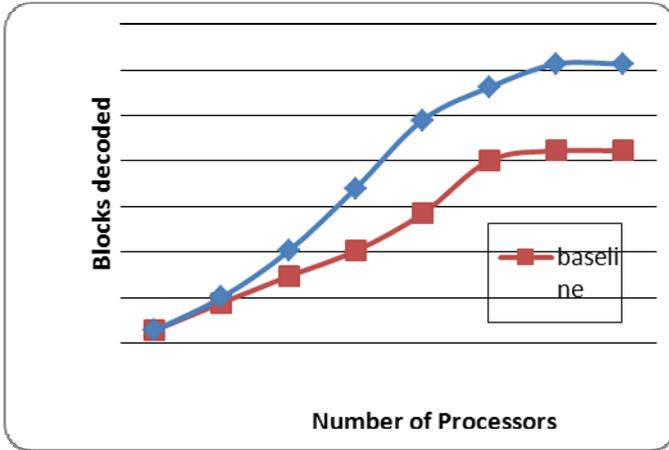


Fig. 4. Execution Performance of JPEG decoder

The Figure 4 shows the execution performance of the JPEG decoding for different number of processors. We have run the decoder under 100MHz system clock. In this example the baseline case allocates the tasks with the invariable macroblocks onto these processors, and the improved case use the proposed hierarchical cluster multiprocessor architecture. Comparing performances of baseline case and improved case, the improved approach can process more JPEG pictures data, and it get higher improvement with increasing of number of processors. The highest speedup ratio is up to 15.34, 143.2% of the baseline case. The improved case can achieve better speedup ratio performance, because it minimize the difference of each processor's executing time so that there are less communications and high utilization.

4 Conclusion

In this paper, a new hierarchical cluster multiprocessor architecture is proposed, which adopts the NoC interconnection and cluster multiprocessors subsystem. And the applications are assigned into the processors by the CTPN dataflow models, which can be used for investigating the architecture design and mapping policies in the MPSoC fast design exploration. The proposed MPSoC architecture has been implemented as a prototype by FPGA board, which integrated many microblaze cpu cores in a chip. And the performance of this prototype is evaluated in a parallel JPEG decoder application design, the speedup ratio is up to 15.34, 143.2% of the baseline.



References

1. Wolf, W.: The future of multiprocessor systems-on-chips. In: Proc. 41st Annual Conference on Design Automation, pp. 681–684 (2004)
2. Baruah, S., Nathan, F.: The feasibility analysis of multiprocessor real-time systems. In: Proc. EuroMicro Conference on Real-Time Systems, pp. 85–94 (2006)
3. Kapasi, U.J., et al.: Programmable stream processors. *IEEE Computer* 36(8) (2003)
4. Maxiaguine, A., Kunzli, S., Chakraborty, S., Thiele, L.: Rate analysis for streaming applications with on-chip buffer constraints. In: ASP-DAC 2004 (2004)
5. Benini, L., De Micheli, G.: Networks on Chips: A New SoC Paradigm. *Computer* 35(1), 70–78 (2002)
6. Yakovlev, A., Furber, S., Bystrov, A.: Design and Analysis of aSelf-Timed Duplex Communication System. *IEEE Transactions on Computers* 53(7) (July 2004)
7. Cortadella, J., Kishinevsky, M., Kondratyev, A., Lavagno, L., Yakovlev, A.: Logic Synthesis of Asynchronous Controllers and Interfaces. Springer, Heidelberg (2002)
8. Graf, S., Saidi, H.: Construction of Abstraction State Graphs with PVS. In: Grumberg, O. (ed.) CAV 1997. LNCS, vol. 1254, pp. 72–83. Springer, Heidelberg (1997)
9. Maciel, P., Barros, E., Silva, M., et al.: Resource Sharing Estimation by Petri Nets in PISH Hardware/Software Co-Design System. In: 2000 IEEE International Conference on Systems, Man, and Cybernetics, October 8-11, vol. 5, pp. 3165–3170 (2000)
10. Cortes, L.A., Eles, P., Peng, Z.: Formal Coverification of Embedded Systems Using Model Checking. In: Proceeding of the 26th Euromicro Conference, vol. 1, pp. 106–113 (September 5-7, 2000)
11. Alur, R., Henzinger, T.A., Ho, P.H.: Automatic Symbolic Verification of Embedded Systems. *IEEE Trans. Software Engineering* 22, 181–201 (1996)
12. Alurand, R., Dill, D.L.: A theory of timed automata. *Theoretical Computer Science* 126(2), 183–235 (1994)
13. Kerbaa, A., Maler, O.: Task graph scheduling using timed automata. In: IPDPS (2003)

Efficient Memory Processors Design of Multiple Applications for Multiprocessor Architecture

Liang Tian¹, Hong Pan¹, and Dong Li²

¹ Xixiang University, Xixiang, P.R. China

² Bureau of Land and Resources, Xixiang, P.R. China
langtianyi@sina.cn

Abstract. Memory access latency and related operations scheduling are often the performances bottleneck in the parallel architecture design of Multiprocessor systems. In this paper we present a model of efficient memory operations, which is an on chip network transaction controller. Utilizing the controller model, we can replace the multiple transactions of memory accesses by the multilayer transactions, which use integrated high level transactions and local-memory computation to reduce the system's communication consumption. We implement a hierarchical memory processor cluster which is located near the memory and executes the memory operations. In our case studies, we applied the design to applications (parallel H.263 decoder) running on a FPGA MPSoC system and found that this transaction approach can improve performance by speedup ratio of 12.59.

Keywords: MPSoC System, Embedded System design, Computer Architecture.

1 Introduction

MPSoC system, featured as high in performance, extensible, small in size and low in power consumption in microprocessor technology is able to meet requirements of complex embedded systems for computing performance, power consumption and cost, thus having become one of the mainstream development trends of embedded systems [1]. Currently, MPSoC is widely used in computer communication, digital media, auto electronics, medical electronics, aviation and aerospace, industrial control and many other fields, and the relevance of research topics on MPSoC has engaged many enterprises and institutes in research on MPSoC, such as programming model, task scheduling, system development, debugging environment and system safety[2]. Recently, there has been a growing interest in design and modeling methods targeted towards the co-design of complex MPSoC system [3], [4].The use of stable formal methods to design and analysis the MPSoC system is a difficult problem. The best way to tackle such a problem is to design the system by communication driven method [5]. Petri net is used as a formal executable method for design and analysis of concurrent, discrete-event dynamic systems [6]. Our design methodology is based on recent advances on asynchronous design with Petri net.

2 Related Work

Recent years Model Checking has been extensively used in the verification of hardware and software of the embedded systems. Predicate Abstraction [7] is modified by many

scientists and used in the hardware and software verification of the embedded systems. The PISH co-design system [8], partitions the hardware and software of the embedded systems with Petri Net.

In this paper, we design and analyze the parallel processing of a MPSoC (multiprocessor system on chip) system architecture, and establish a Event Count Model to capture the timing characters of system, for generally evaluation on the system performances of CPU utilization and processing speed up. We also can analyze the Petri net graph using CPN Tools which will give great help to the designers. We illustrate the function of high level colored time Petri net and CPN Tools in the system design. Petri nets show more advantages to make the complex design and analysis more simple and clear about the process of the video system active figure.

3 MPSoC System Architecture

Our experiments are built on the FPGA, with setting the MPSoC system on FPGA, the heterogeneous multi-core processor architecture has been designed for the special use of mass data computation. The MPSoC has five MBs (MacroBlaze), which is a software IP core of 32bit general CPU. These MBs are connected by the on chip high speed bus directly. The Main Processor is in charge of the system controller and initialization, and other Synergistic Processors execute parallel video data computation separately.

The system interface process will be designed as two modules. One of these modules is the control set, used for transfer various system instructions on registers; and another is data buffer set, which is a non-burst buffer allows data be transferred. The system timing information is described by a Event Count model, and the timing characteristics is translated into a high level timed Petri net description to be transferred with system interactions. In the abstract multiprocessor architecture “Reg” and “Buf” are the places where the interaction instructions and temporary data of executions are stored and judged separately. The execution results will be written back to “Buf”.

4 Event Count Model

An Event Count Model captures the timing characteristics by counting the number of data items in a unit interval time. This count is used to update the relative count variables with the system Petri net. Several guard values can be used to decide at which time the timing characteristics will fire a transition. In this paper, the guard values are denoted as $t \diamond p$. a function of conjunctions of atomic presents of set $\{<, >, =, \leq, \geq\}$. t is a count variable, p may be a constant or parameter of other variable. Further more, in a high level timed Petri net model, a token will be denoted as $1'x @ t$. meaning a colored token x with time label \tilde{t} An Example ECM is shown in Figure 1. The tasks are executed on three modules (A,B,C), In the initialization on A module, the system will execute i items in a unit of time, ($i \in [1, 3]$), and when the total number of tasks on A exceeds 9 then tasks will be transferred onto B if the task count is 10 or onto C if the task count exceeds 10. B module execute 1 to 2 tasks and C execute 2 to 4 tasks in a unit of time, and tasks will return to run on A while the task count exceeds 5. Once C is

entered it never moves to others again. Further system scheduling and communication between A and B costs 5 unit of time, and that between A and C costs 4 unit of time. Thus we can simulate the automaton and get a possible run $A(0@0) \rightarrow A(3@1) \rightarrow A(5@2) \rightarrow A(7@3) \rightarrow A(10@4) \rightarrow B(0@9) \rightarrow B(1@10) \rightarrow B(2@11) \rightarrow B(4@12) \rightarrow B(5@13) \rightarrow B(6@14) \rightarrow A(6@19) \rightarrow A(7@20) \rightarrow A(8@21) \rightarrow A(11@22) \rightarrow C(11@26) \rightarrow A(13@27)$, that executes tasks as: (3@1, 2@2, 2@3, 3@4, 1@20, 1@21, 3@22) on A, (1@10, 1@11, 2@12, 1@13, 1@14) on B, and (2@27) on C.

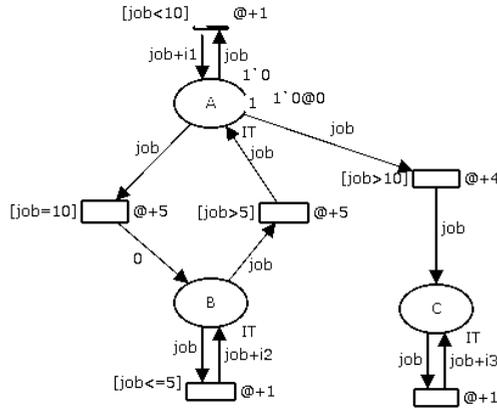


Fig. 1. Example of Event Count Model

5 MPSoC Modeling

The MP and SP cores in MPSoC system architecture run synchronously, so Petri net is just a useful tool to describe this asynchronous and concurrent system. The process of the modeling is generally the following.

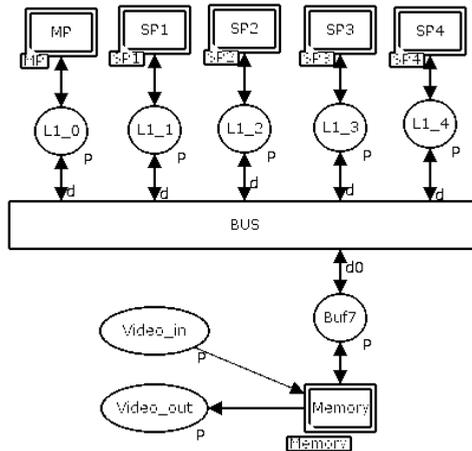


Fig. 2. Petri Net Diagram of MPSoC Modeling



1) Each execution step can be inserted into a hierarchy subpage of transition of component, where the core logic design is described as a subnet in detail. The places “Reg” and “Buf” are the input/output the core, they are displayed in several pages vertically;

2) The timing characteristics are using in the subpage architecture design, the tasks tokens are described with a product color set, <Data, Func, Time>.

3) For the data must be processed in the right order, that need a ID sequence control inserted into the Petri net. “Dataload” represents loading correct data from buffer into the executing cores.

4) The MP core sets the data parsing parameters, and initializes other SP cores data ID, so the each of the cores has its own video data.

With this diagram, the running process of a simple MPSoC architecture can be simulated and each step of the execution is mapped with different state. Some useful CPN Tools can also be used to make some further analysis.

6 Simulation and Analysis

For the Petri net, the CPN Tools supports multiple simulation modes, as a single transition fire, small number transition fire, and fast forward fire, for simulations of different objects, these simulation functions may give great help to the designers for tracing the simulations or analysis on the performance in an interval of time. And the CPN Tools has a powerful simulation and analysis tool. It can calculate state space of the net, and also calculate a strongly connected components (SCC) graph. And further it support queries for investigating the properties of a CP-net. It is possible to investigate the reach ability, boundedness, home, liveness and fairness properties by using standard queries in CPN ML functions. We run the simulation the CTPN model with CPN Tools, with different CPU numbers. Comparing to the results of real MPSoC system on FPGA, Figure 3 shows the speedup of these two methods. The slope of the speedup is nearly linear, and that of real system is a little lower. This is because that the real system’s scheduling is more complex than that supposed in CTPN model.

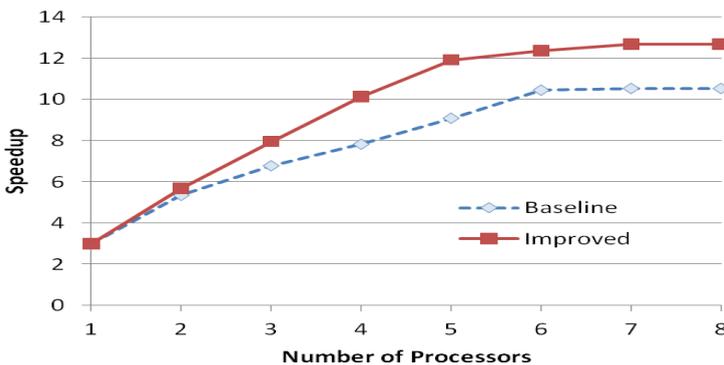


Fig. 3. Speedup ratio of CTPN model

7 Conclusion

Through the above case, which design and analyze a video coding process on the MPSoC system, we illustrate the framework of design the system with a high level colored time Petri net and functions of CPN Tools in the design and simulation. In the more complex design in detail, using Petri net makes the complex design and analysis more simple and clear. And the model can be further extended. The simulation result shows that this design idea of CTPN model is near to the real system. And the future design work is to establish the accurate timing model of detail, and research the methods of description with less Petri net nodes, which will be better for perform formal co-design and co-verification of complex SoC design.

References

1. Wolf, W.: The future of multiprocessor systems-on-chips. In: Proc. 41st Annual Conference on Design Automation, pp. 681–684 (2004)
2. Baruah, S., Nathan, F.: The feasibility analysis of multiprocessor real-time systems. In: Proc. EuroMicro Conference on Real-Time Systems, pp. 85–94 (2006)
3. Kapasi, U.J., et al.: Programmable stream processors. *IEEE Computer* 36(8) (2003)
4. Maxiaguine, A., Kunzli, S., Chakraborty, S., Thiele, L.: Rate analysis for streaming applications with on-chip buffer constraints. In: ASP-DAC 2004 (2004)
5. Benini, L., De Micheli, G.: Networks on Chips: A New SoC Paradigm. *Computer* 35(1), 70–78 (2002)
6. Cortadella, J., Kishinevsky, M., Kondratyev, A., Lavagno, L., Yakovlev, A.: Logic Synthesis of Asynchronous Controllers and Interfaces. Springer, Heidelberg (2002)
7. Graf, S., Saidi, H.: Construction of Abstraction State Graphs with PVS. In: Grumberg, O. (ed.) CAV 1997. LNCS, vol. 1254, pp. 72–83. Springer, Heidelberg (1997)
8. Maciel, P., Barros, E., Silva, M., et al.: Resource Sharing Estimation by Petri Nets in PISH Hardware/Software Co-Design System. In: 2000 IEEE International Conference on Systems, Man, and Cybernetics, October 8-11, vol. 5, pp. 3165–3170 (2000)

Research of Discrete Element Method in Vibratory Feeder

Yaluo Yang, Yayu Huang, and Hao Wu

School of Mechatronic Engineering, Kunming University of Science and Technology, Kunming,
Yunnan, China
luoreninto@sina.com

Abstract. Contact model is the foundation of discrete element method, by comparing the linear contact model and nonlinear contact model of DEM, Cundall spring-damp-sliding contact model was selected as our major studying model. Considering the self-reverence role of granule and the damping effect between granule and vibration plane, we study the elastic collision, elastic-plastic collision and turning of granule on static vibration plane, and comparing the different parameters settings to its effects. At last, the motion state of granule group was simulated after vibration box loaded.

Keywords: Vibratory feeder, Discrete element method, Granule, Contact model.

1 Introduction

There has a profound impact on the study of the granular material in vibratory feeder's operation and development . The first study was similar to the discrete fluid particle continuum for processing, analysis the contact forces between particles and contact-impact relationship, but the discrete particle aggregates in the complex was not a continuous movement, unlike the fluid structure, does not meet the continuous medium model. Later, with the technology to improve and continue to explore and found the use of discrete element method, granular computing contact forces between particles, the motion state of particle collision analysis and simulation studies to meet our demand for granular particles. Discrete element method was based on the characteristics of granular build their own discrete numerical model, consider the factors of discrete plastic body, the interaction between the discrete particles, discrete particles and the interactions between the outer boundary of the research methods, and traditional The continuum model analysis and simulation completely different, and is widely used in mechanical calculation and engineering applications.

Discrete element method (DEM) was developed in the 20th century, used to calculate the age of 70 granular media system (dry granular, wet granular or two-phase suspension) of the mechanical behavior of a numerical method, by Cundall [1] first proposed The basic idea was to study the bulk of the body (or medium) as a separate discrete "unit" or "particles " adjacent "cells" or more of the force between a "unit " of the movement by Newton's second law of domination, by tracking each "unit " of the micro-motion simulation of the object can be mechanical, thermal, physical or chemical state of the distribution and evolution. Discrete element method can be better applied to feeder plane vibration of material particles in the contact model.

2 Discrete Element Model

2.1 DEM Contact Model of Particle Type

1. Cundall's discrete element contact model: Cundall disk element discrete element contact model is a flat disc problems, disc or cylindrical particles in fact very few, but some people do with the disc obtained with the BALL Photo elastic simulation of a similar program Contact force field [1].

2. Thornton's three-dimensional sphere dry contact model: Thornton [2] and other comprehensive ball by predecessors in the normal contact mechanics - the theory of tangential role, but also by the results of contact mechanics model, the normal function of the Hertz theory [3] and tangential role and Deresiewicz Mindlin theory [4]; on the adhesion of spherical particles, the normal contact force based on Hertz theory to consider adhesion forces JKR (Johnson-Kendall-Roberts) theory [5] ; tangential contact force According to Savkoor and Briggs incremental theory and Mindlin and Deresiewicz theory [6] the combination of theory to define the formation of Thornton. Improved model of discrete element.

3. Based on Cundall's DEM contact model, there are a number of improved contact models or algorithm.

4. Non-circular or dry ball of discrete element contact model: for the particles vary widely with the actual disk element, the ball element idealization of particles, there has been non-round DEM contact model.

5. The role of interstitial fluid particle model for the wet: wet particle is a dense granules - interstitial fluid hybrid structure, the creep of this type of multi-phase body, and only caught the main factors to simplify the current model is only approximate results may be obtained.

2.2 The Choice of DEM Model

A current widely used contact model is proposed by Cundall et al DEM used in the first spring - damper - sliding contact model (spring-dashpot-slider model). There is also a contact mechanics studies, using the effective contact model, which Thornton[2] three-dimensional sphere dry contact model, the role of law according to Hertz[3] theory, and according to Midlin tangential role and Deresiewicz[4]the theory of identification, with Cundall [1] compared to discrete element contact model, citing the results of contact mechanics model is nonlinear. Instead of using the linear model was used Cundall reference to the results of the nonlinear contact mechanics model, we need to make a comparison of two models.

The results showed that at 5° and 30° under the action of normal force, shear force and the velocity variation were compared with simulated values from the collision point of view, the tangential force of the linear normal force and non-linear of tangential force variation of normal force is almost identical, complex linear contact model and simple model simulation results were in agreement. At the same time Cleary also believe that: a simple linear model with a more complex nonlinear models to bring the difference between the simulation results, with the use of the ball element and not considering the actual particle shape caused by the simulation results with the actual phase difference between the Q less than a theory, so a contact model which

impact on the simulation results can be ignored. Therefore, using the form of relatively simple linear contact model, the Cundall's spring - damper - sliding contact model (spring-dashpot-slider model), can satisfy the requirements of our simulation, the simulation can greatly reduce the complexity and difficulty. So Cundall spring - damper - sliding contact model (spring-dashpot-slider model) practical and simple with our simulation requirements.

3 Numerical Simulation of Single Granule Motion

Simulate the surface of single particle motion in the state of vibration feeder, it is assumed from the vibration of particles in the surface by its own gravity when the only role in the vibration of the collision contact surface, particle-particle self-gravity and collision with the vibrating surface of the contact force the role of particle collision in the contact force calculated by the discrete element contact model available. The program is based on MATLAB using C + +, compiled with basic data processing functions, text and graphics output, data storage capabilities, and animation capabilities, while the program is running can also be set to any particle parameters such as particle size, density , the initial state, vibration parameters such as surface physics surface vibration amplitude, rotational speed, kinematic parameters such as time step, contact stiffness, coefficient of restitution, particle feeding surface with friction coefficient and so on.

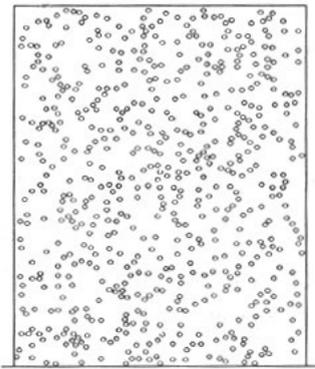
Through the development and debugging run the program, you can test discrete element spring-dashpot-slider model for contact model is analog vibration screening operations, the calculation parameters initially selected to study the movement of particles in the surface state of vibration.

4 Numerical Simulation of Granule Group

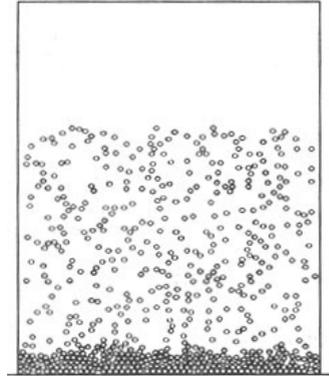
The use of discrete element simulation of particle movement in the state, the first discrete element to take into account the huge data structure, because the simulation program are the input data output data, so the particle simulation of the key point is to effectively organize discrete element of the data structure, How to read the call quickly, how best to save memory space.

Window method eliminates the need for regional law in the "grid" information data structure, where the alternative set off a neighbor element list "grid" of the data structure. Neighbor element in the second part of the list is stored in the new cell coordinates, the unit must meet with the unit particles or a similar distance from the wall equal to or less than the particle radius with distance and control; the third part of the list are in contact three particles are stored contact with the wall, including particles with the wall on the left wall, the wall on the right wall, the wall floor contact state, if the contact is stored in the left side of the wall particles and the wall, the wall on the right wall , the contact coordinates of the wall bottom plate, and stored grain with the wall on the left wall, the wall on the right wall, the wall bottom plate of the contact force, including normal contact force calculation of the tangential contact force information.

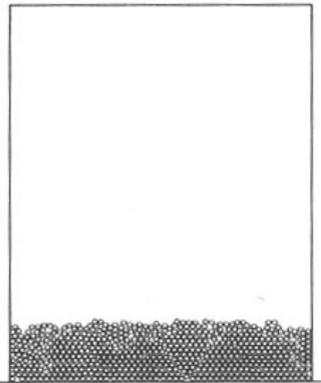
Given a vibrating box, particle number, particle size, particle simulation of the motion state, in this case the vibration of a given particle randomly generated, randomly generated particle centered coordinates, check the coordinates of each particle is generated with the previous particle coordinates repeat, if it is to lay down, if not it will add it to the list of particles until the particles reach a quantitative number, the first case we are not to load on the vibration, just set the particles are gravity, when gravity boxes distributed throughout the vibration of particles in the box slowly deposited on the bottom of vibration, the last particle deposition is completed and all the rest, we set all kinds of information at this time of particles as the initial value, and this initial value is stored in the database, and subsequent operation of the vibration load box is required before the implementation of read particles from the database of information to initialize the initial stored value of the particle array, as shown in Figure 1.



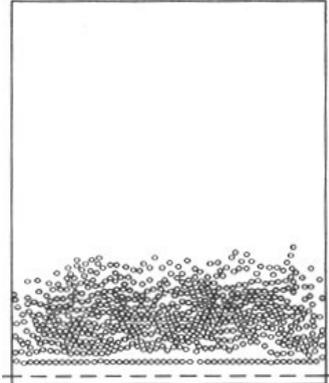
(a) Automatic generation of particles



(b) Grain weight settlement



(c) The natural accumulation of particles



(d) Vibration box loaded

Fig. 1. The initial state of particle size and the movement of vibration box loaded

Simulation of vibration of the plane motion state of particle size, particle size can be studied in the vibration plane loose, layered law, while also studying various parameters on the simulation results and to explore the parameters of different effects can improve the accuracy of the simulation summarized the method increases the time step.

5 Summary

By a brief description of the development of discrete element method, discrete element was introduced and the basic principles of common types of discrete element contact model was compared, and the linear spring - damper - sliding contact model (spring-dashpot-slider model) was decided to solve the feeder in the numerical simulation of dry granular materials problems. Simulation of vibration box loaded particle state of motion, the definition of a reflected particle Porosity expansion coefficient of H , set different parameters were investigated restitution coefficient, friction coefficient, the contact stiffness on simulation results.

Vibration simulation box particle restitution coefficient of expansion coefficient increases with increasing particle size the lower the initial stack height, increasing faster expansion coefficient; expansion coefficient increases with the friction factor showed irregular particles The initial group of stacking height is essentially the same coefficient of expansion is large, the initial accumulation of small height expansion coefficient of the simulation results into the quadratic image; expansion coefficient of contact stiffness showed a very irregular increase of the phenomenon, the contact stiffness is not only bigger and more can increase the time step, the smaller of the value of contact stiffness can increase the step of time.

References

1. Mindlin, R.D., Deresiewicz, H.: Elastic apheres in contact under varying oblique forces. *J. Appl. Mech.* 20(3), 327–344 (1953)
2. Thornton, C.: Interparticle sliding in the presence of adhesion. *J. Phys. D: Appl. Phys.* 24, 1942–1946 (1991)
3. Cundall, P.A.: A computer model for simulating progressive large scale movements in blocky system. In: Muller, L. (ed.) *Proc. Symp. Int. Soc. Rock Mechanics*, vol. (1), pp. 8–12. Balkama AA, Rotterdam (1971)
4. Mindlin, R.D., Deresiewicz, H.: Elastic apheres in contact under varying oblique forces. *J. Appl. Mech.* 20(3), 327–344 (1953)
5. Johnson, K.L., Kendall, K., Roberts, A.D.: Surface Energy and the contact of elastic solids. *Proc. R Soc. Lond. A* 324, 301–313 (1971)
6. Savkoor, A.R., Briggs, G.A.D.: The effect of tangential force on the contact of elastic solids in adhesion. *Proc. R Soc. Lond. A* 365, 103–114 (1971)
7. Oda, M., Iwashita, K., Kakiuchi, T.: Importance of particle rotation in the mechanics of granular materials. In: Behringer, R.P., Jenkins, J.T. (eds.) *Powder & Grain*, vol. 97, pp. 207–214 (1997)

8. Kishino: Computer analysis of dissipation mechanism in granular media. In: Biarrez, J., Gourves, R. (eds.) *Powders and Grains, Proc. Int. Conf. Micromech. Granular Media*, pp. 323–330. Balkema AA. Rotterdam (1989)
9. Di Renzo, A., Di Matio, F.P.: Comparison of contact-force models for the simulation of collisions in DEM - based granular flow codes. *Chemical Engineering Science* 59, 525–541 (2004)

The Research on Type Transformation of Multi-loop Spatial Linkages Kinematics Analysis

Li Wenxing and Lian Liming

Electromechanical Engineering College, Xinxiang University, HenanXingxinag, 451003
LiWenQKL@yeah.net

Abstract. To type into the institutional structure unit cell method has been widely used in motion analysis of planar linkage, which is characterized by the kinematic analysis equations to minimize the dimensions, which greatly simplifies the kinematic analysis, and the body structure, movement and dynamic analysis of unity, to facilitate the achievement of process automation. So that the analysis of institutional analysis into a group of bar, bar group apparently formed by its body than the simple, so this theory to institutional analysis greatly simplified.

Keywords: Type Transformation, Multi-loop Spatial Linkages, Kinematics Analysis.

1 Introduction

Due to Azul bar group a wide range of complex structure, especially for the senior bar group (three and three or more bar group) for motion analysis, its analytic transcendental equation for the trigonometric functions, the need to use Diego on behalf of the method. The structure of the bar group's study of planar linkage mechanism of the type transformation theory that part of the Azul bar group outside the constraint release in the internal movement of the chain outside the same number of virtual constraints, so that rod group appeared outside the chain of internal movement restrictions within the kinematic chain decomposition, senior bar group to degenerate into a series of two-bar group and the superposition of single-component form, by order of the component units to solve agencies, to the analysis of institutions.

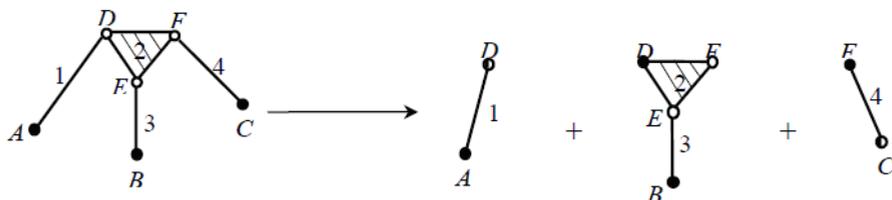


Fig. 1. Decomposition of Conversion type structure

In [1] The specific type of transformation in the use of force in particular, the analysis of motion analysis and force analysis found that the type of transformation can not achieve unity. In [2-3] The essence of the theory of type conversion depth discussions, the generalized type transformation theory, the basic idea is: in the bar group, by the number of parameters within the Deputy Vice-parameters into the outside (ie, virtual) without lifting part of the external constraint.

Group can be converted into a virtual single-component, constraints, and double-pole single-member group. For example, as shown in Figure 1, the three bar groups, using the generalized type transformation theory, including the Vice-D motion parameters on the virtual one is known, is decomposed into a virtual single-pole group member 1, two-bar group 2 and 3, bound Single Component 4.

Formulation of the theory of generalized type transformation not only of body motion analysis and force to achieve reunification, and to establish institutional structures more conducive to learning, kinematics and dynamics of the unified mathematical model and systems approach, easy to implement automated computer-aided analysis, the type transformation theory more perfect. This article will analyze the type of planar linkage and the generalized type transformation into thinking extended to multi-loop spatial kinematic analysis, given the type of transformation space decomposition of multi-loop structure of agency guidelines and structure decomposition method of automatically generated routes; space agencies with the location of single-loop analysis [4-5] and space agencies velocity and acceleration analysis of the influence coefficient method [6-8], the establishment of the movement unit type transformation model, given the space analysis algorithm for multi-loop kinematic to achieve a multi-ring analysis of the structure of space, motion analysis automation.

2 Mathematical Expression of Spatial Linkages

At the institutional analysis, although you can use the movement diagram all the necessary structural parameters that out, but it is difficult to establish mathematical relationships. Analysis of the structure introduced in the topology and the matrix, so that the analysis of research institutions to a new level, and quantitative analysis using graph theory to the loop characteristic institutions. Containing only a single hinge for the linkage, the agency into a topology diagram of the body diagram of just the components used in the top graph theory to represent, vice used to represent the edge map. But with multiple joints of the linkage, direct use of such methods for the topology graph, the graph will appear in the polygon, some theorems in graph theory can not apply. In [9] proposed to solve this problem the concept of two-color diagram, which can effectively solve with multiple joints that the issue of linkage. If the black vertex "•" that the formation of the top member; with white vertex "." Said deputy campaign to form the top, when a deputy campaign on a certain component, then each of them with an edge connecting the corresponding top to body-color topology by $G(V, V, E)$ bw (bw V, V , respectively, in the black for the graph G ,

$$m = \sum_{j=1}^{k^+} |m_j| = \sum_{j=1}^{k^-} |m_j| = \frac{1}{2} \sum_{j=1}^k |m_j| \tag{1}$$

Where k^+ , k^- - a collection of kinematic chains, respectively, ordered in the $UB \Phi > 0$ j_m and < 0 j_m the number of kinematic chains. The m value is the type of transformation of the number of institutions. The space agency, preferably by an ordered set of independent basic loop:

From the above analysis we can see: i_m both kinematic chain $UB_i \Phi$ relative to the activity of the equivalent rack, but also reflects the movement of kinematic chains $UB_i \Phi$ binding characteristics, when $i_m > 0$ the $UB_i \Phi$ join the existing two kinematic chains when a component, and the new campaign to increase chain-chain than the original i_m a sports activity; when the $UB_i \Phi = 0$ i_m already join in the kinematic chain when the two components, neither the original kinematic chain constraints do not increase increased activity; when $i_m < 0$ the $UB_i \Phi$ join in the movement chain has two components, the increase of the original i_m a constrained kinematic chain, and the new movement to reduce chain-chain than the original i_m a sports activity. Agency corresponding to the same multi-loop, often multiple, independent combination of the loop group, is set to n , i -independent of the loop group, taking the ordered loop corresponding to different groups, m values may be different. But not the choice of an ordered set of independent through the loop m above is calculated by multi-loop transfer space agencies, but with that of the decomposition of n kinds of programs in the corresponding m minimum, it is called the machine the number of institutions of transformation, denoted m is a measure of the complexity of multi-loop kinematic an indicator, from the institutional process of the coupling between loops.

3.1 Structural Decomposition and Automatic Generation of Optimal Routes

M value depends not only on each individual the loop itself, but also on their composition and order, so when the preferred order loop group, both to make the selected loop $B_j L$ corresponding to the $j_m UB_j \Phi$ minimum, but also the rest of the loop to minimize the coupling between, which is present to break down the remaining part of the loop to provide as many known variables, so that at least some of the remaining loop is unknown, namely ordered ring Road group preferred the principle of BL. As is the corresponding LB and ΦUB , so the process of optimization of BL is the process of selection ΦUB , but also determine the institutional structure of the process of decomposition route.

- (1) Find the body to form the shortest Pap truss loop (group);
- (2) Election (1) find the shortest loop (group) in the loop composed of the vice of each component and the largest number of the loop group
- (3) Preferred the rack containing the loop as the first loop LB1, and preferred the original frame by frame, as calculated, such as (1) not obtained in the group with the rack loop the loop, Choose one from the First Ring Road as the first loop B1 L, and the largest number of pole selection as vice calculation rack.



3.2 Ordered to Order Cell Decomposition of Kinematic Chains

Body was decomposed by a set of ordered structure of kinematic chains $UB_j \Phi$, according to $UB_j \Phi$ values corresponding to the j_m can be ΦUB_j divided into three categories: ① $m_j > 0$ the ΦUB_j ; ② $m_j = 0$ for ΦUB_j ; ③ $m_j < 0$ ΦUB_j . For $m_j = 0$ of the kinematic chain ΦUB_j , because of its activity is 0, so the movement to determine if the exercise of its analysis, the kinematic chain can be derived for all motion parameters, type conversion will be referred to as the basic unit. For $m_j > 0$ in the kinematic chain ΦUB_j , when $m_j = 1$, if both ends of the movement kinematic chain ΦUB_j Deputy Vice-turn or move for the pair, kinematic chain ΦUB_j will be further decomposed into a virtual unit.

Basic units; Otherwise, the kinematic chain $UB_j \Phi$ is the type of transformation of the virtual unit, the rotational parameters as vice cylindrical dummy variables.

When $m_j < 0$, if the kinematic analysis of the kinematic chain, the kinematic chain will appear on the compatibility condition j_m a movement, called constraint-based unit conversion.

3.3 Structural Decomposition Example

Example 1 (a) shows the body line of the structure decomposition below. Body is first broken down to the rack, active pieces, $1 m_1 =$ the kinematic chain ΦUB_1 (RSSR), $1 m_2 = -1$ the kinematic chain ΦUB_2 (R SSR); and then $1 m =$ of ΦUB_1 (RSSR) and further divided into virtual RS and the unit cell SSR, while the ΦUB_2 is the binding unit.

Example 2 (a) shows body, the structural decomposition line in (b) shows, the body is decomposed into the rack, active pieces, $1 m_1 =$ the kinematic chain ΦUB_1 (SCC), $1 m_2 = -1$ in the kinematic chain ΦUB_2 (SRR), kinematic chain ΦUB_1 and ΦUB_2 virtual unit and were bound unit.

4 Space Agency Type Transformation Motion Analysis Algorithm

For $m = 0$ weak coupling bodies, institutions broken down into the rack, active components and the basic unit of the first order kinematic analysis of active components, and in turn the unit on the analysis is complete.

For $m \geq 1$ the strong coupling body, the analysis The idea is: for $> 0 m_j$ virtual unit, a component of its m_j (or motion pair) in the exercise position (or velocity, acceleration) parameter to the virtual assignment, then Total m a dummy variable, in accordance with the institutional structure decomposition of the positive sequence line, the movement of each unit can be transformed in order to solve alone, but the dummy variable unit of assignment should meet the movement restriction conditions of compatibility, compatibility to meet all the m -sport equation of the dummy variables assigned variables both for the true value of these movements, then, institutional variables may be movement of the other components that facilitate the variable m to get a movement.

Position Analysis. Institutions were established one active input f_i ; m respectively in the virtual unit to take on the virtual value. Unit established in the constraint equations of compatibility: If f_i described as non-negative equation, an arbitrary assignment, the corresponding location of the binding unit may not be able to meet the compatibility condition, that is, when the assignment met, is the true value of the virtual components, so the equation for the position of institutions by the above analysis, solving the virtual component of the process is actually the real value of the $\{\Phi\}$ as variables for the optimization objective function optimization problem.

For $m = 1$ in the body, the position of the dimension of equations is 1, can be obtained through one-dimensional search for the location of institutions solutions.

For $m > 1$ body position analysis, can be viewed as a multi-objective optimization, the constraint functions of the same nature, it can be taken.

Linear weighting method, the weights are set to 1, they are combined into a composite objective function, that is then modified by optimizing the value of $\{\Phi^*\}$, to approach the true value $\{\Phi\}$, to meet the requirements. Need to note meaning that any given set of values $\{\Phi^*\}$, the objective function value of each unit in the order received after solving constraints on a single component of the objective function and a single.

5 Summary

In this paper planar linkage kinematic analysis of the type of transformation and expansion of the generalized type transformation theory to the multi-loop spatial kinematic analysis, conversion of units given to space agencies for the multi-ring structure units The unit method; derived into position analysis algorithm of this module, and then export the transformation space agency position, velocity and acceleration analysis formula; shows the location of the computer automatically solve the optimization method and the velocity and acceleration of the two times the virtual method. Multi-loop spatial kinematic analysis transformation has the following characteristics: (1) position of established institutions of the lowest dimension equations, easily solved, and the elimination method, continuous method compared with the obvious advantages; (2) body velocity analysis, acceleration analysis was explicitly expressed, and institutions to facilitate the process of unifying and simplifying the dynamic analysis of multi-loop dynamic analysis of space agencies; (3) The equation of motion analysis and solution to facilitate the computer automatically generated.

References

1. Prentis, J.M.: The Pole Triangle, Burmester Theory and Order and Branching Problems-I the Order Problem. MMT 26(1), 19–30 (1991)
2. Prentis, J.M.: The Pole Triangle, Burmester Theory and Order and Branching Problems-II the Branching Problem. MMT 26(1), 31–39 (1991)
3. Deckhardt, H.: Kinematic Design of Machines and Mechanisms. McGraw-Hill Education (Asia) Co. and China Machine Press (March 2002)

4. Dhingra, K., Cheng, J.C., Kokli, D.: Synthesis of six-link, Slider-crank and Four Link Mechanisms for Function, Path and Motion Generation Using Homotopy with m-homogenization. *ASME Journal of Mechanical Design* 116(4), 1122–1131 (1994)
5. Foster, D.E., Cipra, R.J.: Assembly Configurations and Branches of Planar Single-Input Dyadic Mechanisms. In: *Proceedings of 1996 ASME Design Engineering Technical Conferences and Computers in Engineering Conference*, Irvine, California (1996); 1996-DETC/Mech-1143
6. Foster, D.E., Cipra, R.J.: Assembly configurations and Branches of Planar Single-Input Dyadic Mechanisms. *ASME Journal of Mechanical Design* 120(3), 381–386 (1998)
7. Foster, D.E., Cipra, R.J.: Assembly Configurations and Branches of Planar Non-Single-Input Dyadic Mechanisms. In: *CD-ROM Proceedings of 1997 ASME Design Engineering Technical Conferences*, Sacramento, CA; Paper No. DAC-3858 (1997)
8. Foster, D.E., Cipra, R.J.: Assembly configurations and Single-loop Mechanisms with Pin Joints and Sliding Joint. *ASME Journal of Mechanical Design* 120(3), 387–391 (1998)
9. Foster, D.E., Cipra, R.J.: An Automatic Method for Finding the Assembly Configurations of Planar Non-Single-Input-Dyadic Mechanisms. *ASME Journal of Mechanical Design* 124(1), 58–67 (2002)

Semi-supervised Data Stream Ensemble Classifiers Algorithm Based on Cluster Assumption

Wang Xuejun

Chengde Petroleum College, Chengde, Hebei, China
jxn_n_a@sina.com

Abstract. Semi-supervised data stream ensemble classifiers algorithm based on cluster assumption was proposed. Although traditional semi-supervised classification algorithm can solve incomplete label data sets classification problem, but it is an unsolved problem that how to use it in data stream environment and how to improve semi-supervised classification algorithm accuracy by using data stream characters. According to analyzing generalization of semi-supervised classifier based on cluster assumption, it indicates that increasing labeled data during training moment can improve semi-supervised classifier accuracy. Making use of this conclusion, a semi-supervised data stream ensemble classifiers algorithm based on cluster assumption was proposed.

Keywords: Data Stream, Ensemble Classifiers Algorithm, Cluster Assumption.

1 Introduction

In the framework of supervised learning, the use of classification algorithms to solve such problems can be divided into two stages: training stage and classification stage. Firstly, the label data set according to the classification model is trained; then use the trained classifier on the unknown class labels for data classification. Therefore, the classification model and use must be labeled with a certain amount of data. However, in many practical applications, the data flow is often not having enough of class labels. First, the data stream is a continuous flow of data ordered sequence, so the artificial 24-hour mark class labels difficult to achieve; Secondly, the manual tagging category label takes some time, in the limited time, the class label number of tags is difficult to meet the needs of the training classifier; Finally, the data flow data with timeliness, when the concept drift occurs when a label marked before the data can not be involved in classifier training[1-3].

2 Semi-supervised Learning Algorithm

2.1 Self-training Models

Self-training Models is a commonly used semi-supervised learning method. The supervised learning method based on labeled data by training the first classifier, the classification of unlabeled data, and then to choose the part of the classification of high confidence data, its classification as a category label, and a joint re-tag data trained

classifier, then the classification of unlabeled data, repeated iteration gives the final classification model. Since the training model approach has the advantage of simple, intuitive and "filter", that classification algorithm model itself on the use of no special restrictions, you can use a simple k-NN algorithm can also use the more complex classification algorithms, such as support vector machines , decision tree and neural network algorithms. The disadvantage is that the early classification errors, the algorithm will be continued to enlarge in the iteration process, ie the introduction of noise in the training process and thus affect the classification of data and produce a vicious circle [4].

Propagating 1-Nearest-Neighbor, PNN is a typical model of self-training method, the algorithm to 1-NN algorithm based algorithm, algorithm specific training as follows:

Algorithm 1.

Input: labeled data set $\{(x_i, y_i)\}_{i=1}^l$, no label data set $\{x_i\}_{i=l+1}^{l+u}$, the distance function

1. set $L = \{(x_i, y_i)\}_{i=1}^l, U = \{x_i\}_{i=l+1}^{l+u}$.
2. while $U \neq \emptyset$
3. selected x from U , making them $x = \arg \min d(x, \tilde{x})$;
4. set x class label $f(x) = \tilde{y}$;
5. added (x, \tilde{y}) to L , removed from U .

2.2 Production Model

Based on production methods (Mixture Models) may be the first semi-supervised learning methods. Located in any instance of an instance space, the maximum value when the conditional probability $P(y_j | x)$, the predicted class label. Clearly for all possible classes, to meet $P(y_i | x) \in [0, 1]$ and $\sum_{y_i} P(y_i | x) = 1$. Inability to effectively calculate directly, so scholars are using Bayesian theory and generate indirect estimates of its type models. Bayesian formula is as follows:

$$P(y_j | x) = \frac{P(x | y_j)P(y_j)}{\sum_{y_i} P(x | y_i)P(y_i)}$$

For any of the categories, the value of the denominator of the formula the same, so only two can be considered and. Based on production methods to meet a particular distribution assumptions, such as image classification, the assumed Gaussian distribution to meet the many variables that

$$P(x | y_j) = N(x; \mu_{y_j}, \Sigma_{y_j}) = \frac{1}{(2\pi)^{D/2} |\Sigma_{y_j}|^{1/2}} \exp(-\frac{1}{2}(x - \mu_{y_j})^T \Sigma_{y_j}^{-1} (x - \mu_{y_j}))$$



And that the mean vector and where the covariance matrix. So the question into how the training data set and the effective estimate. Which the first of which is a specific distribution with parameters such as mean vector and covariance matrix. Then one, the number of categories when the data set is required when a parameter, namely, when the probability that we use the variable and all parameters, and the use and alternative, so the purpose of training is to find a "good" variables value, and thus be estimated. When all the data of the class labels are known, based on maximum likelihood that the compute the value of the maximum value of the following formula is the best [5].

$$\hat{\theta} = \arg \max_{\theta} P(D | \theta) = \arg \max_{\theta} \log P(D | \theta)$$

Which indicated that the training data set. However, the training data set as part of the data class label is unknown,

2.3 Collaborative Training

Cooperative Training (Co-Training, CT) by Blum and Mitchell proposed in 1998 a classic semi-supervised learning methods. This method assumes that the data set with two separate views, each view can be independent while training a classifier. For example, in web application in, you can view web content using the training classifier, hyperlinks can also use the web view of the classifier training. For joint training [6], to any high degree of confidence classifiers with unlabeled data classification to another classification filled with the training data set to add another classification of the next round of training data set size. The specific steps are as follows:

Algorithm 2.

Input: labeled data set $L = \{(x_i, y_i)\}_{i=1}^l$, no label data set $U = \{x_i\}_{i=l+1}^{l+u}$, each with two views of data $x_i = [x_i^{(1)}, x_i^{(2)}]$, constant k .

1. According to data view and data sets will be divided into two labeled training data set;
2. Cycle repeated until the;
3. According to the training data set classifier, classifier based on training data set.
4. Use of classification and separate the remaining unlabeled data classification.
5. Classification results in the confidence to join the largest data set of data, the same confidence classification results in the largest data set of data added, removed from the data.

Blum and Mitchell's collaborative training algorithm, the data set has two independent natural view of the classifier can be trained, but most of the data set and can not meet those requirements. Nigam and Ghain therefore proposed that when the data set has a large property, data sets can be randomly divided into two views. But with a large number of property conditions are relatively harsh, so Goldman and Zhou further proposes that, from the same data set to train two different classifiers for joint training, can effectively avoid the above requirements [7].

2.4 Way of Learning Based on Clustering Assumptions and Popular Assumptions

In general, label-free as long as a reasonable target data and the links between learning, semi-supervised learning classification algorithm can effectively improve the classification performance. For example, the generative model is to use unlabeled data to estimate model parameters on the distribution, in order to better classification of the division plane. Assumptions and popular assumptions based on clustering semi-supervised learning method is to rely on the data distribution is another semi-supervised learning to establish contact method.

The so-called cluster assumption is that the data within the same cluster have the same class label is more likely. Under this assumption, comparing the data dense regions in the same cluster after clustering within the data more likely to be given the same so the possibility of greater class labels; In contrast, in data sparse areas, data clustering After the cluster is more likely different, so the region between the data labels are not necessarily the same class. When the class label is not the same hyperplane will pass through from between the various clusters. Therefore clustering assumptions, unlabeled data is to assist regional and proven data sparse regions dense, so that as much as possible through hyperplane data sparse areas.

Popular assumption is that in a small area of local clinical data with similar class labels. The so-called pop, is in the topological space, for arbitrary, existing neighborhood and integer such that embryos with the same Euclidean space. On this assumption, the role of a large number of unlabeled data is to make the data space is more dense, thus more accurate data set describe characteristics of the local area. And clustering hypothesis focuses on the overall distribution of the data is different from popular pay more attention to the data assuming the local characteristics.

3 Semi-supervised Data Stream Ensemble Classifiers Algorithm Based on Cluster Assumption

This section based on clustering semi-supervised learning algorithm assumption of starting a label data set first analyzed semi-supervised classification algorithms of classification error, and then integrated classifier and the characteristics of data flow is presented based on data clustering assumptions current semi-supervised classification algorithm integration.

Suppose a data set from the unknown distribution and composition, in which. For any one data, the probability distribution described by the form in which the probability density that the data that the return on the value of the data, that is. The purpose of semi-supervised classification algorithm through the training data set and to construct a classification function, so that the average error, and even if the minimum value of the following formula:

$$R(g) = P(g(X) \neq Y)$$

However, I do not know the true distribution of the data subject, can not be directly calculated value, so this classification algorithm by measuring the training data set of unlabeled data classification error, an indirect analysis of semi-supervised classification function of the classification performance.

Definition 1 cluster hypothesis: in the Class 2 data set, data class label is 1 or 0, the mind is a set of clustering clusters, the function taking constant values on each. Where is the indicator function, when the condition is true in brackets is 1, and 0 otherwise.

According to definition, the clustering assumptions, any data within the same cluster, the value or greater than $1/2$, which are less than $1/2$.

Data set a certain way clustering, clustering clusters can be divided into two parts: one is the cluster containing the label data, calling it a label clustering clusters, the other is the data does not contain a label, called no label clustering clusters. With that cluster all the cluster composed of a collection of tags, no tags with that all clusters of the set cluster, which indicated that complement the collection.

In the clustering assumptions, for the collection and the unlabeled data, semi-supervised classification algorithm for the classification process is not the same. Let be an arbitrary set of unlabeled data, according to clustering assumptions, we can use cluster labeled data to classify them by voting method. For an arbitrary set of unlabeled data, since there is no any of its assumptions, it can only use supervised learning algorithm, using the data set all tag data to establish classification model to classify them. The following data sets were considering the size of the label on the two types of cluster cluster unlabeled data classification error.

3.1 Integration of Ordinary Semi-supervised Classification of Data Streams

This paper integrates the data flow classification framework, presents a remarkable integration of semi-supervised classification of data streams, pseudo-code is described as follows:

Algorithm 3.

Type: semi-supervised ensemble classifier E , weight W , the maximum number of individual classifiers n , the most recently arrived data block

Output: A new semi-supervised ensemble classifier \tilde{E} , the value of \tilde{W}

1. Use of an integrated semi-supervised classification and weight of the classification;
2. When the real part of the data obtained in the class label, semi-supervised classification algorithm using the training data block to obtain a new individual classifiers;
3. When the concept drift occurs, remove the integrated classifier classification of all individuals;
4. When the semi-supervised classification of the remaining integration of the individual classifiers is less than, the, or delete individual in the initial training and then joined the classifier;
5. Use of individual classifiers in the classification accuracy rate on the new weights.

3.2 Semi-supervised Classification of Data Flow Integration

TSSEC algorithm can handle data streams Although semi-supervised classification problem, but did not take full advantage of the characteristics of the data stream. According to the foregoing analysis, the semi-supervised classification algorithm classification error with a label data set increases. Therefore, full use of the data stream based on the label data, this paper presents a new semi-supervised clustering based on integrated data flow classification algorithm.

Shown in Figure 1, set by the continued flow of data into the data segment, each data segment contains data sets and in two parts. Which indicates that the data segment was randomly selected set of data marked class labels that are not marked in the data segment of data collection. SSDSEC algorithm uses the data segment in recent months with the label data set has a label composed of a data pool. When training new arrivals segment, the first use of the label data pool combined with the raw data segment a new set of training data, and then training of the new semi-supervised training data set to obtain the individual classifiers.

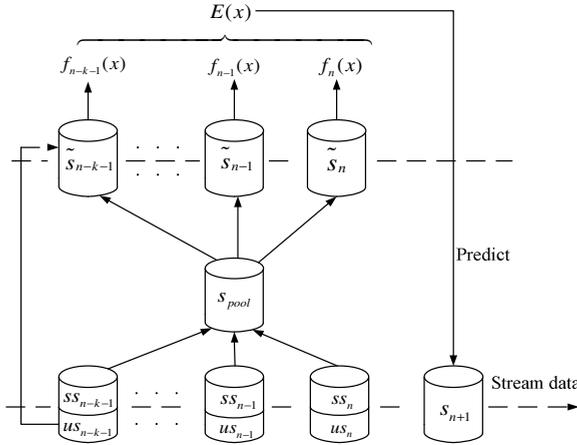


Fig. 1. Semi-supervised Data Stream Ensemble Classifier

We know that the individual classifier accuracy the higher the difference the greater the integration the better the classifier. Individuals in the training of semi-supervised classification, due to the increased concentration of labeled data training data set size, making the individual classifier accuracy improved significantly. At the same time assuming the semi-supervised clustering-based classification of learning, the role of a large number of unlabeled data is to help the data proved the distribution of feature space dense and sparse regions in order to guide the learning algorithm as the sparse data distribution through the region [8]. Therefore, the training data set as unlabeled data are different, semi-supervised classification of individual significant differences still exist between.

4 Simulation and Results Analysis

In order to verify SSDSEC algorithm, this paper in the hyper-plane artificial data sets and Magic Gamma Telescope standard data sets SSDSEC algorithm and TSSEC algorithms comparative experiments. In order to measure the SSDSEC algorithm clearly labeled pool of data, TSSEC algorithm of assigning weights of individual



classification methods and integrated classifier SSDSEC algorithm updates the same way. Experimental environment: windows XP operating system, CPU Pentium IV3.0Ghz, Memory 512M, algorithm matlab 7 writing, the individual classifiers used in [7] proposed semi-supervised classification algorithms.

Hyper-plane experimental data set parameters: the data block size, keep the number of individual classifiers, attribute dimension is 5, the proportion of randomly labeled data labels. Figure 2 is the concept of a smooth, SSDSEC algorithm and the classification error rate TSSEC algorithm comparison. Figure 3 for the two algorithms when the concept drift error rate comparison of classification, in order to simulate concept drift, data generator generates 20 data blocks each time the value changes. All data for the figure 30 times the average of duplicate runs.

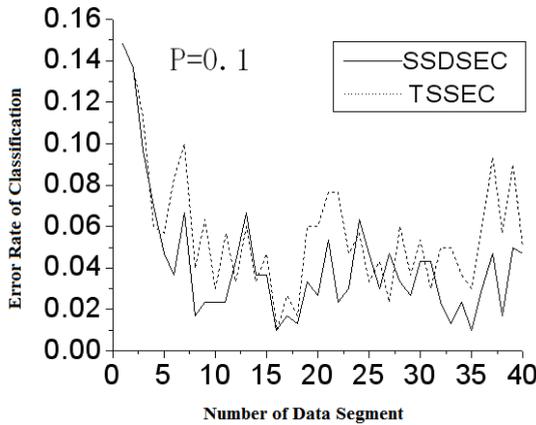


Fig. 2. Compare classification error between two algorithm without concept drift

Figure 2 and Figure 3 show, SSDSEC algorithm classification error rate was less than TSSEC algorithm. In the beginning of the experiment stage, the two algorithms classification error rate was relatively large, with the increasing flow of data blocks, the two algorithms significantly reduce the classification error rate. The reason is that the beginning of the influx of experimental data blocks less training in the use of individual data blocks less classification, classification accuracy is relatively low, with the Geti classifier growing, integrated classifier accuracy improved significantly. Figure 3 is the concept of drift classification error rate of two algorithms compared can be seen from Figure 3 in the first data blocks 20,40, the two algorithms of the classification error rates are suddenly increased, then rapidly shrinking because of the way the data generated two conceptual drift, use the "old concept of" individual classifier generated can not necessarily accurate classification of the newly generated data, but the use of updated guidelines for integrated classifier to remove "outdated" individual classifiers, so the integration of the two algorithms generate classifiers can quickly apply the concept of the new data.

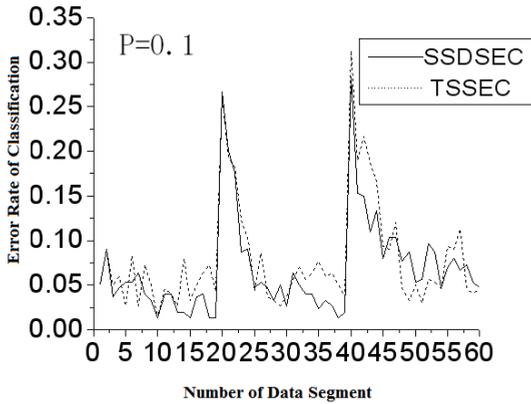


Fig. 3. Compare classification error between two algorithm with concept drift

Magic Gamma Telescope data set from the standard database UCI, a total of 19 020 data, the purpose is to determine the cosmic rays from gamma rays. After the mixed data set composed of random sampling of data blocks. Experimental parameters: the data block size, keep the number of individual classifiers, the proportion of randomly labeled data class object. It can be seen from Figure 4, the standard data set, SSDSEC classification error rate algorithm not only significantly smaller than TSSEC algorithms, classification performance is also more stable (less variance), because the data pool of data caused by weakening of the sample data randomness. Although the data obtained a sample of blocks of time may deviate significantly from the original data distribution, data pool their data will also be amended to reduce the training of the individual data block classifier error variance.

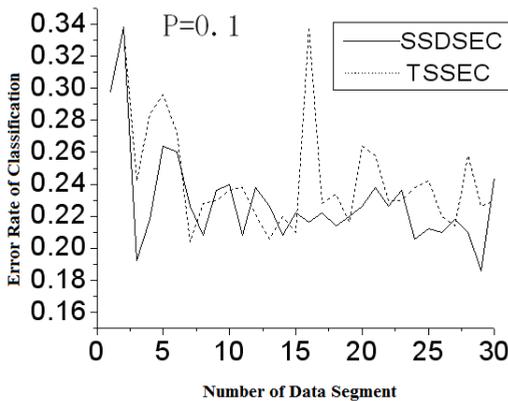


Fig. 4. Compare classification error between two algorithm on UCI data set



5 Conclusions

Based on the analysis of data sets labeled cluster-based semi-supervised classification algorithms assume that the impact of classification error starting the data stream is proposed integration of semi-supervised classification algorithm. The algorithm in training the individual classifiers to take full advantage of the characteristics of data flow, increasing the training data set the ratio of labeled data to improve the collection of semi-supervised classification accuracy device, while under the no-label data in the semi-supervised classification algorithm role, that remains between the individual classifiers are quite different. Finally, the artificial data sets and data sets will be compared with the TSSEC algorithm Experimental results show that the algorithm improves the classification accuracy, but also increased the stability of classification algorithms.

References

1. Tsoumakas, G., Vlahavas, I.: Random k -Labelsets: An Ensemble Method for Multilabel Classification. In: Kok, J.N., Koronacki, J., Lopez de Mantaras, R., Matwin, S., Mladenič, D., Skowron, A. (eds.) ECML 2007. LNCS (LNAI), vol. 4701, pp. 406–417. Springer, Heidelberg (2007)
2. Zhou, Z.H., Wu, J., Tang, W.: Ensembling neural networks: Many could be better than all. *Artificial Intelligence* 137(1-2), 239–263 (2002)
3. Shen, D., Yang, Q., Chen, Z.: Noise reduction through summarization for web-page classification. *Information Processing and Management* 43(6), 173–174 (2007)
4. Wu, X.D., Zhu, X.Q.: Mining with noise knowledge: error aware data mining. *IEEE Transactions on Systems Man and Cybernetics* 38(4), 917–932 (2008)
5. Zhu, X.Q., Wu, X.D., Chen, Q.J.: Eliminating class noise in large datasets. In: Proceedings of the Twentieth International Conference on Machine Learning, Washington DC, USA, pp. 920–927 (2003)
6. Sofie, V., Anneleen, V.A.: Ensemble methods for noise elimination in classification problems. In: Fourth Workshop on Multiple Classifier Systems, United Kingdom, pp. 317–325 (2003)
7. Tumer, K., Ghosh, J.: Analysis of decision boundaries in linearly combined neural classifiers. *Pattern Recognition* 29(2) (1996)
8. Rigollet, P.: Generalization error bounds in semi-supervised classification under the cluster assumption. *The Journal of Machine Learning Research* 8, 1369–1392 (2007)

Object-Oriented Design and Implementation of Embedded System IDE Software

Guo Wushi and Yi Xin

Sichuan Engineering Technical College, Sichuan, China
wwseaww@sina.cn

Abstract. The rapid development of embedded hardwares makes the proportion of software design increase gradually. This paper proposed one method of desining embedded software IDE based on object oriented technology because of the shortcomings of current IDE. Software desingers can use these reusable components modules to infrastructure software system which meets the hardware system. This greatly improves the efficiency and quality of the development of software.

Keywords: Object-Oriented, Embedded software IDE, Cross Compile, Kernel Porting.

1 Introduction

Highly integrated microprocessors, operating systems and the introduction of embedded system applications require people to improve, leading to the development of embedded system software design of the difficulty of unprecedented complexity. From the software architecture, operating systems and middleware such as the complexity of multiple levels plagued software developers. Ideally, developers do not have to worry about the next operating system, function integration and software platform, but rather to focus on product differentiation and competitive advantage. But the reality is that most of the energy device software developers spend the tools, methods, and integration aspects. Basically, the developer has no way to reuse the design of the original design results and intellectual property rights [1]. Embedded software is limited in concept to the major operating systems and development tools.

The current implementation of object-oriented software technology is modular, the best way to improve software reusability. In this paper, object-oriented technology to design an embedded software IDE, the development of a common model architecture, hardware configuration automatically generated according to the basic system software programs. Only need the system for each object class attributes and operation set to automatically build a cross compiler environment, reducing the reduction of the operating system and the difficulty of migration, resulting in the basic framework of the driver. Developers only need to configure the basic information of each module, will be more focus on the system's applications[2].

2 Object-Oriented Analysis and Design of IDE

Embedded Linux operating system in embedded development to occupy the mainstream, this article focus on a Linux operating system for portable embedded development process for discussion. According to their development process[3], the use of object-oriented technology to design six basic categories. These basic class encapsulates the details of a large number of complex, embedded software development through the important part.

Hardware system information class, bearing the description of the hardware system, the user by setting the parameters of other objects for which the information will be configured.

Environment to build cross compiler chain management class source code, according to the information architecture automatically run a script directory in the development of cross compiler generates a complete chain, which is the basis for embedded development.

Project management class, integrated development environment is the main purpose of the unified organization of the document, save labor management, automatic compile and run the source code, which contains the file operations and organize the relationship between source code, compiled using make technology, involve a large number of global environment parameter.

Control the font type as processor family includes the common control character register, simply select a specific function, can generate source code need instructions. The same driver development, but also frequent use of the c language on the register operation.

System Migration category covers the entire system needs to transplant the module, the boot loader, operating system kernel and file system.

The most important feature of object-oriented is "high cohesion and low coupling", each object encapsulates the details of the operation, only a limited exchange of information on the external interface, such flexible thinking improve the maximum extent of software reusability, This design IDE, according to the different architecture, only the individual objects need to modify the internal data and operations, without any impact on the overall structure [2-3].

3 System Mode

Integrated development environment running on Windows platform, but requires a powerful Linux command set, in this, this paper presents techniques to simulate the Cygwin Linux operating environment, which makes the underlying by the powerful IDE command support, you can run a variety of script file.

Cross compiler tool chain design module is the core of IDE, this cross-compiler tool-chain module based on the GNU GCC tool chain available, not a separate structure for a system, you can redirect based on configuration parameters to generate a series of compilation tools . Management by means of engineering source code, is the IDE design of the bridge. Makefile use this technology to manage the whole project, mainly drivers and applications for the organization and compilation. Cygwin

integrated Gnumake, so long as the preparation of a good Makefile, you can complete compilation of all the connections work.

A key technique is to display information redirected to the IDE output window specified, all information is the standard script console output, so you need to compile the standard information and error messages redirect. Usually by pipeline technology to solve this problem, but the embedded development process of many long hours compiling the information end of the process often required to be fully displayed. This IDE is used to read asynchronous pipeline technique and technology, makes the redirection and information display at the same time, timely feedback in the process of compiling a variety of information.

Bootloader is a boot loader for embedded systems, its role is to initialize the necessary hardware equipment, build the kernel needs some information and the information is passed to the kernel through the relevant mechanisms, the system software and hardware to bring a suitable environment state, the final call to the operating system kernel, really played the role of guide and load the kernel [5].

Bootloader depends on the hardware implementation, in addition to architecture, but also depends on the specific configuration of the embedded board-level device, so Bootloader is embedded developers the most difficult part of the transplant, in general, is to modify existing code. In this paper, U-Boot 1.2.0, is more common for the Bootloader.

To the maximum extent possible to help users complete the code modifications, IDE to provide different types of architecture need to modify the source of the message, and the need to modify some configuration message, the need for a help file database. Add needed module by controlling the source font. Most important thing is to modify two aspects, one is the manner in which start, if it is Nandflash start, need to define the chip configuration function, increased Nandflash to copy data from the memory code. In addition, tft agreement for the use of burning core to the flash, need to be modified in response to a chip card.

In short, the kernel configuration is nothing more than a configuration file generated by the Makefile to read, different architectures, it will choose a different source directory structure. However, relatively complicated kernel configuration, difficult to sort out the general staff source the relationship between the configuration is not good, not only does not compile, even if reluctantly compiled successfully, it may not be able to function properly. Moreover, the kernel is not just a configuration problem, part of the source must be modified as needed, such as memory Nandflash if it is necessary to establish Nandflash partition table[4-5].

For the kernel configuration, this IDE provides a template to use a second configuration, create the template database, the user's parameter settings, to provide qualified template, if the basic template to meet the requirements, you can use the template to compile details of the changes if needed, is called the kernel configuration tool to configure the kernel using make xconfig configuration tools in the form, produce a graphical interface, of course, require the support of X system, you need to install Cygwin, modified.

Root file system is a Linux / Unix system starts an important part of the operating system to work is a necessary part of the time, when the kernel needs to start with the file system to mount. In modern Linux operating system, kernel code image file in the

root file, the system will boot process equipment manufacturers from the root to the kernel code is loaded into memory to run.

Embedded root file system has RomFS, JFFS2, NFS, EXT2, RAMDISK, Cramfs etc., each system has its own characteristics, the most commonly used is Cramfs file system, Cygwin toolkit provides Cramfs tools, mkcramfs.

The bottom of the embedded root file directory, have their specific uses and purposes, it is necessary to establish the most basic root. And some of the root directory of the file must be fixed, such as / bin, / sbin, / usr / bin, / usr / sbin need to put Linux command tool, / etc directory requires system configuration files, including startup files. Therefore, this module to complete the following functions: to provide the necessary root, root directory files, the root made into a root file system, this temporary provision Cramfs file system. Embedded Linux command tools provided by the BusyBox tool, it contains a lot of standard Linux tools to achieve a single executable, this paper provides a BusyBox IDE standard configuration, only need to use the cross tool chain for compiling shown as Fig.1.

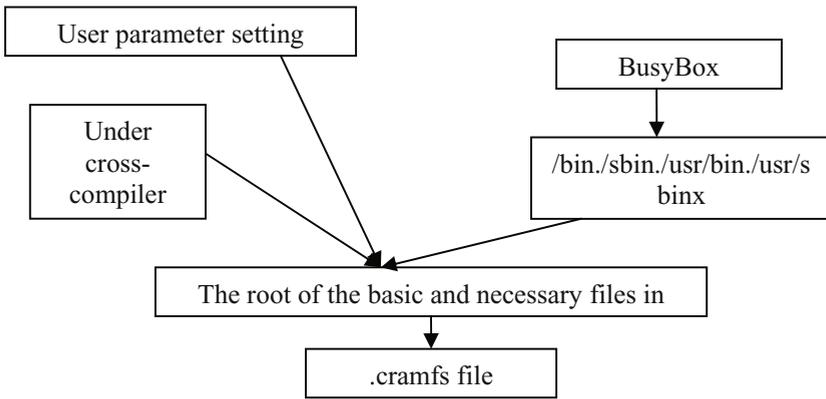


Fig. 1. Module of boot files making

Driver modules and application programming modules as long as the preparation of a text editor, the use of project management to organize the source code to compile some. The focus here to talk about writing drivers, Linux drivers have a fixed pattern, such as character, block-based, network-based data structures are different, but each type of driver architecture are the same. Another driver is the most important is the preparation of the control word read and write, even if the developer under the control of a system of very familiar characters, but also frequent reference manual to the accuracy of [4].

This article was written proposed driver model is to choose a type of automatic generation of some type of template, add this template to change; the second is to use the control word module, the driver in the development process, developers can choose under a system structure for the operation of a specific register, and then automatically generates the necessary code, simply embed the code to the driver



source code. If the driver is going to be added to the kernel, need to reconfigure the kernel.

Finally, the environment of embedded software IDE for ARM 2410 series platform for the U-Boot, Linux2.6 kernel, root file system configuration and migration, the preparation of a basic network card driver, the test is successful.

4 Summary

In this paper, the lack of embedded software development tools and the lack of object-oriented technology to design an integrated development environment, developers use a combination of the modules, build their own system software. Object-oriented technology increases the low coupling IDE reusability; and the high degree of cohesion, then make the developers do not care about the internal details of the complex will be more energy to focus on application logic. Unified development environment, eliminating the compatibility issues between the last tool. Currently, the IDE architecture also slightly rough, technology is not perfect, to be further addressed.

References

1. Li, Y., Ou, W.: ARM embedded Linux system development, vol. 1-4, pp. 34–35. Tsinghua University Press, Beijing (2007)
2. Zhu, C.: Object-oriented embedded system development, pp. 25–30. Beijing University of Aeronautics and Astronautics Press, Beijing (2007)
3. Lan, G.: The principle of object-oriented methods and practice, 3rd edn. Mechanical Industry Press, Beijing (2003); Shan, Y.S. (translation)
4. Sun, C.: Embedded Linux driver development guide the design. Electronic Industry Press, Beijing (2005)
5. Simon, D.E., Chen, X. (translation): Embedded system software tutorial. Mechanical Industry Press, Beijing (2005)

Development and Application of Cloud Computing in China Information Resource Management Field

Pang Jingxia

Lenovo (Beijing) Co. Ltd., China

ji.kam@qq.com

Abstract. Cloud computing has a broad application prospect, with information resource management as its core issue. This paper introduces the cloud computing performance from the perspective of information resource management, and discusses how to use cloud computing to improve the sharing and building ability of information resources. At the same time it also analyzes the influence and reformation of cloud computing on information resource management, and puts forward the basic application principles of cloud computing in information resource management, providing new approach for the application of cloud computing in information resource management.

Keywords: Information resource management, Application, Development, Cloud computing.

1 Introduction

With the development of the technologies, e.g. multinuclear processor, virtualization, distributed storage, broadband internet and automated management, a new style computing paradigm-cloud computing emerges, which is able to arrange computing resources as required, while the users only need to pay for the used resources [1].Essentially cloud computing refers to that the user terminals obtain the computing resources, e.g. storage, calculation and database, etc. through long-distance connection. Since the concept of cloud computing is put forward, cloud computing has become the hot topic in computer science and network field, and triggered a series of exploration and revolution. Compared with the concept of grid computing put forward previously, cloud computing is not only the conceptual discussion, it places more on realization [2,3].It not only changes the usage mode of computers, but also fundamentally changes the mode of information acquisition, information storage and information communication.

The core of cloud computing is providing users with the services in the form of renting computing resources after the virtualization of computing resources in a particular or several data centers [4].Briefly, it gathers the computers located in the same place or different places to provide various services to enterprises or personal users. There are three basic forms of cloud computing service, i.e. infrastructure as a service (IAAS), platform as a service (PAAS) and software as a service (SAAS), forming the pyramid of cloud computing from the bottom up (shown in Figure 1).

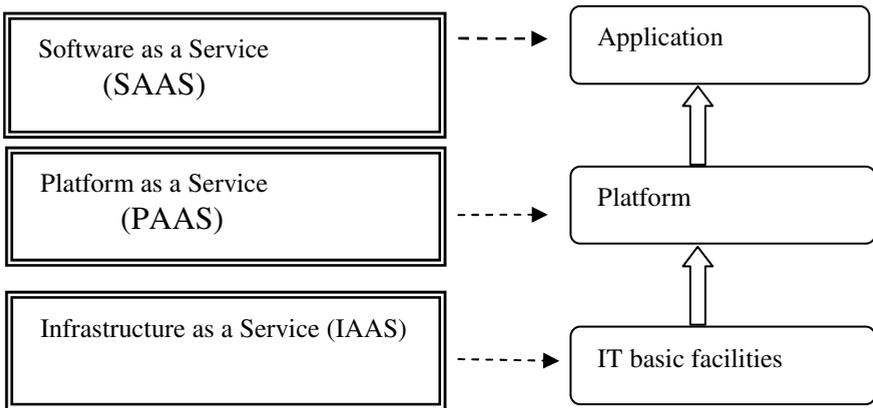


Fig. 1. Three basic forms of cloud computing service

Seen from the development history of information technology, it is a process from information scarcity to material abundance. When the information was scarce, people focused on technology and hoped they could produce more information. The researchers in our country unanimously believe that the basic element of cloud computing is information [5]. Information management science, computer science and network techniques are complementary to each other. Applying the relevant principles and theories of information resources management to cloud computing will bring opportunities for information management subject, and meanwhile the development demands of information resource management are also able to improve the performance requirements of cloud computing correspondingly. At present, all countries in the world pay attention to the great influence of cloud computing on information management, the new chief information officer (CIO) of American federal government Vivek Gundra said the government would be pushed to adopt cloud computing, and he believed although cloud computing may have potential privacy or safety dangers, the speed and efficiency of cloud computing can not be neglected for this reason. Currently the development and application of cloud computing in China information resource filed is still in the primary stage. The application and principles of cloud computing are discussed in the paper first based on information resource management.

2 Analysis of Cloud Performance and Characteristics

First, it can rent server hardware and software without the need to invest any costs for file storage hardware in early stage. Reduce the costs of every month to manageable scale, and in addition many companies (e.g. Google) provide its Web application program free of charge. Second, it can choose certain location for data storage, while cloud computing is able to lock the data to some designated physical locations, and furthermore cloud computing provides unlimited storage ability. Third, the data can be obtained without the limit on location. Through cloud computing, there is no need for users to carry the files with them. All data are in the cloud, and can be obtained only

with a computer and network connection. In addition, the control function of cloud computing storage can enable you to share the files only with the person you designate, whether they are your staff or outsiders. You can control visitors' privilege, while you do not need to allow the outsiders to visit your company network.

3 Use Cloud Computing to Improve the Sharing and Building Ability of Information Resources

China is a country with abundant information resources, while the information resources produced, used and occupied by public sectors, e.g. governments, approximately account for 80% [6] of total amount of information resources in the whole society, containing great social and economic values. However, in the process of information construction, there are many problems in information resources application and utilization, mainly reflected in: under-exploitation, insufficient utilization and low efficiency of information resources, relatively lagging behind information infrastructure construction; the sharing of information resources is difficult and the collection is repetitive. Furthermore, it is difficult for traditional information technology and network environment to meet of intelligent terminal user demands for information. Therefore, utilizing cloud computing mode to solve the resource storage and sharing, unifying the data standard in the industry, reducing costs and improving efficiency, promoting the value-added development and utilization of information resource is the important measure [7] for improving the sharing management ability of information resources in China. Figure 2 is "Information Resource Model of Cloud Computing".

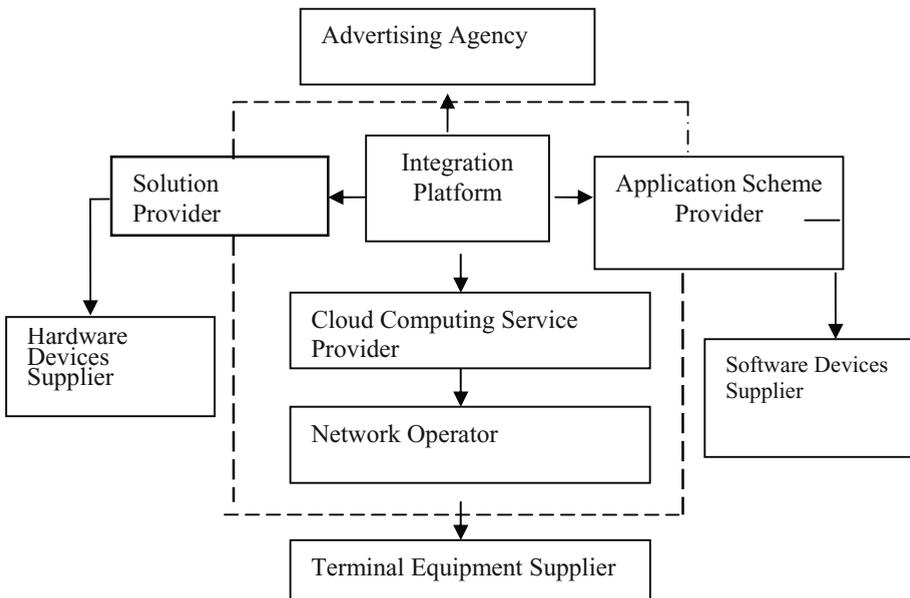


Fig. 2. Information Resource Model of Cloud Computing

Cloud computing users: usually they include the individual consumers using cloud computing services to meet the individual demands, and also include the enterprise consumers using cloud computing to improve enterprise informatization, etc.

Cloud computing service providers: they are the sponsors and service providers of cloud computing. After excavating and identifying the demands of users or user groups, the cloud computing service providers provide particular cloud computing products and services for user groups, and further integrate their own resources and the resources and abilities of various parties interested in cloud computing, build the integration platform of cloud computing service, form cloud computing value network geared to the customers value, enabling other enterprises participating in the market to excavate the real customer needs on the service integration platform of cloud computing, and finally realize the mutual benefits of users, partners and themselves. For example, Google is the typical representative of cloud computing service providers, whose aim is integrating all information around the world for analysis and procession and finally providing to the users by a simple “plug”.

Participants: Solution providers, application and content providers, software, hardware and equipment suppliers and advertising agencies, etc. Solution providers: they provide more targeted, more personalized and more specialized solution to user groups on the basis of value-added development integration platform of cloud computing information resources, and the network resources, software and hardware resources, products and services combination provided by other participants and organizers. For example, IBM life cycle management of cloud service.

Application and content provider: Products and services are the materialized form and carrier of values. While application and content is the principal form of products and services. Application and content providers excavate various kinds of network resources for the demands of particular industry businesses and user groups.

Software, hardware and equipment providers: they are the indispensable participants in information technology application field. For example, Dell provides end-end total solution for users when assisting enterprises in implementing cloud computing, including hardware equipments, software and services.

Advertising agencies: Currently advertising agencies create the most direct incomes for many internet enterprises, and they are the most direct value sources in internet business mode. In the economic benefits of cloud computing, the advertising based revenue mode is the major profit model of cloud computing operators.

4 Cloud Computing Influences and Reformation on Information Resources Management

It can be seen from the above-mentioned feature analysis of cloud computing, cloud computing can realize the mutual benefits of end users and cloud service provider. Therefore, studying and spreading cloud computing in information resource management is inevitable trend of development. The advent of cloud computing age will bring the following significant change for information resource management activities:

Motivating the Advent of New-generation Pervasive Computer age. With the rapid development of mobile devices, smart phones and sensor technologies, ubiquitous

computing is entering people's daily life gradually. One of important features of ubiquitous computer is that the computers and information system can perceive the change of physical world, and then make initiative adjustment of their behaviors in order to provide better service for human beings [9]. Although academic circles have researched ubiquitous computing technology for many years, the forthcoming cloud computing age brings new challenges and opportunities for the fields. For example, based on the accumulation and excavation of physical world circumstances knowledge from a large number of customers and equipments, we can develop more innovative internet service. The comprehensive connection and blending of human society, information space, physical world are becoming the important feature and the major trend of the next generation ubiquitous computing technology. The essence of human, machines and objects mix environment is providing more transparent, more intelligent, more ubiquitous, greener and safer integrated service, and the environment is highly harmonious.

Changing the Storage Technology and Management Mode of Information Resources In cloud computing environment, the information resources of individual customers and enterprise customers are kept in clouds with unified storage and management by cloud server cluster. The quantity of cloud information resource, handling capacity of visiting I.nJ, computation speed, fault-tolerant ability, safety and concurrent processing requirements exceed the information storage and management ability scope of server end in server computing mode. So further development of cloud computing needs storage technologies and management methods of information resource management.

5 Basic Application Principles of Cloud Computing in Information Resource Management

In the environment e.g. cloud environment where the information highly concentrates, the basic principles regarding information organization are still applicable, and also can become the quality standard for judging the information organization of cloud computing.

Modernization Principles. The principles of information organization modernization include the two aspects, i.e. the concept modernization and technological means modernization. The concept modernization of information organization is mainly reflected in the standardization of information organization. The technological means modernization of information organization has fully proved that the wide popularization and application of modern information technology form the irreversible automation direction of information organization. Automatic indexing, secondary automatic information generation and database construction have become relatively mature automatic results of information organization at present. The technological means modernization of information organization changes the traditional manual mode, greatly improves work efficiency and work quality and better meets the diversified information demands of users, and conforms to the general trend of

information management. In cloud environment where information and services are extremely abundant and complex, the technological means modernization of information organization seems very important.

Pertinence Principles. Because the customers' information demands differ in thousands ways, and vary dynamically with time. For this reason, providing users with satisfactory information service is not easy. The enterprises providing cloud services should seriously study users' information demands and the demands variation, master users habits of using information and the characteristics, and provide users with well-directed information services.

Benefit Principles. Whether for cloud computing enterprises or users, certain costs are needed (time costs and fund costs), we should ensure that the maximum benefits of information services are obtained with the least cost.

Systematization Principles. In cloud computing, storage resources and databases expand continuously. If information organization work can not be grasped effectively as a whole, the system organization information and cloud storage efficiency will have a greatly reduced quality. So the relation between macroscopic information organization and microcosmic information organization, the relation between information organization department and other departments, the relation between various sections of information organization work and the relation between different information processing methods must be grasped.

6 Summary

The emergence of cloud computing brings new hot topic for computer world. It receives extensive praise from various trades in society because it can use the minimized client side to realize complex and effective procession and storage, bringing us tremendous development space. If cloud computing network is popularized in information resource sharing management, there is no doubt the ability of information resource sharing will be improved greatly, and furthermore it can bring new application fields for cloud computation, making cloud computing service more close to our life and study, and making us experience tremendous revolution the technology brings us. Although at present the research on cloud computing is not mature, we have good reason to imagine that the technologies in information resource management field can become more scientific and more modernized by utilizing the efficient calculating procession ability and unlimited storage ability of cloud computing.

References

1. Twenty experts define cloud computing, http://cloudcomputing.sys-con.com/read/612375_p.htm
2. Buyya, R. (ed.): High Performance Cluster Computing: Architectures and Systems, vol. 1. Prentice-Hall, Upper Saddle River (1999)
3. Stuer, G., Vanmechelena, K., Broeckhovea, J.: A commodity market algorithm for pricing substitutable grid resources. *Future Generation Computer Systems* 23(5), 688–701 (2007)

4. Zeng, R., Lin, Y., Lv, Y.: Design and Realization of Management System of General Question Bank. Journal of Fujian Agriculture and Forestry University (Natural Science Edition) 37, 78–80 (2008)
5. <http://cloud.it168.com/a2009/1013/757/000000757668.shtml>
6. <http://www.amazon.com/ec2/> (July 18, 2008)
7. Zeng, R., Lin, Y., Lv, Y.: Design and Realization of Management System of General Question Bank. Journal of Fujian Agriculture and Forestry University (Natural Science Edition) 37, 78–80 (2008)

The Comparison between Two Control Methods about Simulation of Traction Load

Jingyu Hu, Min Xu, Kai Guo, and Yongqiang Cui

School of Electrical Engineering, Southwest Jiaotong University, Chengdu, China

Abstract. This paper proposed an enhanced solution, which adopted ways of changing methods for controlling current on inverter side, adding practical current after enabling D-Q transform and adopting triangular carrier control instead of hysteresis control, for simulating traction loads based on PWM converter. Simulation result figured that increasingly close loops on inverter side and reduction on the risk of instability may be caused by exceeded switch frequency, comparing with existing solution using hysteresis and direct current control, can be achieved in the advantages of solution proposed in this paper. Therefore, energy saving and better traction load simulating have been implemented.

Keywords: Cophase power supply, Simulating load characteristics, PWM, Hysteresis control, D-q transform.

1 Introduction

The traction load is single-phase AC load, resulting in three-phase unbalanced in power system. Cophase power supply system is the best solution to solve this problem. By using the cophase power supply technology in Substation, the primary side of traction transformer can not rotate, cophase power supply substation to cancel the phase insulator. The testing system of Cophase power supply is the necessary means to study the technology of cophase power supply. Simulating traction load is an important component in the system; it can simulate the AC - DC and AC - DC - AC traction load.

According to the main circuit AC / DC / AC conversion structure of the simulation traction load, analysis of the rectifier side and inverter side under different control strategies in the simulation results. Rectifier side is mainly the comparison of hysteresis control and triangular carrier control on PWM. Inverter side is mainly the comparison of direct current loop control and improved d-q transformation of the double-loop control.

2 The Simulation of Traction Load and the Working Principle of the Main Circuit

The main circuit consists of 10 anti-parallel IGBT with diodes composed the rectifier and the inverter parts. On the one hand, it is to be able to simulate various operating

conditions of the traction load, On the other hand, it also can feedback energy for testing by a high power factor into grid. Input inductor is added at both ends of the power to balance voltage, Rectifier and the inverter parts are separated by regulator, used to adjust the energy exchange between the rectifier side and inverter side and stable DC voltage and inhibit the harmonic voltage on DC side [1-2].

Set the single-phase AC power supply voltage of the test the standard sine wave, the load simulates the impedance load.

$$\begin{aligned} u_{s1} &= U_{m1} \sin \omega t \\ i_{s1} &= I_{m1} \sin(\omega t - \phi_1) \end{aligned} \tag{1}$$

The part of the energy's feedback is an inverter device and whose active power factor is 1, a phase on inverter side as that

$$\begin{aligned} u_{s2} &= U_{m2} \sin \omega t \\ i_{s2} &= I_{m2} \sin(\omega t - \phi_2) \end{aligned} \tag{2}$$

To achieve the PWM rectifier to work, the capacitor voltage must be met

$$u_{dc} \geq U_{m1}, u_{dc} \geq \sqrt{3}U_{m2} \tag{3}$$

Steady-state operation, the active power must be balanced

$$P_1 = P_2 \tag{4}$$

$$I_a = \frac{U_{m1} I_{m1} \cos \phi_1}{3U_{m2}} \tag{5}$$

Because this article focuses on comparing the two control methods, we are no longer to describe some of the circuit parameters.

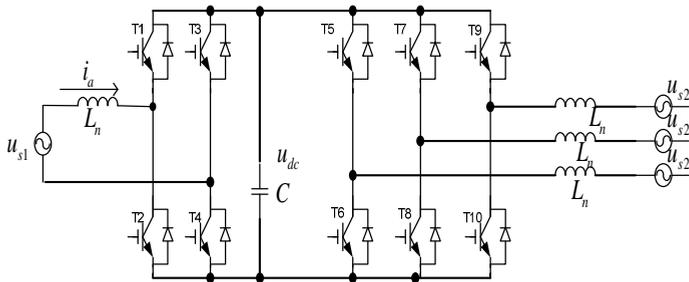


Fig. 1. Main circuit based on single-phase--three-phase PWM rectifier simulation of the traction load



u_{s1} : the voltage of substation reduced by a certain percentage ; u_{s2} : the voltage of the power grid in testing; u_{dc} : the voltage of the regulator; T1~T10: anti-parallel IGBT with diodes.

3 Comparisons of Two Control Strategies

\tilde{i} : the specified current ; i^* : the measured current;
 \tilde{u}_c : the specified voltage of the regulator; u_c^* : the measured voltage of the regulator.

3.1 Rectifier Side

1) The block diagram of the control system by hysteretic on rectifier side[3].

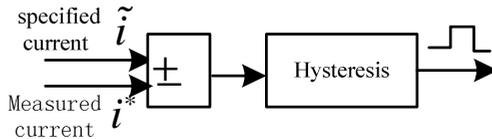


Fig. 2. Block diagram of current control on rectifier side

2) The block diagram of the control system by triangular carrier on rectifier side.

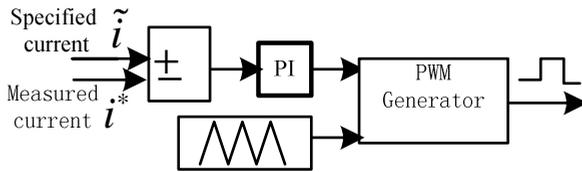


Fig. 3. Block diagram of current control rectifier

3.2 Inverter Side

D-q transformation: Symmetric positive-sequence three-phase sinusoidal current i_a, i_b, i_c transformed through d-q transformation and the d, q axis components namely [4]:



$$\begin{bmatrix} i_{d+} \\ i_{q+} \end{bmatrix} = T_+ \begin{bmatrix} \sqrt{2}I_+ \cos(\omega t + \varphi_+) \\ \sqrt{2}I_+ \cos\left(\omega t - \frac{2\pi}{3} + \varphi_+\right) \\ \sqrt{2}I_+ \cos\left(\omega t + \frac{2\pi}{3} + \varphi_+\right) \end{bmatrix} \tag{6}$$

1) Direct current control in loop control

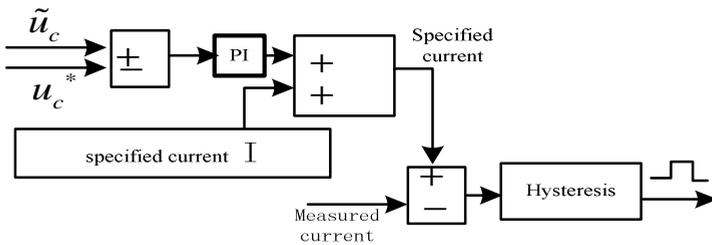


Fig. 4. Block diagram of loop control on the inverter side in method I

The specified voltage of the capacitor being subtracted by the actual voltage and the result through the PI links, and then add the specified current, and phase determined by the phase of the voltage in feedback grid. Designated as the final specified current [3].

2) Controlled under d-q transformation in loop control

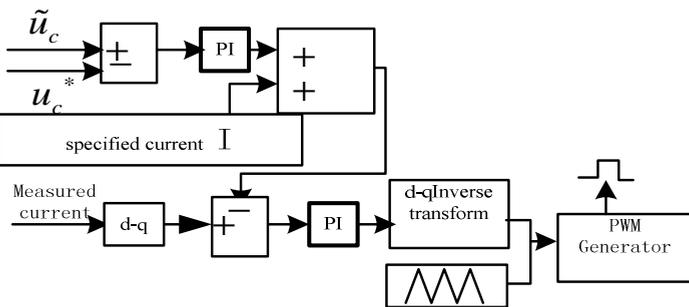


Fig. 5. Block diagram of double loop control on the inverter side in method II



The specified voltage of the capacitor being subtracted by the actual voltage and the result through the PI links, and then add the specified current, and then subtracted by the active component of the three-phase current transformed by d-q transformation on inverter side, the result converted from voltage to current after PI links. After d-q inverse transformation, then compared with the triangular carrier. It can control the IGBT on inverter side.

4 The Results of the Simulation

4.1 Rectifier Side

1) The THD of the waveform from the substation's original data:

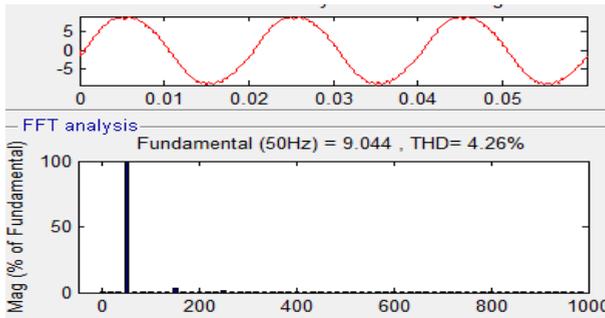


Fig. 6. The THD of the specified current on rectifier side

2) The THD of the tracking current Observed in the FFT in method I.

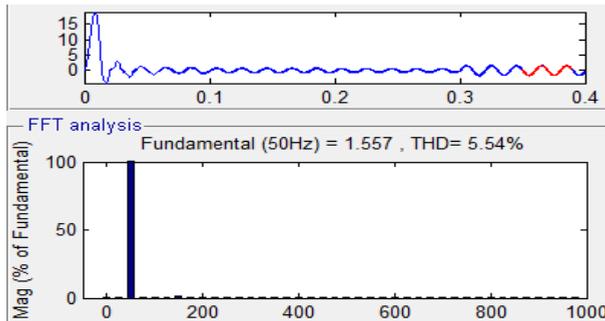


Fig. 7. The THD of the tracking current in method I

3) The THD of the tracking current Observed in the FFT in method II.

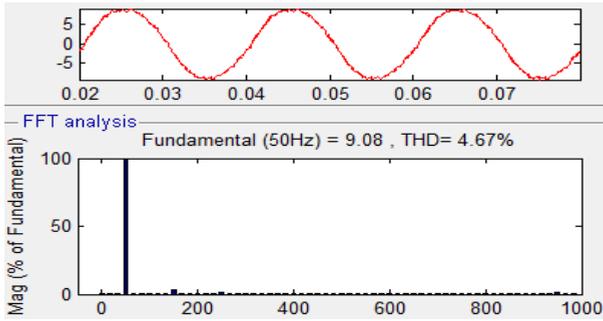


Fig. 8. The THD of the tracking current in method II

4) The specified current and the tracking current.

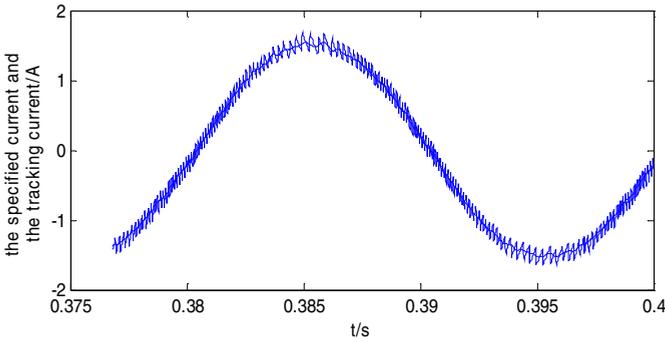


Fig. 9. The specified current and the tracking current

4.2 Inverter Side

1) The THD of the current on inverter side Observed in the FFT in method I.

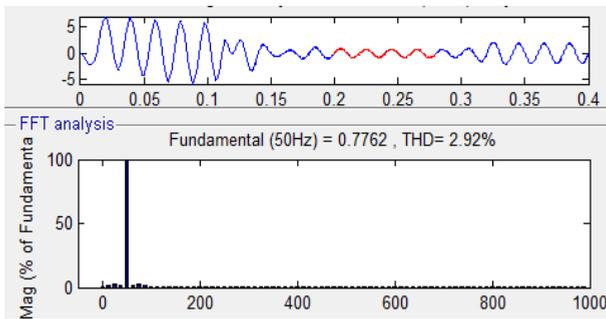


Fig. 10. The THD of the current on inverter side in method I

2) The THD of the current on inverter side Observed in the FFT in method II

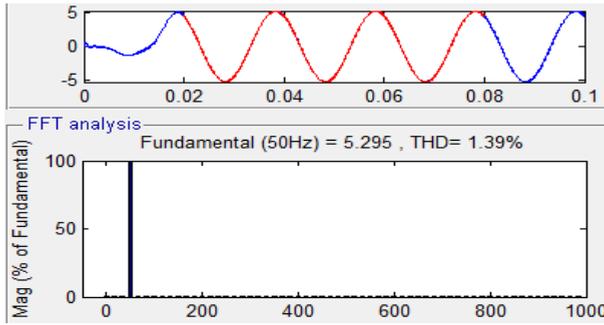


Fig. 11. The THD of the current on inverter side in method II

3) The response of the current on inverter side in method I

the response of the current on inverter side /A

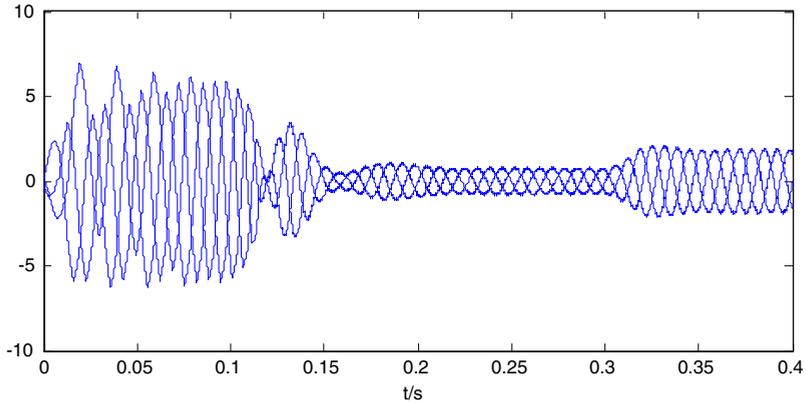


Fig. 12. The response of the current on inverter side in method I

4) The response of the current on inverter side in method II.

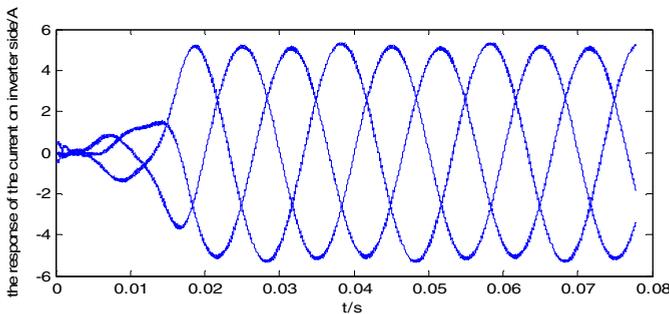


Fig. 13. The response of the current on inverter side in method II

5) The phase of the current and voltage on inverter side.

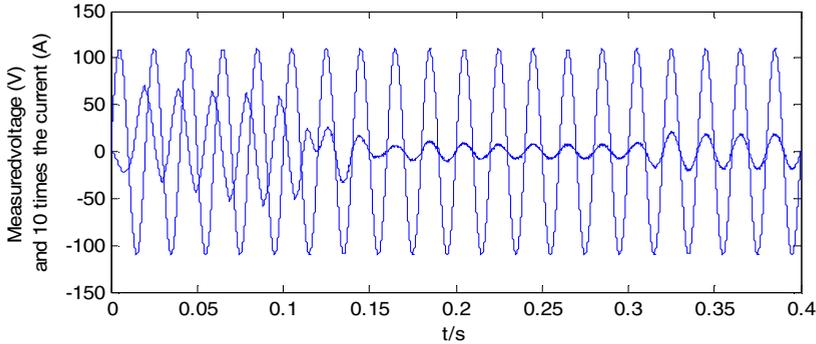


Fig. 14. The phase of the current and voltage on inverter side

5 Conclusion

As the simulation experiments with two control strategies.

A. Rectifier Side

Compared with hysteresis control, under triangular carrier comparison control the THD of the tracking current is less, and the power switching device's frequency is certain, which is equal to the carrier frequency (in the experiment is 3kHz), and it's less than the switching frequency under the hysteresis control. This advantage makes the device a better implementation and more energy efficient.

B. Inverter Side

The inverter side under the double loop control combination of d-q transformation and d-q inverse transformation. The current on inverter side has a higher response speed, and the THD of the current is less, too. It makes the rectifier and the inverter parts a better contact, and reduces the harmonics' passing in the direct current control.

References

1. Li, F.: Study of AC Electronic Load based on PWM converter. *High Voltage Engineering* 34(5), 930–934 (2008)
2. Zeng, G.-H., Hao, R.-T.: A novel three-phase balanced traction supply system based on active power filter. *Journal of the China Railway Society*. 5(1), 48–53 (2003)
3. Feng, J.: Study on simulated scheme of traction load in Electrified Railway Cophase Power Supply Test System. *Journal of the China Railway Society* 32(6), 29–33 (2010)
4. Du, X.: Short-circuit fault classification causing dual voltage sags based on double d-q transformation. *Automation of Electric Power Systems* 34(5) (2010)

A Notable Problem on t-Distribution

Shen Maoxing, Li Jun, and Zhang Yaoping

Department of Management Science, Xijing University, Xi'an Shaanxi,
P.R. China, 710123
shenyu_ze@21cn.com

Abstract. A notable problem on student's distribution is presented due to it is a widely used in applied statistics. It comes from a well known relation between F-distribution and t-distribution with them definitions. It pointed out that all related problems will solving if this problem could solve. This is have a great meaning for the theoretic and applied statistics.

Keywords: Student's distribution, Probability distribution function, Probability distribution density function, Applied statistics.

1 Introduction

In the statistics, the student's distribution $t(n)$ takes an important role in the theoristis researches and wide fields of applications such as quality control and management, reliability and maintenance eEngineering, total quality management etc.. There is a well known fact about the relation between student's distribution and F-distribution that $X^2 \sim F(1, n)$ if $X \sim t(n)$. In the mean while, we know the

definition of t-distribution $\frac{X_1}{\sqrt{Y_1/n}} \sim t(n)$ if $X_1 \sim N(0,1)$ and $Y_1 \sim \chi^2(n)$

are independent each other, then we can proved a conclusion easily that

$$\frac{X_1^2}{Y_1/n} \sim F(1, n).$$

Hence, we should meet a class problem in calculations of random variables spontaneously that is as below:

Suppose $X \sim N(0,1)$, X and Y are independent each other, wheather $Y \sim \chi^2(n)$ is hold when $\frac{X}{\sqrt{Y/n}} \sim t(n)$? If no, what is the distribution of

Y and what is the distribution density fuction of it? This notable problem can be consider as a think of converse distribution of the t-distribution.

2 Derivation of the Distribution Function

Based the consideration as above mentioned, let $Z = \frac{X^2}{Y/n}$ i.e. $Y = \frac{nX^2}{Z}$, suppose

$X \sim N(0,1)$, $Z \sim F(1,n)$, X and Z are independent each other. We consider the distribution of Y .

Hence, the union probability density distribution function (p.d.d.f.) of (X, Z) is as follows

$$f(x, z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})} \frac{1}{n} z^{-\frac{1}{2}} (1 + \frac{z}{n})^{-\frac{n+1}{2}}, \quad (z \geq 0) \tag{1}$$

Sequentially, the probability distribution function (p.d.f.) of Y is as follows

$$\begin{aligned} F_Y(y) &= P\{Y \leq y\} = \iint_{\frac{nx^2}{z} < y} f(x, z) dx dz \\ &= \frac{1}{\sqrt{2\pi}} \frac{\Gamma(\frac{n+1}{2})}{n\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})} \iint_{\frac{nx^2}{z} < y} e^{-\frac{x^2}{2}} z^{-\frac{1}{2}} (1 + \frac{z}{n})^{-\frac{n+1}{2}} dx dz \\ &= \frac{1}{\sqrt{2\pi}} \frac{\Gamma(\frac{n+1}{2})}{n\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \int_{\frac{n}{y}}^{+\infty} \frac{1}{\sqrt{z}} (1 + \frac{z}{n})^{-\frac{n+1}{2}} dz \quad (y > 0) \\ &= \frac{1}{\sqrt{2\pi}} \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n}\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \int_{1 + \frac{1}{y}}^{+\infty} \frac{u^{-\frac{n+1}{2}}}{\sqrt{u-1}} du \quad (u = 1 + \frac{z}{n}) \\ &= \frac{1}{\sqrt{2\pi}} \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n}\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \int_{1 + \frac{1}{y}}^{+\infty} \frac{u^{-\frac{n+1}{2}}}{\sqrt{u-1}} du \end{aligned}$$



$$\begin{aligned}
 &= \frac{2}{\sqrt{2\pi}} \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n}\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \int_{\arctg \frac{|x|}{\sqrt{y}}}^{\frac{\pi}{2}} \cos^{n-1} t dt \quad (u = \sec^2 t) \\
 &= \frac{2}{\sqrt{2\pi}} \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n}\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})} \left[\frac{\pi}{\sqrt{2}} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n}{2} + \frac{1}{2})} - \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \int_0^{\arctg \frac{|x|}{\sqrt{y}}} \cos^{n-1} t dt \right] \quad (y > 0) \\
 &= \frac{1}{\sqrt{n}} - \frac{1}{\pi} \sqrt{\frac{2}{n}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \int_0^{\arctg \frac{|x|}{\sqrt{y}}} \cos^{n-1} t dt \quad (2)
 \end{aligned}$$

So, the p.d.f. of Y is

$$\begin{aligned}
 f_Y(y) &= F_Y'(y) \\
 &= -\frac{1}{\pi} \sqrt{\frac{2}{n}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \left[\int_0^{\arctg \frac{|x|}{\sqrt{y}}} \cos^{n-1} t dt \right]'_y \\
 &= \frac{1}{\pi} \sqrt{\frac{2}{n}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{1}{2y\sqrt{y}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} \frac{|x|}{(1 + \frac{x^2}{y})^{\frac{n+1}{2}}} dx \\
 &= \frac{1}{\pi} \sqrt{\frac{2}{n}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{1}{y\sqrt{y}} \int_0^{+\infty} e^{-\frac{x^2}{2}} \frac{x}{(1 + \frac{x^2}{y})^{\frac{n+1}{2}}} dx
 \end{aligned}$$

but

$$\begin{aligned}
 \int_0^{+\infty} e^{-\frac{x^2}{2}} \frac{x}{(1 + \frac{x^2}{y})^{\frac{n+1}{2}}} dx &= \int_0^{+\infty} e^{-w} \frac{dw}{(1 + \frac{2}{y}w)^{\frac{n+1}{2}}} \quad (w = \frac{x^2}{2}) \\
 &= \frac{y}{2} e^{\frac{y}{2}} \int_1^{+\infty} u^{\frac{n+1}{2}} e^{-\frac{y}{2}u} du \quad (u = 1 + \frac{2}{y}w)
 \end{aligned}$$



$$\int_1^{+\infty} u^{-\frac{n+1}{2}} e^{-\frac{y}{2}u} du = \left(\frac{2}{y}\right)^{\frac{n-1}{2}} \int_{\frac{y}{2}}^{+\infty} t^{-\frac{n+1}{2}} e^{-t} dt \quad (t = \frac{y}{2}u)$$

$$= \left(\frac{2}{y}\right)^{\frac{n-1}{2}} \int_{\frac{y}{2}}^{+\infty} t^{-\frac{n+1}{2}-1} e^{-t} dt$$

Hence, the p.d.d.f. of Y is

$$f_Y(y) = F_Y'(y)$$

$$= -\frac{1}{\pi} \sqrt{\frac{2}{n}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{y}{2} e^{\frac{y}{2}} \left(\frac{2}{y}\right)^{\frac{n-1}{2}} \int_{\frac{y}{2}}^{+\infty} t^{-\frac{n+1}{2}-1} e^{-t} dt$$

$$= -\frac{1}{\pi} \sqrt{\frac{2}{n}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \left(\frac{y}{2}\right)^{\frac{n+1}{2}} e^{\frac{y}{2}} \int_{\frac{y}{2}}^{+\infty} t^{-\frac{n+1}{2}-1} e^{-t} dt$$

$$= -\frac{1}{\pi} \frac{\Gamma(\frac{n+1}{2})}{(\sqrt{2})^n \sqrt{n} \Gamma(\frac{n}{2})} e^{\frac{y}{2}} y^{\frac{n+1}{2}} \int_{\frac{y}{2}}^{+\infty} t^{-\frac{n+1}{2}-1} e^{-t} dt \tag{3}$$

3 Conclusion and Notes

There is a close related problem or an equivalent problem with our mentioned question that could be described as follows, suppose $T \sim t(n)$ and $X \sim N(0,1)$

are independent each other, $Y = \frac{nX^2}{T}$ (i.e. $T = \frac{X}{\sqrt{Y/n}}$), then, whether is

$Y \sim \chi^2(n)$ hold? If no, what is the distribution function or p.d.d.f. of Y ? This proposition can be stated as follows too: when $S \sim \chi^2(1)$, $T \sim t(n)$, S and

T are independent each other, then $Y = \frac{nS}{T} \sim \chi^2(n)$? These propositions could be solved all when one of these is quite solved.



References

1. Editor Group of 《Probability and Statistics》 of the mathematics and dynaminc department of Zhongshan University, Probability and Statistics. People Education Press, Beijing (1980) (in Chinese)
2. Sheng, J., Xie, S., Pan, C.: Probability and Statistics, 2nd edn. High Education Press, Beijing (1989) (in Chinese)
3. Olkin, I., Gleser, L.J., Derman, C.: Probability Models and Applications. Macmillan Publishing Co., Inc., New York (1980)
4. Shen, M.: Relationships among the Common Statistical Distributions. Journal of Northwest University (Natural Science Edition) 29(S.I.), 35–39 (1999) (in Chinese)

SVM-Based Pornographic Images Detection

Haiming Yin, Xiangqiong Huang, and Yuanwang Wei

Computer Application Research Lab, Jiaxing University, Jiaxing 314001, China
hai_tao_123@sina.com

Abstract. In this paper, we present a new automatic system for pornographic image detecting. An improved skin model in RGB color space is employed to identify the rough skin regions. An original texture filter is used to modify these regions. Then the useful information of skin regions is extracted. Close-up face images are excluded by the face detection. At last all the information of skin and face is all fed to the SVM Classifier to tell whether the image is pornographic or not. Our experiments on real-world web image data indicate that our system can improve the accuracy of pornographic content detection significantly.

Keywords: pornographic image detection, skin identifying, texture filtering, fractal dimension.

1 Introduction

With the increase of network bandwidth and hardware storage capacity, thousands of millions of images have appeared on the internet, which enable people conveniently achieve useful information and entertain their lives. However images can contribute in many useful fields, it also can bring much negative influence, especially the pornographic images harm children's physical and mental health. So developing efficient applications of image filtering for the illegal pornographic content has become one important task to researchers.

The pioneer work for identifying pornographic images by the analyzing the image content are proposed by [1-9]. In [1-3] color and texture properties are used to obtain a mask of skin regions. A specialized grouper attempts to group human figures using geometric human body structure constraints. But the false negatives ratio of these algorithms is always high for several reasons. Some close-up or poorly cropped images do not contain arms and legs, vital to the current geometrical analysis algorithm. Regions may have been poorly extracted by the skin filter, due to desaturation. The edge finder may fail due to poor contrast between limbs and their surroundings. In [4], Xu et al. present an approach which extracts color and texture features from arbitrary-shaped segmented regions. In the classification phase, Gaussian Mixture Models (GMM) are built for skin and non-skin region classification. In [5,6], the online sampling algorithm is employed to construct the positive and negative set, and then use it in the image filtering. In [7], Shen et al. present an erotogenic-part detection based method for the image filtering. In [8], Hakan et al. present an adult image content classification model which uses global features and skin region. In [9], a statistical

color models for skin detection are developed. They get the probability distribution of the color of human skin from extensive samples and then use some math model such Bayesian model to detect the skin.

In this paper, we present a SVM-based automatic system for pornographic image detection. In the first phase, the image features include skin, face, shape are extracted. In the second phase, all features are fed to a SVM-based classifier to tell whether the image is pornographic or not.

2 Skin Detection

Detecting Skin Using Color Information. Generally, big skin regions are the most evident feature for a pornographic image. Skin detection is the first step in pornographic image detection. There are many skin models which cover several color spaces presented to deal with this problem [1, 3, 4, 6].

Peer [11] introduced one skin model for face tracking which describes the skin cluster in the RGB color space which has very fast speed and can detect most skin regions in images. Inspired by Peer's algorithm, we present a subsection model in RGB color space for skin detection which model skin according to the luminance of the pixel.

Definition:

$$\text{Gray} = 0.5R + 0.35G + 0.15B \quad (1)$$

The rules for skin detection according to the luminance are as follows:

IF Gray < 70 , R-G > 5 AND R-B > 5 AND max(G,B) > 20

ELSE IF Gray < 140, R > G AND R > B AND R-min(G,B) > 15 AND R < G+B

ELSE IF Gray < 210, R > G AND G > B AND R-G > 25 AND R < 1.2G+B

ELSE R > B AND G > B |R-G| ≤ 15

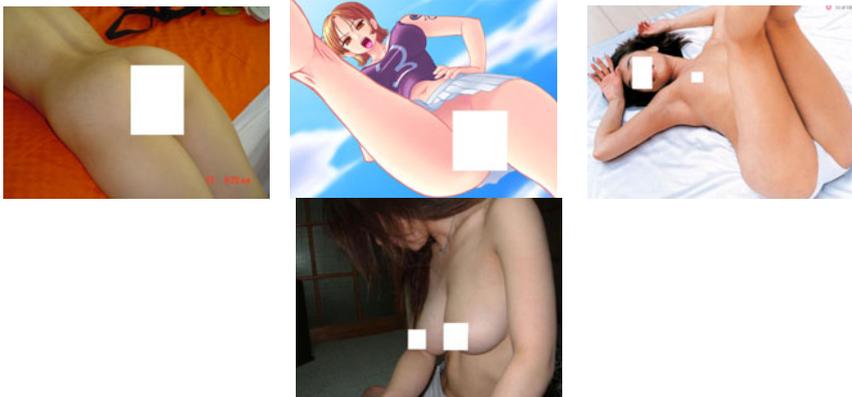


Fig. 1. Skin detection result

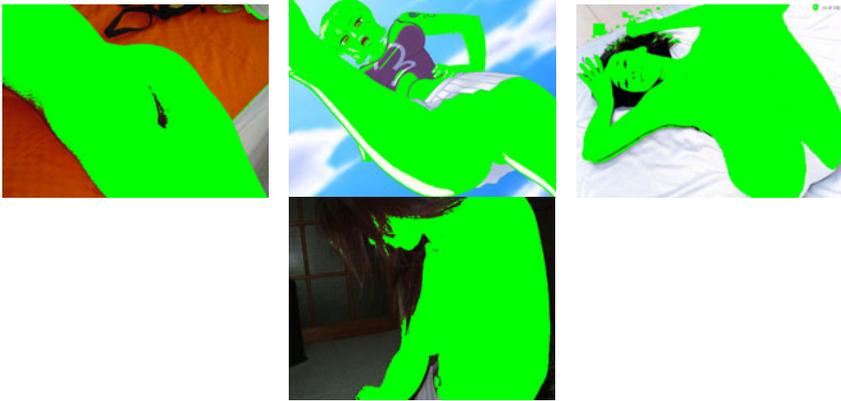


Fig. 1. (continued)

Modify Skin Region with Texture. Because there are many things (such as lion's fur or a pile of soybean) which color is close to the skin, this makes it is very difficult to detect the skin precisely with color information only. Since human skin has an obvious property—smooth, skin detection algorithms always use texture to modify skin regions detected using color information.

In this paper, we propose a coarse-degree-based skin texture filtering algorithm, computing every pixel's color coarse degree and then get the whole region's texture coarse degree around that pixel. One threshold is used to tell whether one pixel is skin pixel.

The equation for calculating the color-coarse-degree of pixel \mathbf{P}_x is computed as.

$$\bar{v} = \frac{1}{N} \sum_{i=1}^N v_i, \quad R_x = \left[\frac{1}{N} \sum_{i=1}^N \|v_i - \bar{v}\|^2 \right]^{\frac{1}{2}} \quad (2)$$

Here N is the number of skin pixels in the window within a radius of r , v_i is the color vector of the i th pixel, $\| \cdot \|$ means the Euclidean distance.

After getting every pixel's texture coarse degree in a skin region, one statistical work is accounted and the 20% pixels with biggest coarse degree will be discarded, and this skin region's texture coarse degree will be computed by:

$$R_{\text{patch}} = \frac{1}{M} \sum_{x=1}^M R_x \quad (3)$$

Because shot distance will affect skin texture, usually the skin texture in a close-up image is coarser than that in a long-distance image. Through experiments we found that most of close-up pornographic images have more than 60% skin pixels and the long-distance images always have skin pixels less than 20%. Therefore the radius of texture filtering should be dynamically adjusted according to the ratio of the skin color pixels of the image.

$$r = \alpha^2 \frac{\min(H, W)}{Thre} \tag{4}$$

Here α is the ratio of skin pixel in the image, and H, W are the image height and width, $Thre$ is a const and it should be changed with the change of image size, and in our experiment all images are scaled to 320*240 and it is set to 50.

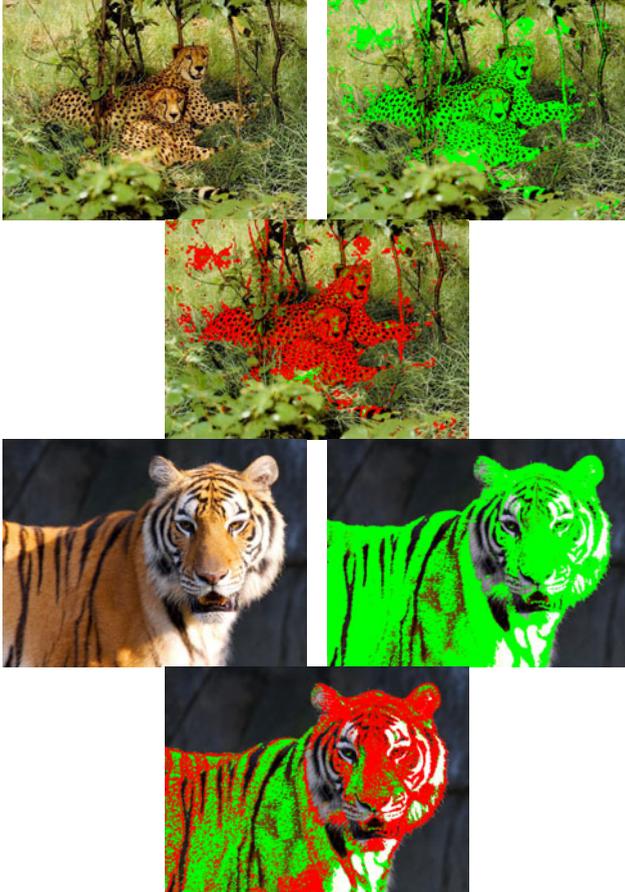


Fig. 2. Texture modifying result (Left: original images, Middle: skin detection, Right: texture modifying)

3 Geometry Feature Analysis of Skin Regions

Geometry is important information when detecting pornographic content in an image. Fractal geometry is presented by Mandelbrot [11], mainly studies irregular and disordered figures which can't be described by Euclidean geometry. As a measure of the object contour's irregular degree, fractal dimension is consistent with human's

perception. Geometries with similar regularity have similar fractal dimension and the fractal dimension between differently regular geometries always have large difference. In this paper we use fractal dimension of the skin regions to filter some scenery pictures which are difficult to distinguish using normal feature (such deserts or beaches which all have the similar color and texture with real human skin).

There are many methods to compute the fractal dimension of a curve [12], in this paper we choose the simple perimeter-area fractal dimension considering of the computation task. (See details in [12])

Let L be the boarder of an object in an image and Ω the region, and define functions $f(x, y)$ and $g(x, y)$ as follows:

$$f(x, y) = \begin{cases} 1, & \text{pixel } (x, y) \text{ is on the edge} \\ 0, & \text{otherwise} \end{cases}$$

$$g(x, y) = \begin{cases} 1, & \text{pixel } (x, y) \text{ is in the region} \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

Then the fractal dimension of the object can be computed by equation (11).

$$D = 2 \frac{\ln\left(\sum_{(x,y) \in L} f(x, y)\right)}{\ln\left(\sum_{(x,y) \in \Omega} g(x, y)\right) + (\ln \alpha)^2} \tag{6}$$

The skin regions after texture modifying will be processed in morphology through erosion and dilatation, and then the fractal dimensions of the regions will be computed.

We perform the experiments on 244 adult images and 64 desert images and 47 beach images. Fig3 describes the distribution of the fractal dimension of real skin regions and false ones in the desert and beach images. It is obvious that the fractal dimensions of these two kind regions are statistically separable if one suitable threshold is chosen.

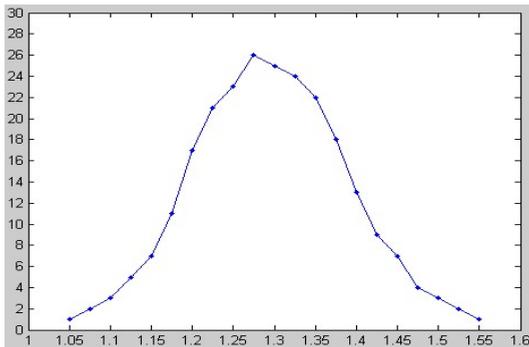


Fig. 3. Fractal dimension distribution (horizontal ordinate is the fractal dimension and vertical coordinates is the image num). Left: fractal dimensions of skin regions. Right: fractal dimensions of desert and beach regions.

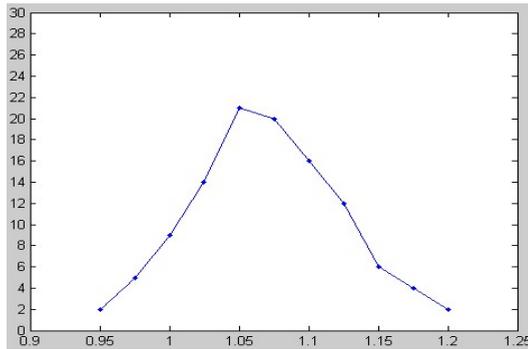


Fig. 3. (Continued)

4 Face Detection

Since the close-up face photos always contain big skin region, which make this kind of image is easily detected as pornographic image. Face detection is a feasible way to solve this problem. Meanwhile the face information can help to modify the skin regions detected by skin detection.

There are many algorithms for face detection [13, 14]. In these algorithms, AdaBoost-based algorithm is the most popular one for its stability and precision [14]. AdaBoost, from adaptive boosting, was rapidly made popular in the machine learning community after it was presented by Freund [26]. AdaBoost is a learning method for solving classification problems which combines an ensemble of weak classifiers into a strong classifier. In this paper, we use this kind of method to detect faces in the image.



Fig. 4. Face detection result

5 Use SVM to Classify the Images

The main idea of Support Vector Machine is to construct a nonlinear kernel function to map the data from the input space into a possibly high-dimensional feature space and then generalize the optimal hyper-plane with maximum margin between the two classes [15]. Given an N-element training set $\{(x_i, y_i) \mid x_i \in R^d, y_i \in \{0, 1\}, i = 1, \dots, N\}$, a linear SVM classifier can be described as :

$$F(x) = \sum_i a_i y_i x \cdot x_i + b = x \cdot \sum_i a_i y_i x_i + b = x \cdot w + b \quad (7)$$

Where $a_i \geq 0$, b is a bias term and d is the space dimension of the input data; \cdot is a dot-product operator, and w is the normal vector of the classification hyper-plane. The a_i has zero values for the training set except a small sub-set; this sub-set is called the support set and its elements is called the support vectors. The optimal hyper-plane is found such as to maximize the classification margin. The soft-margin optimization task can be formulated as:

$$\text{Minimize: } \sum_i C_i \xi_i^p + \frac{1}{2} \|w\|^2 \quad \text{subject to:}$$

$$y(w \cdot x + b) \geq 1 - \xi_i \quad (8)$$

Where $C_i \geq 0$, $p \geq 0$ and $\xi_i = 1 - y(w \cdot x + b)$; The slack variables ξ_i take non-zero values only for bound support vectors. In eq.(8) the accuracy over the training set is balanced by the "smoothness" of the solution. In our system, we set p to 1, and in this case a number of highly efficient computational methods have been developed.

The features fed to SVM-classifier include: 1) the pixel number of the biggest skin region; 2) the number of skin regions; 3) the mean of the pixel number of skin regions; 4) the center coordinates of the biggest skin region; 5) the fractal dimension of the biggest three skin regions; 6) the face number and the locations; 7) the skin pixel number in the face areas.

6 Experimental Results

In our experiments, the database consists of 2000 pornographic images and 5000 normal images which are all loaded from the internet randomly. The topics of the 5000 normal images include animals, scenes, close-up faces, things and humans. The experimental results indicate that the proposed system is effective with the detection of pornographic images.

Table 1. Pornographic image detection results with a variety of images.

| Image type | Pornographic | Animals | Close-up faces | Scenes | Humans | Things |
|------------|--------------|---------|----------------|--------|--------|--------|
| Image num | 2000 | 1000 | 1000 | 1000 | 1000 | 1000 |
| Alarm num | 1892 | 21 | 62 | 46 | 165 | 87 |

7 Discussions and Conclusions

In this paper, we have proposed an automatic system for pornographic image detection. Through a quick skin detection algorithm with section function in RGB color space, a coarse-degree-based texture modifying, we get the skin regions and then calculate the useful information from it. Fractal dimension is utilized to distinguish the desert-type images from the pornographic images. The close-up face images are excluded by the face detection operation. The experimental results show our system can detecting pornographic images with high recall ratio and precision ratio.

Acknowledgment. This work is supported in part by the subject of Zhejiang Education Plan, China (Grant No. SCG215) and in part by Students Science and Technology Innovation Activities Program of Zhejiang Province (Grant No. 2010R417008) and in part by Nonprofit Technology and Application Research Program of Zhejiang Province, China (Grant No. 2010C31083).

References

1. Fleck, M.M., Forsyth, D.A., Bregler, C.: Finding Naked People. In: Buxton, B.F., Cipolla, R. (eds.) ECCV 1996. LNCS, vol. 1065, pp. 593–602. Springer, Heidelberg (1996)
2. Drimborean, A.F.: Image processing techniques to detect and filter objectionable images based on skin tone and shape recognition. In: Proceedings International Conference on Consumer Electronics, pp. 278–281 (2001)
3. Zheng, H., Liu, H., Daoudi, M.: Blocking objectionable images: adult images and harmful symbols. In: Proceedings of the IEEE International Conference on Multimedia and Expo. (ICME), pp. 1223–1226 (June 2004)
4. Xu, Y., Li, B., Xue, X., Lu, H.: Region-based pornographic image detection. In: IEEE 7th Workshop on Multimedia Signal Processing (MMSP), pp. 1–4 (November 2005)
5. Lee, J.-S., Kuo, Y.-M., Chung, P.-C.: Detecting Nakedness in Color Images. In: Sencar, H.T., Velastin, S., Nikolaidis, N., Lian, S. (eds.) Intelligent Multimedia Analysis for Security Applications. Studies in Computational Intelligence, vol. 282, pp. 225–236. Springer, Heidelberg (2010)
6. Lee, J.-S., Kuo, Y.-M., Chung, P.-C.: The adult image identification based on online sampling. In: Proceedings of the International Joint Conference on Neural Networks (IJCNN), pp. 2566–2571 (July 2006)
7. Shen, X., Wei, W., Qian, Q.: The filtering of internet images based on detecting erotogenic-part. In: Proceedings of the Third International Conference on Natural Computation (ICNC), pp. 732–736. IEEE Computer Society, Washington, DC, USA (2007)
8. Sevimli, H., Esen, E., Ateş, T.K., Ozan, E.C.: Adult Image Content Classification Using Global Features and Skin Region Detection. LNEE, vol. 62(1), pp. 253–258 (2010)
9. Jones, M.J., Rehg, J.M.: Statistical color models with application to skin detection. In: Proceedings of Computer Vision and Pattern Recognition, pp. 274–280 (1999)
10. Kováč, J., Peer, P., Solina, F.: Human Skin Colour Clustering for Face Detection. In: Proceedings of EUROCON, pp. 144–148 (2003)
11. Mandelbrot, B.B.: The fractal geometry of nature, p. 125. W H Freeman and Company, New York (1983)

12. Cao, L.: Machine recognition of citrus variety based on the fractal dimensions of perimeter-area. Transactions of the CSAE 26(2), 351–355 (2010)
13. Liu, C.: A bayesian discriminating features method for face detection. IEEE Trans. Pattern Analysis and Machine Intelligence 25, 725–740 (2003)
14. A jörgensen. AdaBoost, Histograms for Fast Face Detection. Master of Science Thesis Stockholm, Sweden (2006)
15. Guo, G., Li, S.Z., Chan, K.L.: Support vector machines for face recognition. Image and Vision Computing 19(9-10), 631–638 (2001)

Secure Map Reduce Data Transmission Mechanism in Cloud Computing Using Threshold Secret Sharing Scheme

Hua Yi Lin¹, Che-Yu Yang¹, and Meng-Yen Hsieh^{2,*}

¹Department of Information Management, China University of Technology,
Taiwan, ROC

²Department of Computer Science and Information Engineering,
Providence University, Taiwan, ROC

Abstract. With the fast development of the internet network, combining the structure of distributing operation is paid attention to gradually. Especially on Internet, numerous personal computers share their resources such as CPU, memory, disk and network with each other. Cloud computing exploits this advantage of integrating with many operational resources over Internet to reach borderless services. Additionally, cloud computing owns strong arithmetic capacity. Through login, users perform any program [1, 2] and carry out the expected operation result. However, personal computers are exposed to open environments, and arbitrarily accept anonymous users to sign in and take personal information away. Therefore, the security issue of cloud computing becomes an important challenge. At present, most of cloud computing exploit Map and Reduce functions to perform divide and merge message respectively that lack information protection mechanisms. This study proposes a secure Map Reduce mechanism using a threshold secret sharing scheme to enhance the security of transmitted data.

Keywords: Cloud computing, Map, Reduce, Threshold secret sharing.

1 Introduction

Cloud computing is based on unreliable computers providing parallel computing, and consists of three types of computers. There are Master, Mapper and Reducer computers. In cloud computing, computers accept Map/Reduce task assignments, cooperatively perform Map/Reduce operations, and deliver data. The Master computer is responsible for receiving requested tasks of users, dividing tasks into pieces of processes, and assigning processes to Mappers and Reducers. The entire cloud computing infrastructure is as shown in Fig. 1.

Different from traditional program developments, programmers must define parallel processing datagram, and develop Map programs to perform computing on numerous computers. Eventually, each Reducer collects intermediate data from

* Corresponding author.

Mappers and integrates them into a complete message [3, 4]. However, malicious computers probably impersonate Mappers (or Reducers) to modify or send incorrect messages to users. Therefore, this study proposes a secure cloud computing infrastructure which provides digital signature and secure data transmission agreements using threshold secret sharing mechanisms to achieve secure Map Reduce operations.

Since Master is responsible for determining which computers have to perform Map and Reduce operations. Initially, Master dispatches the Map task to Mappers, and divides the message (input data) into several datagrams. After calculation, Mappers save the results into their storage. Eventually, Reducers read intermediate data and perform reduce functions to merge intermediate data into a complete message.

This study concentrates on exploiting threshold secret sharing scheme to enhance Map/Reduce secure functions, and exploits hash message authentication code (HMAC) to verify the integrity of transmitted information. Moreover, this study develops several mechanisms to ensure identities among participating in computers and enhance secure data transmission.

The rest of this study is structured as follows. Section 2 introduces the threshold secret sharing mechanism. Section 3 describes the proposed secure Map/Reduce data transmission scheme. Section 4 presents the analyses of the proposed scheme and computing evaluation. Conclusions and future work are finally drawn in Section 5.

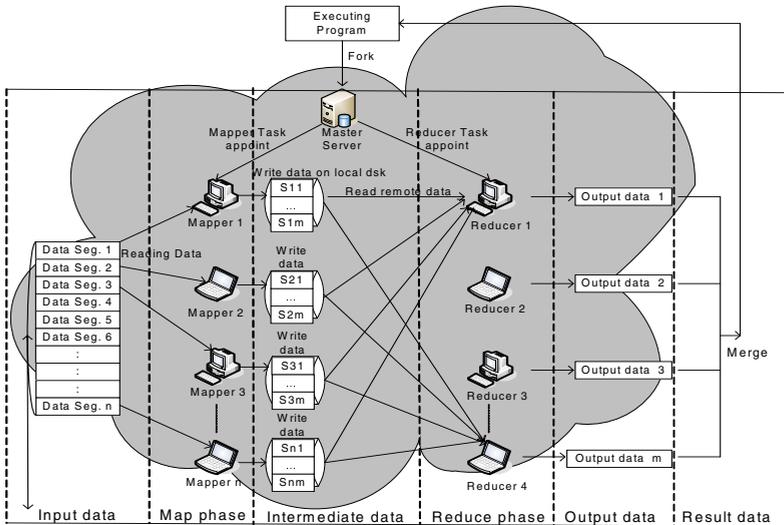


Fig. 1. Map and Reduce operation phases infrastructure

2 Threshold Secret Sharing Scheme

Currently, most cloud computing environments are based on virtual machine (VM) infrastructures to achieve distributed computing, as shown in Fig. 2. Therefore, this



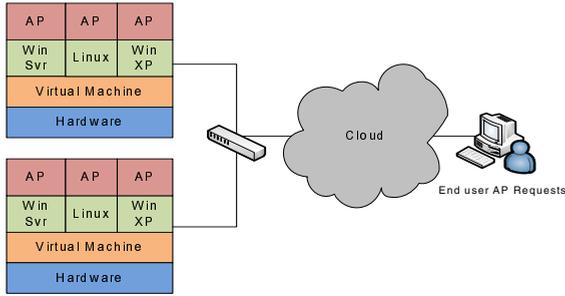


Fig. 2. Virtual machine infrastructure in cloud computing

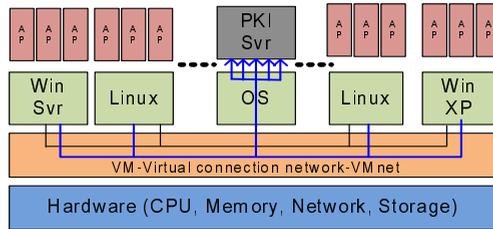


Fig. 3. PKI in virtual machine infrastructure

study builds up a VM named PKI VM as a Master, a VM as a Reducer VM, and several VMs as Mappers VMs. The entire VM system is as shown in Fig. 3. Additionally, our proposed secure operations are based on improved threshold secret sharing scheme to achieve secure cloud computing. This mechanism can solve the aforementioned issues and supply more secure environments. The following presented schemes attempt to intelligently distribute the authority functionality among a number of nodes. Although this concept has been proposed earlier, its application on cloud computing is novel.

To achieve fault tolerance of security and perfect secret in cloud computing, a threshold scheme [5, 6] is used. A (n, k) threshold cryptography scheme allows n Mapper VMs to share the ability performing a cryptographic operation (such as creating a digital signature), so that any k Mapper VMs can perform this operation jointly, whereas it is infeasible for a maximum of $k-1$ Mapper VMs to do so, even if they collude.

In the proposed model, the n Mapper VMs share the ability to recuperate the system-secret-key SK . To ensure the service can tolerate $k-1$ compromised Mapper VMs, a (n, k) threshold cryptography scheme is employed and SK of the service is divided into n shadows (secret sharing keys) $(Sh_1, Sh_2, \dots, Sh_n)$, and each shadow is assigned to each Mapper VM. Figure 4 illustrates a $(n, 3)$ threshold cryptographic scheme. Each Mapper VM_i has shadow sh_i . Correct Mapper VMs (1, 3 and n) submit their own shadow Sh_i to PKI VM (a dealer and combiner). Even if other Mapper VM_i ($i \neq 1, 3$ and n) fail to submit their shadows, PKI VM can generate the correct SK . However, compromised Mapper VM_i (of which there are a maximum of two) cannot generate SK correctly, because they can generate a maximum of two partial shadows.

Subsequently, we adopt the aforementioned module on Map/Reduce operations in cloud computing. Initially, PKI VM calculates SK , and then divides SK into n secret sharing key $(Sh_1, Sh_2, \dots, Sh_n)$. PKI VM is responsible for distribute n pieces to n Mapper VMs. When the entire system collapses, the new PKI VM merely collects $k-1$ secret sharing keys from Mapper VM_i , thus can reconstruct the original system key SK . After the secret sharing operations, each Mapper VM_i owns secret sharing key sh_i .

In our secret sharing design, the system key pair is denoted as (PK, SK) , where PK represents the system public key, SK is the system secret key, and PKI VM generates the (PK, SK) . Once a node participates in or departs from the system, PKI VM regenerates a new system key pair (PK, SK) according to the proposed scheme. The newer PK is well known to all Mapper VM_i , and SK is divided into n shadows (secret sharing keys) $(Sh_1, Sh_2, \dots, Sh_n)$ by the secret sharing approach. Initially, PKI VM encrypts Sh_i using the public key Pch_i of Mapper VM_i , and then sends the encrypted Sh_i to Mapper VM_i . After receiving, Mapper VM_i decrypts the encrypted Sh_i using the corresponding public key Pch_i . In the proposed approach, SK is shared by arbitrary Mapper VMs using a secret polynomial $f(x)$. If the degree of $f(x)$ is $k-1$, then any k Mapper VMs can reconstruct the secret key using Lagrange Interpolation, while any numbers of Mapper VMs less than k reveal no information regarding SK . In the scenario considered here, PKI VM generates the secret key $SK = S_d$ and randomly chooses a polynomial $f(x)$ of degree $k-1$, $f(x) = S_d + f_1 \cdot x + \dots + f_{k-1} \cdot x^{k-1}$, where S_d and f_i ($i = 1, 2, \dots, k-1$) are less than random prime p , and the shared secret is $f(0) = S_d \pmod p$. Each Mapper VM_i ($i = 1, 2, \dots, n$) holds a shadow (secret sharing key) $Sh_i = (f(VM_i) \pmod p)$. With any construction of k Mapper VMs ($VM_1, VM_2, VM_3, \dots, VM_k$), this study derives that

$$S_d \equiv [\sum_{i=1}^k Sh_i \bullet l_{VM_i}(0)] \pmod p \equiv [\sum_{i=1}^k SK_i] \pmod p ,$$

where the $l_{VM_i}(x) = \prod_{j=1, j \neq i}^k \frac{(x - VM_j)}{(VM_i - VM_j)}$ is the Lagrange Coefficients, each share holder VM_i can calculate SK_i from its shadow Sh_i via Lagrange Interpolation, and then the $SK (=S_d)$ is reconstructed from the sum

$$S_d = [\sum_{i=1}^k SK_i] \pmod p .$$

The polynomial secret sharing system should be protected against attacks because adversaries could break into k or more Mapper VMs in enough attack time. The system has to refresh periodic secret sharing updates with different polynomials using a proactive secret sharing mechanism. This study constructs another polynomial f_{k+1} from $f_k, f_{k+1} = f_k + g_k$, where g_k is a random degree $(k-1)$ polynomial, and the new secret share $f_{k+1}(VM_i) = f_k(VM_i) + g_k(VM_i)$ can reconstruct the S_d . If the shadow of a Mapper VM expired or compromised, the Mapper VM is considered a suspect. The system will store the shadow and the Mapper VM in its CRL using a suspect accusation factor and forward the accusation to other Mapper VMs, which store it in their revocation list with a suspect accusation and decrease the secure weight. The suspect is convicted when a



Mapper VM acquires k accusations. Based on this model, this study provides robust PKI VM fault tolerance. Once PKI VM collapses or any k Mapper VMs convict it of malicious behaviors, a election algorithm elects a new PKI VM and then cooperates with other Mapper VMs to reconstruct the system-secret-key SK from the k shadows $(Sh_1, Sh_2, \dots, Sh_n)$, after which the system regains normality.

3 Map/Reduce Secure Data Transmission Agreement

Since the Map/Reduce operations in cloud computing probably encounter malicious attacks, we need a secure mechanism to protect the entire operations among PKI VM, Mapper VMs and Reducer VMs. Simultaneously, we must ensure the identity and security of the joining Mapper VMs, and prevent from malicious nodes impersonating Mapper VMs. Additionally, Reducer VMs receive intermediate data from Mapper VMs to reconstruct the entire data, must confirm the accuracy of Mapper VMs and ensure the integration of transmitted data to avoid receiving tampered data. Based on the security issues, this study exploits threshold secret sharing scheme to achieve secure Map/Reduce operations. The proposed secure data transmission agreement is as shown in Fig. 5, and the detailed procedures are presented as follows.

(1)Initial phase: When PKI VM receives requests from users, it sends assigned tasks to Mapper VMs. For further ensuring the accuracy of the sender, PKI VM has to sign the request message, PKI VM ID and transmitted data for Mapper VM later verifying the identity of PKI VM.

$$[\text{Request} | \text{ID}_{\text{PKI-VM}_i}, \text{Data}_{\text{req}}]_{\text{Sig-of-PKI-VM}}$$

After receiving the transmitted request, Mapper VM_i exploits the public key of PKI VM to verify the identity of PKI VM, and then signs the reply message, Mapper VM_i ID and transmitted data.

$$[\text{Reply} | \text{ID}_{\text{mapper-VM}_i}, \text{Data}_{\text{req}}]_{\text{Sig-of-mapper-VM}_i}$$

(2)Subsequently, PKI VM takes the data segment $Data_{\text{segi}}$ as input, and calculates $\text{HMAC}(Data_{\text{segi}})$ with original data $Data_{\text{segi}}$ and *Time stamp* and Mapper VM_i ID. Eventually, PKI VM signs the entire data using the secret sharing key Sh_i of the receiver, and then sends the results to Mapper VM_i .

$$[Sh_i, [\text{ID}_{\text{map-VM}_i}, \text{Time}, \text{Data}_{\text{segi}} || \text{HMAC}(Data_{\text{segi}})]]_{\text{Sig-of-PKI-VM}}$$

Once Mapper VM_i receives the transmitted data, it decrypts the encrypted data and verifies the Sh_i and the integrity of HMAC.

(3)Second Phase: When Reducer VM receives assigned tasks from PKI VM to perform reducing procedures. To ensure the accuracy of the sender, PKI VM must sign the request message, PKI VM ID, and transmitted data for Reducer later verifying the identity of PKI VM.

$$[\text{Request} | [\text{ID}_{\text{PKI-VM}}, \text{Data}_{\text{req}}]]_{\text{Sig-of-PKI-VM}}$$

After Reducer VM receiving the transmitted data, the Reducer VM confirms the identity of PKI VM using the public key of PKI VM, and then signs the reply message, Reducer ID, and transmitted data.

$$[\text{ReplyID}_{\text{reducer}}, \text{Data}_{\text{req}}]_{\text{Sig-of-reducer}}$$

(4)Reducer VM receives transmitted data segments $\text{Data}_{\text{Seg}1} \sim \text{Data}_{\text{Seg}n1}$ from Mapper VM_i , which are encrypted with *Time stamp* and *SeqNo* using the secret sharing key Sh_i of Mapper VM_i .

$$\begin{aligned} &\text{Mapper VM}_{(1 \sim n)} \rightarrow \text{Reducer VM}_i \\ &[\text{Data}_{\text{Seg}1}, \text{Time}, \text{SeqNo}]_{\text{Sig-using-Sh}_i} \end{aligned}$$

(5)When Reducer VM receives the transmitted data from Mapper $\text{VM}_{(1 \sim n)}$, Reducer VM asks the corresponding public key of Sh_i of Mapper VM_i from PKI VM to decrypt the received data.

$$\begin{aligned} &\text{Reducer VM} \rightarrow \text{PKI VM} \\ &[\text{RequestPub}_{\text{Sh}_i}]_{\text{Sig-of-Reducer-VM}_i} \end{aligned}$$

Eventually, Reducer VM obtains the corresponding public key of Sh_i , then decrypts the encrypted data, and merges each piece of received datagram into a complete message. Consequently, Reducer VM sends the complete message to the demanding user, and finishes the secure Map/Reduce operations.

The proposed scheme is rapid and efficient, and is capable of fault tolerance. When PKI VM collapses, the system only needs to collect k secret sharing key from Mappers VM_i , and then can reconstruct the system secret key SK . Additionally, Mapper VM_i and Reducer VM verify the integrity of transmitted data using the secret sharing key. *Time stamp* and *SeqNo* can assist system in avoiding malicious reply attacks. During the assigned task phases, PKI VM exploits signature to ensure the accuracy of Mappers VM_i and Reducer VM, thus can avoid impersonation attacks performing deny of service (DoS).

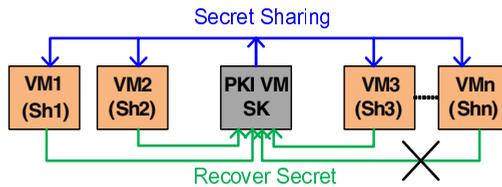


Fig. 4. Using a $(n, 3)$ threshold cryptographic scheme to generate the correct SK

4 Security Analyses and Performance Evaluation

This section provides several security analyses of the proposed method, and evaluates the performance of the proposed scheme.

4.1 Fault Tolerance of Threshold Secret Sharing Mechanism

According to the threshold secret sharing protocol with (n, k) threshold values, if PKI VM collapses, the system collects at most k sharing keys from Mapper VMs, and then



conserve a key pair to perform a few encrypting and decrypting operations, and preserves a hash function, such as HMAC-160 or RIPEMD-160, to verify the integrity of transmitted data. The entire system consumes a few resources. Consequently, the plain operations are highly suited for cloud computing environments.

4.5 Performance Evaluation

This study simulates the response time from the user sending requests to receive data. Also, we evaluate the execution time for the system processing variable data length. As shown in Fig. 6, when the data length increases, the transmission time increases, and the Map/Reduce operation time increases desirably. Moreover, Figure 7 demonstrates that although increasing the number of tasks, the Map/Reduce response time increases stably. The proposed scheme indeed performs well, even if a Map/Reduce operation with security functions.

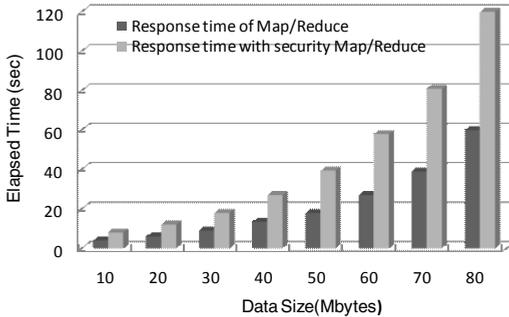


Fig. 6. The comparison of response time for variable data length

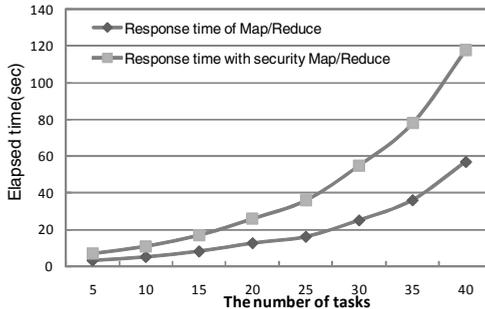


Fig. 7. The comparison of response time for increasing tasks



5 Summary

Cloud computing is the future tendency. How to provide a secure and efficient mechanism is a challenge. In Internet, since personal computers are exposed to open environments, anonymous users can randomly access the other computer resources, however, the owners don't realize that Map and Reduce operations probably reveal personal information. As a result, the security issue affects the intentions of using cloud computing for users. Therefore, given a secure data transmission mechanism is significantly important.

This study proposes a secure cloud computing architecture that provides security and fault tolerance Map/Reduce operations. The proposal can efficiently detect malicious attacks, ensure the integrity of transmitted data, and avoid identity impersonation. Additionally, this study exploits digital signature to ensure the identity of PKI VM, Mapper and Reducer, and the threshold secret sharing scheme is adopted to protect and verify the accuracy of transmitted data. Eventually, the simulation result demonstrates that the proposed scheme merely needs a few computing resources to achieve security operations, and is highly suited for cloud computing environments.

References

1. Devanbu, P., Gertz, M., Stuart, C.M., Stubblebine, G.: Authentic Third-party Data Publication. In: 14th IFIP 11.3 Working Conference in Database Security, Netherlands (August 2000)
2. Du, W., Jia, J., Mangal, M., Murugesan, M.: Uncheatable grid computing. In: The 24th International Conference on Distributed Computing Systems, pp. 4–11 (2004)
3. Zhao, S., Lo, V.: Result Verification and Trust-based Scheduling in Open Peer-to-Peer. In: Proceeding of the 5th IEEE International Conference on Peer-to-Peer Computing, pp. 31–38 (2005)
4. Itani, W., Kayssi, A., Chehab, A.: Privacy as a Service: Privacy-Aware Data Storage and Processing in Cloud Computing Architectures. In: the Proceeding of the 8th IEEE International Conference on Dependable, Autonomic and Secure Computing, pp. 711–716 (2009)
5. Desmedt, Y., Frankel, Y.: Threshold Cryptosystems. In: Brassard, G. (ed.) CRYPTO 1989. LNCS, vol. 435, pp. 307–315. Springer, Heidelberg (1990)
6. Stinson, D.R., Wei, R.: Unconditionally Secure Proactive secret Sharing Scheme with Combinatorial. In: Heys, H.M., Adams, C.M. (eds.) SAC 1999. LNCS, vol. 1758, p. 200. Springer, Heidelberg (1999)

Superior-in-Status Analysis of Improved Genetic Algorithm for GTSP

Tan Yang¹, Hao Zhi-feng^{1,2}, Cai Zhaoquan³, and Huang Han¹

¹ School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China

² School of Computer, Guangdong University of Technology, Guangzhou 510641, China

³ Educational Technology Center, Huizhou University, Huizhou 516007, Guangdong China
tjgdwjq@sina.com

Abstract. Superiority in status relation (\succ) can be used to rank the given EAs in terms of convergence capacity. The performance of an EA can be improved if it is modified to be superior in status to its original version. In this paper, the \succ relation model is applied to analyzing the improvement of generalize-chromosome genetic algorithm (GCGA) for generalized traveling salesman problem (GTSP). Hybrid-chromosome genetic algorithm (HCGA) is superiority to GCGA. The numerical results also indicate that HCGA performs better and more steadied than GCGA in solving several GTSP instances. The case is the application example of the proposed relation model.

Keywords: Swarm-based optimization methodology, Genetic algorithm, Superior-in-status analysis.

1 Introduction

Since genetic algorithm [1] was first proposed, several studies have been published on the convergence analysis of EAs. Some preliminary results like pattern theorem [1], Walsh analysis [2-3], and blocking assumption [4-5], were devoted to the idiographic performance of genetic algorithms (GAs). Later, Markov chain analysis was done on the convergence of EAs, which mentions whether an evolutionary algorithm can converge to the global optimal solution.

Most of the convergence results [10-13] mostly focus on the case of known optimal solution. When the optimal solution of the tackled problem is unknown, an approach to convergence contrast is necessary to EAs. However, few studies have touched upon the theory for the convergence performance comparison of EA. Most of the research on the topic was done and verified experimentally without a strictly-defined theoretical framework or model. Therefore, a relation model was proposed for analyzing the convergence capacity of evolutionary algorithm [6]. We will use a superior-in-status relation to analyze an improvement of genetic algorithm for GTSP.

2 Genetic Algorithm for GTSP and Its Improvement

A GCGA was proposed for GTSP problems. The testing results show that GCGA performed much better than RGA did with less computing cost and higher steady efficiency [7]. The framework and the encoding of GCGA were described in Fig.1 and Fig.2.

The crossover of GCGA contains crossover operators on the head and the body which can be found from Ref. [6] in detail. The following Fig. 3 is an example of the crossover on the head.

All of the offspring mutate by a small probability according to the random mutation point. Then, inversion operator runs on the body encodings of all the solutions.

1. Encoding.
2. Initialization. $t=0$.
3. Evaluation and Selection. $t=t+1$.
4. If the stop criterion is satisfied then
 Output the best-so-far solution.
 Else
 Go to step 5.
 Endif
5. Run crossover by probability 0.89
6. Run mutation on the solution generated by step 5 with probability 0.003
7. Run inversion on every offspring solution, go to step 3.

Fig. 1. The framework of GCGA

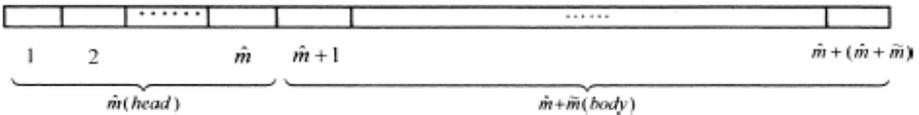


Fig. 2. The encoding of GCGA

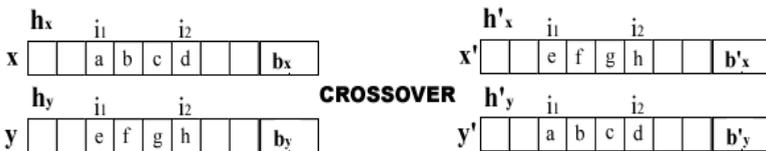


Fig. 3. An example of crossover on the head

Crossover is crucial to GCGA according to the framework. We modified the encoding (Fig. 4) and crossover operator on the head to design another genetic algorithm HCGA as an improvement of GCGA.

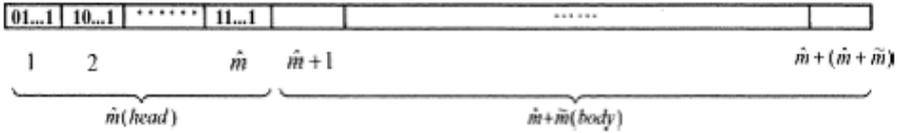


Fig. 4. The hybrid encoding of HCGA

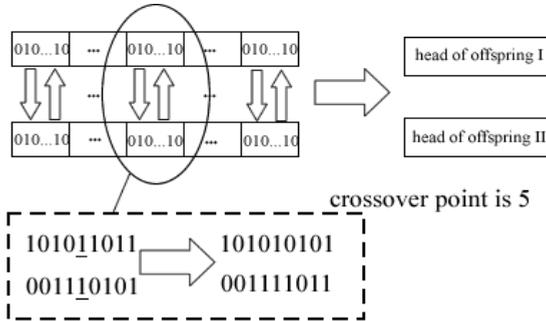


Fig. 5. An example of novel crossover on the head

The encoding of HCGA is hybrid including the head encoded with binary number and the same body as GCGA. The framework and operators of HCGA are the same as that of GCGA, except for the crossover on the head (Fig. 5). In the aforementioned crossover, single point crossover runs on \hat{m} positions (binary number strings) with \hat{m} random crossover points belonging to $\{1, \dots, \hat{m}\}$.

3 Improvement Analysis by Superior-in-Status Model

Based on the relation of equivalence in status [6], the section proposes an ordering relation to compare the given EAs in convergence capacity.

Definition 1. Given a set of EAs signed as S_{EA} and two Markov chains $\{\xi_t^a\}_{t=0}^{+\infty}, \{\xi_t^b\}_{t=0}^{+\infty}$ for $\forall a, b \in S_{EA}$, it is said that b is superior in status to a if $\Xi_t^a \cap Y^* \subseteq \Xi_t^b \cap Y^*$ for $t = 1, 2, \dots$, signed as $b \succ a$ or $a < b$.

$b \succ a$ means that EA b can attain all of the solutions found by EA a , and some of the solutions are not found by EA a . The following theorem and corollary [6] indicate the convergence relation between a and b if $b \succ a$. Their proofs can be found in Ref. [6].

Theorem 1. Given a set of EAs signed as S_{EA} and $\forall a, b \in S_{EA}$, b converges, if $b \succ a$, the stochastic process of b is an absorbing Markov chain and a converges.



Corollary 1. Given a set of EAs signed as S_{EA} and $\forall a, b \in S_{EA}$, a cannot converge if (1) $b \succ a$, (2) the stochastic process of b is an absorbing Markov chain and (3) b cannot converge.

This subsection will use theorem 1 and corollary 1 to prove that HCGA \succ GCGA by the following theoretical results.

Theorem 2. HCGA \succ GCGA.

Proof. Given two stochastic processes $\{\xi_t^g\}_{t=0}^{+\infty}$ and $\{\xi_t^h\}_{t=0}^{+\infty}$ for GCGA and HCGA, ξ_t^g and ξ_t^h belong to the discrete status set Y_{gtsp} because the algorithms all use discrete encoding chromosome.

An iteration of genetic algorithm can be considered as a status mapping $F = f^c \circ f^m \circ f^r \circ f^s$, where f^c, f^m, f^r, f^s stand for the operators of crossover, mutation, inversion and selection respectively, i.e. $\xi_{t+1}^g = F(\xi_t^g)$ and $\xi_{t+1}^h = F(\xi_t^h)$. The f^m, f^r, f^s of GCGA and HCGA are the same, and their f^c mappings are different because the algorithms have different crossover operation in the head of chromosome.

According to Fig.4 and Fig.5, the analysis can be given in the following. The head crossover of GCGA has only one step, but HCGA has two steps including the crossover in the head chromosome and the crossover in each binary crossover gene (shown as Fig.5). For the same parent chromosome and crossover point in the head, HCGA can generate the same possible offspring as GCGA if its crossover point in the gene is 0. However, HCGA can generate different offspring from GCGA according to other crossover points. Therefore, HCGA can attain all of the status which GCGA can attain and other status which GCGA cannot attain.

Given $\forall \xi'$ such that $\xi' \in Y_{gtsp} \wedge \xi' \notin Y_{gtsp}^*$ where Y_{gtsp}^* is the optimal status set,

$$\{\xi^x \mid P\{\xi_{t+1}^g = \xi^x \mid \xi_t^g = \xi'\}\} \subseteq \{\xi^x \mid P\{\xi_{t+1}^h = \xi^x \mid \xi_t^h = \xi'\}\}$$
 $(t = 0, 1, \dots).$

Because ξ^x can be arbitrary, $\Xi_t^g \subseteq \Xi_t^h$ ($t = 0, 1, \dots$) according to [6] definition 5. Thus, $\Xi_t^g \cap Y^* \subseteq \Xi_t^h \cap Y^*$, i.e. HCGA \succ GCGA. \square

Since HCGA is superior in status to GCGA, we have further conclusion in the following.

Corollary 2. HCGA converges if GCGA converges.

Proof. According to [6] theorem 3 and 4, HCGA converges if GCGA converges since
 HCGA \succ GCGA.
 \square

Corollary 3. GCGA cannot converge if HCGA cannot converge.

Proof. According to corollary 2 and [6] theorem 4, GCGA cannot converge if HCGA cannot converge. \square



The theoretical results indicate that HCGA has better convergence capacity than GCGA if $\forall \alpha \in Y^* \wedge \alpha \notin \Xi_t^g$ and $\exists \gamma \in Y^* \wedge \gamma \in \Xi_t^h$ ($t = 1, 2, \dots$), which is also verified by the numerical results in the following subsection.

4 Numerical Results

A conclusion can be drawn that the convergence capacity of HCGA is not worse than GCGA. Furthermore, HCGA can be better than GCGA in the case that GCGA cannot attain the optimal status which belongs to the possible status set of HCGA. The conclusion can be verified by the experimental results (table 1) of computing some benchmark problems [9]. Both of the algorithms ran 30 times in the experiment.

Table 1. Comparisons of results in 30 runs

| GTSP Problems | Optimal Length | HCGA Min | GCGA Min | HCGA Ave | GCGA Ave |
|---------------|----------------|----------|----------|----------|----------|
| 30KROA150 | 11018 | 11018 | 11018 | 11018 | 11022 |
| 30KROB150 | 12196 | 12196 | 12196 | 12196 | 12314 |
| 31PR152 | 51576 | 51576 | 51586 | 51576 | 53376 |
| 32U159 | 22664 | 22664 | 22664 | 22664 | 22685 |
| 40KROA200 | 13406 | 13408 | 13408 | 13408 | 13617 |
| 40KROB200 | 13111 | 13113 | 13120 | 13119 | 13352 |
| 45TS225 | 68340 | 68340 | 68340 | 68426 | 68789 |
| 46PR226 | 64007 | 64007 | 64007 | 64007 | 64574 |
| 53GIL262 | 1013 | 1013 | 1013 | 1013 | 1057 |
| 53PR264 | 29549 | 29549 | 29549 | 29549 | 29791 |
| 60PR299 | 22615 | 22631 | 22638 | 22638 | 22996 |
| 64LIN318 | 20765 | 20788 | 20977 | 20940 | 22115 |
| 80RD400 | 6361 | 6456 | 6465 | 6468 | 6604 |
| 84FL417 | 9651 | 9663 | 9663 | 9663 | 9725 |
| 88PR439 | 60099 | 60184 | 61273 | 60370 | 62674 |
| 89PCB442 | 21657 | 21768 | 21978 | 21825 | 22634 |

HCGA can find the optimal solution in the problems that the best solution can be attained by GCGA, which can verify the conclusion of corollary 3. HCGA can obtain the optimal length in some instances which GCGA cannot solve well, which can prove the correctness of corollary 4. The average performance of HCGA is better than GCGA in all of the testing benchmark problems.

The following figures indicate the comparison results of convergence trend between GCGA and HCGA. For the limit of the space, only the experimental results of five problems are indicated in detail.

The results of Fig.6-10 indicate that HCGA can converge to the better GTSP tour than GCGA at the end of the iterations. The numerical results confirm that HCGA has better convergence capacity based on $HCGA \succ GCGA$.

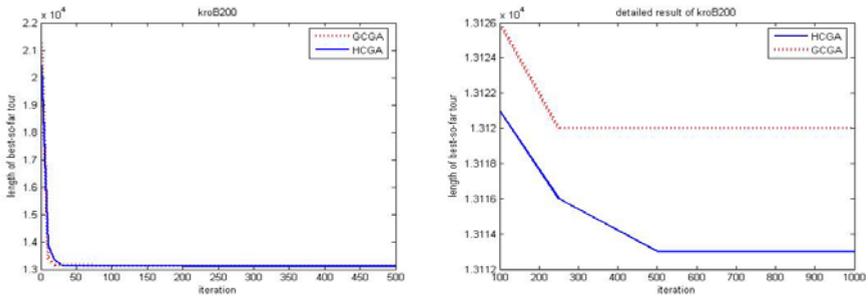


Fig. 6. Comparison results between GCGA and HCGA in the TSP problem kroB200

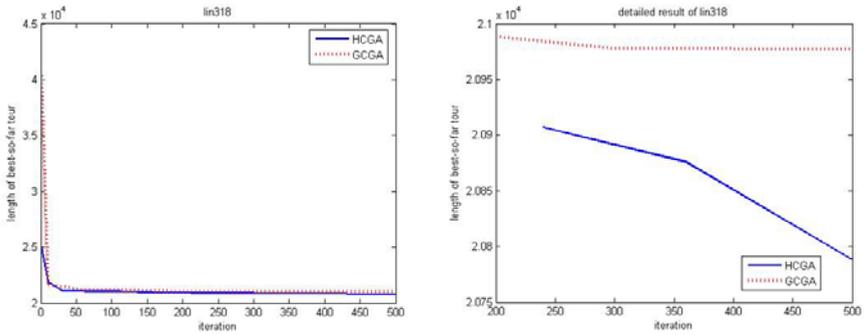


Fig. 7. Comparison results between GCGA and HCGA in the TSP problem lin318

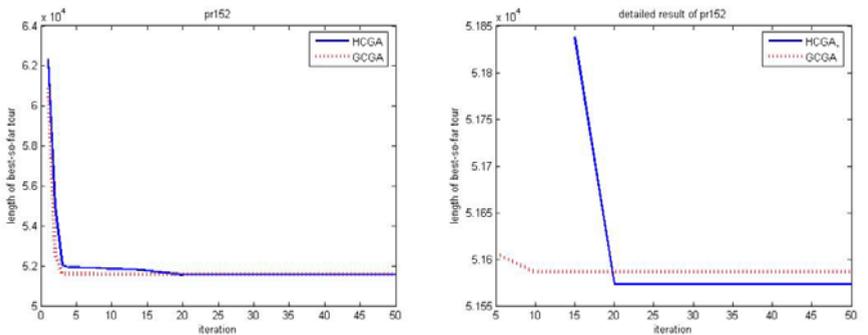


Fig. 8. Comparison results between GCGA and HCGA in the TSP problem pr152

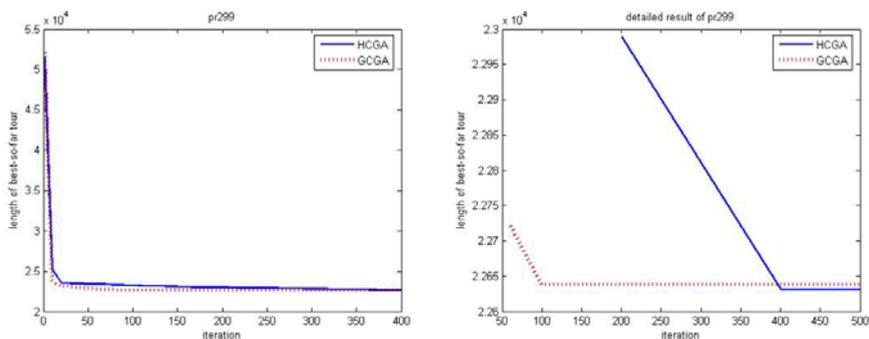


Fig. 9. Comparison results between GCGA and HCGA in the TSP problem pr299

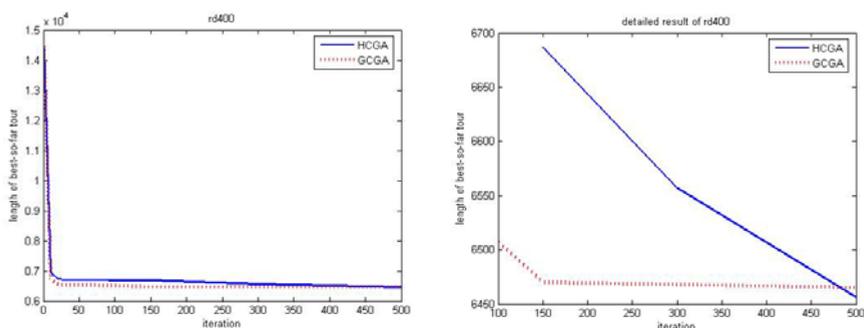


Fig. 10. Comparison results between GCGA and HCGA in the TSP problem rd400

5 Conclusions and Future Work

A case study of the improvement of genetic algorithm (GCGA) for GTSP problem is presented with the given method. GCGA is modified to be another genetic algorithm HCGA which has been proved to be superior to GCGA. The experimental results of computing several GTSP benchmark instances verify that HCGA has better convergence capacity than GCGA, which confirms the conclusion of the proposed theory. In the future study, the improved relation model will be studied to discuss the convergence and computational time in order to make more fundamental improvement of EAs. Moreover, we will use the improved model to study more instances of evolutionary algorithms.

Acknowledgement. This work is supported by Natural Science Foundation of China (61070033), Doctoral Program of the Ministry of Education (20090172120035), Natural Science Foundation of Guangdong province (9251009001000005, 9151008901000165, 10151601501000015), Science and Technology Planning Project of Guangdong Province (2010B050400011, 2010B080701070, 2008B080701005, 2009B010800026), “Eleventh Five-year Plan” Program of the Philosophy and Social

Science of Guangdong Province (08O-01), Opening Project of the State Key Laboratory of Information Security (04-01) and Fundamental Research Funds for the Central Universities, SCUT(2009ZM0052).

References

1. Holland, J.H.: *Adaptation in Natural and Artificial Systems*, 1st edn. MIT press, Cambridge, MA (1992)
2. Goldberg, D.E.: *Simple Genetic Algorithms and the Minimal Deceptive Problem*. In: Davis, L. (ed.) *Genetic Algorithms and Simulated Annealing*, pp. 74–78. London Pitman (1987)
3. Bethke, A.D.: *Comparison of Genetic Algorithms and Gradient-Vased Optimizers on Parallel Processors: Efficiency of Use of Processing Capacity (Technical Report No.197)*. University of Michigan, Login of Computers Group, Ann Arbor (1976)
4. Grefenstette, J.J.: *Deception considered harmful*. In: Whitley, L.D. (ed.) *Foundations of Genetic Algorithms*, Morgan Kaufmann, pp. 75–91. Morgan Kaufmann, San Mateo (1993)
5. Yu, Y., Zhou, Z.H.: *A new approach to estimating the expected first hitting time of evolutionary algorithms*. *Artificial Intelligence* 172, 1809–1832 (2008)
6. Hao, Z.-F., Huang, H., Li, H., Ling, S., Li, B.: *A Relation-Based Model for Convergence Analysis of Evolutionary Algorithm*. In: Panigrahi, B.K., Das, S., Suganthan, P.N., Dash, S.S. (eds.) *SEMCCO 2010. LNCS*, vol. 6466, pp. 190–197. Springer, Heidelberg (2010)
7. Wu, C.G., Liang, Y.C., Lee, H.P., Lu, C.: *Generalized chromosome genetic algorithm for generalized traveling salesman problems and its applications for machining*. *Physical Review E* 70(1), 16701-1–16701-13 (2004)
8. Laporte, G., Asef-Vaziri, A., Sriskandarajah, C.: *Some applications of the generalized traveling salesman problem*. *J. Oper. Res. Soc.* 47, 1461–1467 (1996)
9. Fischetti, M., Salazar, J.J., Toth, P.: *Branch-and-cut algorithm for the symmetric generalized traveling salesman problem*. *Operations Research* 45(3), 378–394 (1997)
10. Zhou, Y., He, J., Nie, Q.: *A comparative runtime analysis of heuristic algorithms for satisfiability problems*. *Artificial Intelligence* 173, 240–257 (2009)
11. Oliveto, P.S., He, J., Yao, X.: *Analysis of the (1 + 1)-EA for Finding Approximate Solutions to Vertex Cover Problems*. *IEEE Transactions on Evolutionary Computation* 13(5), 1006–1029 (2009)
12. Chen, T., Tang, K., Chen, G., Yao, X.: *Analysis of Computational Time of Simple Estimation of Distribution Algorithms*. *IEEE Transactions on Evolutionary Computation* 14(1), 1–22 (2010)
13. Oliveto, P.S., Witt, C.: *Simplified Drift Analysis for Proving Lower Bounds in Evolutionary Computation*. In: Rudolph, G., Jansen, T., Lucas, S., Poloni, C., Beume, N. (eds.) *PPSN 2008. LNCS*, vol. 5199, pp. 82–91. Springer, Heidelberg (2008)

The Development of Software and Its Influence on Digital Art

Cheng Yin¹, Shengyong Mu, Jing Liu, and Maopo Yin

¹ Shandong Water Polytechnic, Rizhao, China

² Rizhao Polytechnic, Rizhao, China

³ Jing-Yi Primary School, Rizhao, China

chen-hy654@tom.com

Abstract. The development of computer technology not only made great contribution to the development of art disciplines, but also brought some new digital artistic practices. The Free and Open-source software are widely spread on the Internet nowadays. A large number of digital art practitioners downloaded them to use or modify according to their individual needs. In this paper, we examine the development of free and open-source software and its characteristic, summarize the advantages of free and open-source software technology that allow digital art practitioners to have more freedom and flexibility, emphasize the influence of its openness and cooperation design philosophy on the dispersion characteristics of digital art.

Keywords: Digital art, free software, open-source software, free and open-source software, generative art.

1 Introduction

The term “digital art” has become widespread and popular in these years. Actually, digital art is a general term for a range of artistic works and practices that use digital technology as an essential part of the creative and/or presentation process. Early pioneers and researchers in digital arts can be traced back to the 1950s and 1960s. The most well-known digital art may be digital music which was one of the first suggested applications of computers when they were first invented.

Digital art, in a broad, is the digitalize of art discipline, such as utilizing digital technology as a means of graphic design, web-media communication and even mobile phone ring tones. As long as the digital technology was used as the carrier, it can be classified into digital art and possesses its independent aesthetic value. In the creative process, digital arts will fully or partially use the digital technology means. Nowadays, as a creative discipline, digital art includes a vast variety of forms and activities such as interactive media design, digital video art, virtual reality and new media arts. Interactive media design indicates the design that use interactive media as a carrier, such as the World Wide Web as the carrier of web design, web game design and mobile phone as the carrier of WAP design, mobile game. Digital video art includes digital animation, DV film and digital video advertisement. Virtual reality

refers to the digital museum, digital mall that involves virtual space design. New media art corresponds to the traditional fine arts, sculpture, which is a pure art form using the digital technology as the means and materials.

In a narrow sense, digital art generally refers to the art-related design, video, animation or other artistic work in which involves computers in their creative processing. Digital art, as opposed to the traditional art, has many advantages in the dissemination, storage, copying and other irreplaceable advantages in many aspects.

2 Categorize of Digital Art Practices

Digital art is based on computer hardware and software. Many software companies, such as the Adobe and Autodesk, are in the forefront and provide a good learning platform for digital art.

For many contemporary artists, the computer has already become a vital apparatus in their art practice. The impact of digital technology has changed the art activities such as painting, drawing and sculpture; while the new art forms, such as net art, digital installation art, and virtual reality, have become recognized artistic practices.

The use of computer in digital art practice, especially the software, can be categorized into two different kinds. One is the simulation based practice, in which computer is used to simulate traditional art environment and instruments. The other is the programming based practice, in which computer is used as a new instruments to explore new creation possibilities through programming.

2.1 Simulation Oriented Practice

Similar with other new technologies, at first, the computer tried to find its application in improving the efficiency of the previous technologies. For art discipline, the computer as a new technology was utilized to improve the way artists worked, and to enhance the skills and domain knowledge they were already familiar with. With the variety of software, the computer can almost emulate all the procedures that can be formally expressed, and rapidly became a popular way to assist creating art works.

Simulation based practice focus on the recreation of the traditional apparatuses through the computer, which can give the artists an efficiency, flexibility and practicality way of art work creation. In simulation based practice, the key point still remains at its traditional methods and aesthetic values. The fundamental artistic principles are unchanged, while the crating process is greatly improved. Meanwhile, the computer provides a significant continuity in the transition from the physical environment to virtual environment. Hence, artists can easily transfer their skill in digital world and the only change happens at the way in which they express their creativity.

In simulation based practice, the relationship between artists and the computer are very simple. Artists, as the end user of computer software, do not need lots of specific knowledge about the computer and have little influence on the development of the software.

2.2 Programming Oriented Practice

The programming ability is the most significant characteristics of the computer technology. For this reason the computer distinguishes from any other past artifact. In programming based practice, artists aimed to create artistic works through the programming, they use the computer as a tool trying to explore new creation possibilities of digital art.

By programming, artists can overcome the limitation of the traditional tools what they used. Artists can translate all kinds of ideas into codes which can be implemented by the computer very quickly. In this way, artists can obtain great freedom and possibilities for art creation. The computer as a new media of art practice greatly change artists work habit and ideology. Moreover, as a new media for art, it will inevitably create its fire-new and unique aesthetics. The “Software art” movement, which was popularized in the late 1990s, is a good example. In the “Software art”, artist’s concern shifted to the quality and beauty of codes, which means they thought more about the structure and efficiency of the code. The notion of ‘Software Art’ has been introduced as an attempt to describe a practice that is artistic, non-functional, reflexive and speculative with regard to the aesthetics and politics of software, and that takes the computer programming as the material proper of the artistic practice. Another example is the “Live coding” movement, in which the process of programming is considered as a part of art creation. The process of programming becomes as equally important as the final results generated by programming.

In programming based practice, the relationship between artists and the computer is more complex than that in simulate based practice. Artist’s role is not only the art creator but also the programmer.

3 The Development of Free and Open-Source Software

From the previous analysis, we can find a very strong relationship between digital art and computer software. Essentially, software is a kind of logical instructions. And for digital artists, it is more critical and complex than computer hardware. Moreover, software can be modified upon artist’s will to fulfill his idea, which can make a more directly impact on digital art.

Commercial software and the Free and Open-Source software(FOSS) are two different kinds of software that digital artists can have access to. Commercial software are usually developed by software companies and its distribution and source-code are all protected by intellectual property. Comparing with commercial software, the Free and Open-Source software has gained increasing attention in the public and digital art community in recent years. FOSS refers to the software that is liberally licensed to grant the right of users to use, study, modify, and improve its design through the availability of its source code. This approach has gained both momentum and acceptance because of its potential benefits which have been increasingly recognized by both individuals and corporations.

3.1 Free Software and Open-Source Software

Free and open-source software is an inclusive term which covers both free software and open-source software, despite describing similar development models, have differing cultures and philosophies.

Free software and open-source software are two influential movements which both have much influence on the development of Free and open-source software. Although they have some common issues and these two terms are often used interchangeably, but their ideology are totally different from each other.

The key point of free software movement is the ethical issues related to user's freedom when he or she uses the software. Free software focuses on the philosophical freedoms it gives to the users while open source focuses on the perceived strengths of its peer-to-peer development model. Beside the technical concern, free software movement also pays high attention to the social issues and tries to remedy the unfair in the use of software.

On the other hand, the open-source software movement only focuses on the technique issues. The developers believe that the development of software is based on its technical advantages and development model. In other words, the open-source strategy actually becomes the means by which to provide best software for the users.

The interpretation of the freedom in these two movements is very critical to understand the purpose of them. The Free Software Foundation, an organization that advocates the free software model, suggests that, to understand the concept, one should "think of free as in free speech, not as in free beer". As Stallman pointed out, "For the Open Source movement, non-free software is a suboptimal solution. For the Free Software movement, non-free software is a social problem and free software is the solution". In the context of free and open-source software, free refers to the freedom to copy and re-use the software, rather than the price of the software.

3.2 Development of Free and Open-Source Software

The development of Free and open-source software was highly influenced by the hacker culture, Free Software Foundation and Open Source Initiative.

The notion of exchanging and sharing software and its source codes can be traced back to the hacker culture in 1960s. One of the motivations of the early hackers was to develop a liberal approach towards the knowledge on the Internet. Furthermore, the "gift culture" embedded in hacker's practice that made great contribution to the hacker culture had a sense of community and freedom philosophy. To gain the acceptance of peers or increase one's reputation, one should make contributions to the community. Similarly, the influence of "gift culture" still can be found in free and open-source software.

The hacker culture gradually disappeared due to the commercial software company's prosperity. In order to preserve the original hacker culture, Richard Stallman started the GUN project and Free Software Foundation (FSF). The goal of these movements is to produce a complete operation system in which both the software and its source codes are free from commercial restrictions. The initiation of FSF marked the beginning of the free software movement, and the consequent development of FOSS.

The open source initiative was formed in 1998 to facilitate and encourage the transition from “free software” to “open-source software”. The change proved to be an unprecedented success because it gained a lot of supports from commercial software companies. The community and the industry built a new relationship which provided external support for the development of open source software.

3.3 Characteristic of Free and Open-Source Software

FOSS, as an influential movement, has its unique ideologies. These ideologies can be identified as follow: freedom in software, mutual collaborations, resource sharing and voluntary and distributed efforts. Accordingly, we can summarize the characteristic based on its ideologies.

Freedom in the software is the most important ideology and characteristic of FOSS. In FOSS, freedom contains the use, modification, distribution and redistribution of software and its source codes. In details, as the free software's definition which is provided by the Free Software Foundation, the freedom includes:

- 1) The freedom to run the program for any proposes.
- 2) The freedom to study how the program works, and adapt it to you needs. Access to the source code is a precondition for this.
- 3) The freedom to redistribute copies so you can help your neighbor.
- 4) The freedom to improve the program, and release your improvements to the public, so that the whole community benefits. Access to the source code is a precondition for this.

The mutual collaborations are considered as the key elements to propel the development of FOSS. Meanwhile, the notion of collaboration also influences the digital art community.

4 The Influence of Free and Open-Source Software on Digital Art

Free and open-source software can provide a reliable software platform to support digital artists to convey their creativity. And based on the aforementioned analysis, we can see that it fully meets the criteria that excellent software should possess for digital art:

- 1) It encourages flexible use of software.
- 2) It allows artists to influence the development of software design.
- 3) It enables artists to customize the software according to their own requirements.
- 4) It provides a platform for digital art practitioners, such as artists, engineers and programmers.

For its unique characteristic, FOSS becomes more and more popular in the digital art community in recent years. The influence of FOSS on digital art can be summarized as follow:

4.1 Decrease the Cost of Art Practice

By using FOSS, artists do not need to buy commercial software such as Photoshop, AutoCAD and so on. They can utilize their limited budget on some other more important things. Moreover, the update of the software is also free to them.

As to the hardware, FOSS can support all kinds of hardware, from traditional servers and desktop to the new generation of personal devices. This virtue means digital artists who use FOSS in their creation do not need to upgrade their hardware too frequently. For the old hardwares, FOSS can support them without additional requirements. For the new emerging hardwares, such as hand-held devices, many manufacturers started to adapt the operation system for their products.

4.2 Increase the Freedom in the Art Practice

With the computer technology becomes more miniaturized and high-speed Internet becomes more available, digital artists are now experiencing more freedom in the virtual world instead of the traditional physical world. Freedom of expression is central to any of the creative practices, and digital art is no exception. For digital artists, the most important is the tools and media though which artists can express their creative ideas clearly and thoroughly.

Using commercial software, artists are limited by the functions that provided by software company and they cannot change any core part of the software. However, FOSS provides the fully freedom to the artists. The more important thing is that the freedom provides by FOSS is not free from the physical and practice limitation, but the usage of creative tools and media. For example, computers may have some limitations on processing speed and memory size; artist's work cannot exceed the physical ability of the hardware. The freedom of the artists embodies when they can use the computer in any way as they wish.

The most used FOSS softwares in digital art community are Pure Data, SuperCollider, pure.dyne, PacketForth, Fluxus and so on. All of them have their own user community and online support resources.

5 Enlarge the Digital Art Community

With the development of Internet and computer technology, anyone who is familiar with the computer can easily create graphics, music and animation with only a little effort involved. Moreover, digital artists are now more connected as a world community; ideas are created and communicated globally very quickly.

The ideology of FOSS also affected digital art community, which propeled artists to create a more flexible way to cooperate, exchange ideas and distribute art works. More and more collaborative groups come out; they work together and share information and resources with each other. These collective groups are independently formed without formal support form external organization. The most popular and influential groups are Goto10, OpenLab, Dyne, Bek and so on. The collective environment provides the opportunity for the artists to exchange their ideas and skills, and get feedback through peer view.

6 Conclusion

We can see that FOSS is highly suitable for digital art practices. FOSS provides a stable and sustainable platform for the digital art creation. Nowadays, many FOSS tools have been used in digital art community. The success of FOSS in digital art domain will continue and it is foreseeable that FOSS will be one of the mainstreams in the digital art.

References

1. Davis, D.: The Work of Art in the Age of Digital Reproduction. *Leonardo* 28(5) (1995)
2. Marchese, S.M., Marchese, F.T.: Digital Media and Ephemeralness: Art, Artist, and Viewer. *Leonardo* 28(5) (1995)
3. Jones, S.: Towards a Philosophy of Virtual Reality: Issues Implicit in "Consciousness Reframed". *Leonardo* 33(2), 125–132 (2000)
4. Yellowlees Douglas, J., Hargadon, A.: The Pleasure of Immersion and Interaction: Schemas, Scripts, and the Fifth Business. *Digital Creativity* 12(3), 153–166 (2001)
5. Gombrich, E.H.: *Art and Illusion*. Princeton University Press, Princeton (1960)
6. Danto, A.C.: *The Philosophical Disenfranchisement of Art*. Columbia University Press, New York (1986)
7. Holtzman, S.: *Digital Mosaics: The Aesthetics of Cyberspace*. Simon&Schuster, New York (1997)
8. Darley, A.: *Visual Digital Culture: Surface play and spectacle in new media genres*. Routledge, London (2000)
9. Manovich, L.: *The Language of New Media*. The MIT Press, Massachusetts (2001)

An Evaluation Model for Synchronized Supply Chain Based on Entropy Weight and TOPSIS I

Li Zhiyong, Wang Xiang, and Zheng Nan

Zhengzhou Institute of Aeronautical Industry Management,
Zhengzhou 450015, China
leesilee@126.com

Abstract. A detail comparison of many relevant indexes is made towards the performance evaluation of synchronized supply chain. Combining to the design ideal which supply chain need to considerate short-term goals and strategic goals, the indexes system based on balanced scorecard is picked up. Because of the characteristics of synchronized supply chain, Entropy weight and TOPSIS is selected to do the performance measurement, A performance evaluation model is established based on Entropy weight and TOPSIS.

Keywords: Supply Chain, Synchronized, Balanced Scorecard, Entropy Weight, TOPSIS.

1 Introduction

Peter (1983) [1] proposed the concept of the supply chain. Faced with the situation of global economic integration and diversification demand of consumer, as well as the need to reduce product costs, improve product quality and accelerate product cycles, many suppliers, manufacturers and distributors formed a need-supply cooperative relationship, which is abstracted to supply chain management. It is generally believed that supply chain management is the applications of a group of methods to integrate suppliers, manufacturers, warehouse and inventory, which is used to minimize the total cost and satisfy service levels so as to product and distribute goods in right number, place and time .

With the development of the supply chain concept, Anderson & Lee (1996) [2] proposed the concept of Synchronized supply chain, giving more emphasis on Synchronized work of partners to quickly deal with customers requirement and maintain every member's competitive advantage to gain greater profits. Collaboration is a physics concept, and focused on the evolution of system from disorder to order. But introducing collaboration to management science is to achieve the value of the whole organization greater than the value of the independent components; Similarly, introducing the concept of collaboration to supply chain is to maximize the overall efficiency.

Synchronized Supply Chain is currently divided into three levels: strategic level, policy level and technical level. Based on the technical level of synchronized supply chain, this paper is mainly to solve the performance evaluation of synchronized supply

chain. The practical significance of the study is that only the proper evaluation on the performance of synchronized supply chain is made can its strategic and policy level adjusted properly. However, the study on this aspect at home and abroad is fewer, only Shi Chengdong et al (2009) [3] used rough sets and neural networks to solve the problem, but this evaluation method has problem of needing historical training data and complex of implementation. This paper attempts to find an easy and feasible method to evaluate the performance of synchronized supply chain, which can be more instructive and workable to the problem.

2 Index System of Synchronized Supply Chain Performance Evaluation

Synchronized supply chain performance evaluation is a new problem. But this paper believes that it is equivalent to the evaluation of supply chain performance after the concept of collaboration is introduced to supply chain. However, supply chain performance evaluation is studied widely. Hong Weimin et al (2006) [4] think that the evaluation index system of supply chain can be divided into three categories: the evaluation indexes system based on supply chain operations reference-model, the evaluation indexes system based on balanced scorecard and the evaluation indexes system based on resource-output and flexibility.

At present the supply chain operations reference-model is the most influential and the most widely used reference model, therefore we can utilize this model to carry on supply chain performance evaluation. Based on this model, "Chinese Enterprise Supply Chain Performance Evaluation Reference Model (SCPR1.0) Constitute Plan" was issued in October 2003. It is based on not only the internal business processes but also the external business processes to evaluate supply chain. Such result shows that there are too many indexes, and operability is not strong. SCPR1.0, for example, has a total of 45 third-level indexes, which makes the evaluation difficulty increase greatly.

The Balanced Scorecard techniques proposed by Kaplan & Norton (1992) [5] are widely used to various performance evaluation. Its biggest feature is to combine short-term goals and long-term goals, financial indicators and non-financial indicators, lagging indicators and leading indicators, internal performance and external performance, and makes the manager's attention transfer from implementing short-term goals to giving consideration to short-term goals and strategic goals. This feature is very fit to design concept of supply chain system, so Ma Shihua et al (2003) [6] and Ye Chun-Ming (2005) [7] apply the basic principles of balanced scorecard to supply chain performance evaluation, and combine with its feature to establish evaluation system of supply chain balanced scorecard. This system establishes indexes from the financial value perspective, business process perspective, future development perspective and customer service perspective respectively.

We can see from this table, the indexes system has total of only 16 third-level indicators. It is significantly less than the 45 third-level indicators of SCPR1.0 system, and the 16 indicators is easy to obtain from the actual operations in the supply chain, so

Table 1. Performance evaluation system of supply chain based on balanced scorecard

| | | |
|---|--------------------------------|---|
| Performance evaluation system of supply chain | Financial value perspective | return on capital of supply chain |
| | | Inventory day of supply chain |
| | | Cash turnover ratio |
| | | Sales growth |
| | Operation flow perspective | Effective advance time ratio |
| | | Time flexibility |
| | | achieve Target cost ratio of supply chain |
| | | sales ratio of new product |
| | Future development perspective | holding cost of supply chain |
| | | final main decoration of products |
| | | Information sharing ratio |
| | Customer service perspective | Team participation |
| | | Order circle |
| Customer response time recognition | | |
| customer value ratio | | |
| | customer retention ratio | |

the indicator system has simplicity and simple operation. Furthermore, because of the characteristics of the balanced scorecard, the indicator system expresses not only the performance of the strategic level, but the performance of practical operation level.

Beamon (1999) [8] established a performance evaluation system from the strategic perspective of the supply chain. It built 18 three-level indicators from the resources, output and flexible aspects. The indicators system has the advantage of focusing on strategic objectives. But the drawback is that it only considers the strategic goals but ignores short-term goals, and its third-level indicators are also more than indicators of balanced scorecard-based system.

In summary, this study focused on feasibility and simplicity, so the selection is based on balanced scorecard indicators system.

3 Performance Evaluation Method of Synchronized Supply Chain—Entropy Weight and TOPSIS

To determine weight—Entropy weight. Evaluation algorithm includes two steps, and the first step is to determine the weight. Only the weight is accurately determined, the results of evaluation is convincing. The most common methods for determining the weights are the AHP and pairwise comparison. The basic ideas of the two methods are all to invite experts to compare pairwise between indicators, and then the results of comparison are transferred into weights. The participation of experts makes the two methods have strong subjectivity; the indicators of pairwise comparison can not be too many, generally no more than 9. The Entropy method is objective, and can determine weights on any number of indicators, so it is adopted in this paper.

The concept of entropy comes from thermodynamics. It is a measure of the uncertainty about system state [9]; and information is a measure of system order degree in information theory, their absolute value are both equal, but sign is opposite. The basic idea which Entropy method determines weights is based on the above nature.

Entropy is defined as follows, let P_j denote j -th information uncertainty degree (that is the emergent probability), then the uncertainty degree of whole information (suppose that there are n items), is entropy. It can be presented by formula (1):

$$S = -K \sum_{j=1}^n p_j \ln p_j \tag{1}$$

Where K is positive constant, extremely, various information is equal to the probability, $P_j=1/n$, then the entropy S is maximized, while the weight of indicator is the smallest.

The basic process of Entropy method determining the weight is that, first, it determines the entropy sum on each attribute for all programs, then determines the weight according to entropy sum. If entropy sum is large to this attribute, it means that for all programs has small changes in this attribute, then this attribute has small contribution to the final evaluation results. Else, the result is opposite.

The specific operations are as follows. *First*, we assumed decision matrix is:

$$M = \begin{matrix} & \begin{matrix} A_1 & A_2 & \dots & \dots & A_m \end{matrix} \\ \begin{matrix} A_1 \\ A_2 \\ \dots \\ \dots \\ A_m \end{matrix} & \begin{bmatrix} x_{11} & x_{12} & \dots & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & \dots & x_{2n} \\ \dots & \dots & & & \\ & \dots & \dots & & \\ & & \dots & \dots & \\ x_{m1} & x_{m2} & \dots & \dots & x_{mn} \end{bmatrix} \end{matrix}$$

Let formula (2) denote contribution degree of j -th attribute X_j , i -th program A_i , then the total contribution of all programs (total is m programs) on the properties X_j can be expressed as:

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} \tag{2}$$

From the view of entropy, p_{ij} denotes information uncertainty degree of j -th attribute X_j -th program A_i , the sum of information uncertainty degree of j -th attribute X_j is 1.

$$E_j = -K \sum_{i=1}^m p_{ij} \ln p_{ij} \tag{3}$$

where, the constant K can be taken as: $K = 1 / \ln m$, then $0 \leq E_j \leq 1$, i.e. maximum of E_j is 1.

According to the basic idea of entropy method, the attribute weight depends on differences of all major programs on this attribute. So define d_j as consistency of each program's contribution degree on j -th attributes.

$$d_j = 1 - E_j \tag{4}$$

Definition of d_j is based on two considerations, first is the entropy $E_j \leq 1$; second is that the greater entropy E_j is, the smaller the program changes are, and similarly, the



more smaller d_j is according to the formula (4). In other words, d_j is a yardstick which measures the change of each program on attribute indicator.

Then, the weight of each attribute w_j can be denoted:

$$w_j = d_j / \sum_{j=1}^n d_j \tag{5}$$

This step is the normalization process, the purpose is to let the weight of each attribute be comparable, and the sum of all the attribute weights is 1.

The evaluation method of Samples—TOPSIS. TOPSIS is a program comparison algorithm of high objectivity. TOPSIS and Entropy methods are all objective measurement for the whole solution set, and aren't dependent on the characteristics of the problem itself, so the entropy method and TOPSIS are often used in conjunction. TOPSIS has good precision and simple operability when solutions are sorted and compared. The goal of this paper is to put the performance of many synchronized supply chain solutions in order, rather than a solution. Based on the above two points, this paper selects TOPSIS.

The basic idea of TOPSIS (close to the ideal solution for sorting methods) is as follows: first to find the solution set of the ideal solution and negative ideal solution, then to determine the Euclidean distance from each program to the ideal solution and negative ideal solution respectively, and finally to put all the programs in order according to Euclidean distance [9].

Assumed that ideal solution is denoted as A^* and negative ideal is A^- , specific algorithm process of TOPSIS is as follows:

Step one: construct normalized decision matrix R

$$R = [r_{ij}], \quad (r_{ij} = x_{ij} / \sqrt{\sum_{i=1}^m x_{ij}^2}) \tag{6}$$

Where i denotes i -th program, j represents j -th attribute. The purpose of this step is to let attributes of a program be comparable.

Step Two: Construct weighted normalized matrix $V = [v_{ij}]$

$$V = R \cdot W = \begin{bmatrix} w_1 r_{11} & w_2 r_{12} & \cdots & w_n r_{1n} \\ w_1 r_{21} & w_2 r_{22} & \cdots & w_n r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_1 r_{m1} & w_2 r_{m2} & \cdots & w_n r_{mn} \end{bmatrix} \tag{7}$$

Where w_i represents weights determined by Entropy method, r_{ij} represents i -th program j -th attribute after step one. The purpose of this step is to add the weights of attributes based on original normalized decision matrix.

Step Three: Determine the ideal solution and negative ideal solution

When the attribute value is efficiency, the ideal solution is the max value in each column, the negative ideal solution is the min value in each column; when the attribute

value is loss, the ideal solution is the min value in each column, and the negative ideal solution is max value in each column. The idea is specifically expressed as follows:

$$A^* = [(\max_i v_{ij} \mid j \in J), (\min_i v_{ij} \mid j \in J')] = [v_1^*, v_2^*, \dots, v_j^* \dots v_n^*]$$

$$A^- = [(\min_i v_{ij} \mid j \in J), (\max_i v_{ij} \mid j \in J')] = [v_1^-, v_2^-, \dots, v_j^- \dots v_n^-] \tag{8}$$

$i=1, 2, 3 \dots m$

Where: $J = (j = 1, \dots, n \mid j \text{ is the efficiency target attribute}); J' = (j = 1, \dots, n \mid j \text{ is the loss target attribute}).$

Obviously, the ideal solution is consisted by the optimal value of each attribute; the negative ideal solution is consisted by the worst value of each attribute.

Step Four: calculate the distance

Including distance of the ideal solution S_i^* and distance of the negative ideal solution S_i^- , formula is as follows:

$$S_i^* = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^*)^2}$$

$$S_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2}$$

(9)

Where v_{ij} denotes the value of i -th program, j -th attribute after Step Two, v_j^* express the value of j -th attribute in ideal solution, v_j^- express the value of j -th attribute in negative ideal solution. Obviously S_i^* is Euclidean distance between the i -th program and the ideal solution, but S_i^- is Euclidean distance between the i -th program and the negative ideal solution.

Step five: Calculate the relative proximity degree

After having obtained Euclidean distance, you can calculate the relative proximity degree C_i , which is showed as follows:

$$C_i = \frac{S_i^-}{(S_i^- + S_i^*)}, \text{ where } i = 1, \dots, m \quad 0 \leq C_i \leq 1 \tag{10}$$

In extreme cases, when $S_i^- = 0, C_i = 0, A_i = A_j^-$, i.e. when i -th program A_i is the negative ideal solution A^- , the proximity C_i is 0; when $S_i^* = 0, C_i = 1, A_i = A_j^*$, that is, when i -th program A_i is the ideal solution A^* , the proximity C_i is 1. C_i clearly expresses the program advantages and disadvantages. At last, according to C_i , programs are sorted.



4 Conclusion

Synchronized supply chain is a new hotspot of supply chain. This paper uses Entropy Weight and TOPSIS which is an objective and easy operative way to evaluate its performance, and provides a suitable idea for practical application, which has the following two contributions:

First, an indicator system of synchronized supply chain performance evaluation is established and its advantages compared with other indicators system are presented.

Second, the entropy and TOPSIS is used to evaluate performance of synchronized supply chain.

References

1. Peter, K.: Purchasing must become supply management. *Harvard Business Review* (5), 109–117 (1983)
2. Anderson, D., Lee, H.: Synchronized Supply Chain: The New Frontier. *ASCET* 6(1) (1996)
3. Shi, C., Chen, J.: The performance prediction of supply chain collaborative management based on heuristic attribute reduction and neural network. *Science and Technology Management Research* 29(3), 283–286 (2009)
4. Hong, W., Liu, J.: A Survey on Supply Chain Performance Measurement. *Commercial Research* (8), 18–21 (2006)
5. Kaplan, R.S., Norton, D.P.: The Balanced Scorecard—Measures That Drive Performance. *Harvard Business Review* (1), 71–79 (1992)
6. Ma, S., Li, H., Lin, Y.: An Application Research of Supply Chain Performance Measurement Using Balanced Scorecard. *Industrial Engineering and Management* (4), 5–9 (2002)
7. Ye, C., Ma, H., Li, D., Liu, Y.: Application Research of Backpropagation Neural Network in Supply Chain Management Performance Index. *Industrial Engineering and Management* (5), 35–37 (2005)
8. Beamon, B.M.: Measuring Supply Chain Performance. *International Journal of Operations & Production Management* 19(3/4), 275–292 (1999)
9. Yang, B., Zhang, K.: Multi-objective decision analysis theory, method and application. Donghua University Press, Shang Hai

An Evaluation Model for Synchronized Supply Chain Based on Entropy Weight and TOPSIS II

Li Zhiyong, Wang Xiang, and Zheng Nan

Zhengzhou Institute of Aeronautical Industry Management,
Zhengzhou 450015, China
leesilee@126.com

Abstract. A detail comparison of many relevant indexes is made towards the performance evaluation of synchronized supply chain. Combining to the design ideal which supply chain need to considerate short-term goals and strategic goals, the indexes system based on balanced scorecard is picked up. Because of the characteristics of synchronized supply chain, Entropy weight and TOPSIS is selected to implement the performance evaluation, The performance evaluation model based on Entropy weight and TOPSIS is applied to a case. Case shows the detail of this method, and its advantages are convincing by comparing this method with other methods.

Keywords: Supply Chain; Synchronized, Balanced Scorecard, Entropy Weight, TOPSIS.

1 Introduction

We have given An Evaluation Model for Synchronized Supply Chain Based on Entropy Weight and TOPSIS. In this model, Entropy weight and TOPSIS which is an easy and feasible method are selected to evaluate the performance of synchronized supply chain. Its basic process is first to determine the weight by using entropy weight, then to evaluate the sample by using TOPSIS which is a evaluation method of samples. The basic operation idea of TOPSIS is to find the solution set of the ideal solution and negative ideal solution, then to determine the Euclidean distance from each program to the ideal solution and negative ideal solution respectively, and finally to put all the programs in order according to Euclidean distance. After having obtained Euclidean distance, we can calculate the relative proximity degree C_i , C_i clearly expresses the program advantages and disadvantages. According to C_i , programs are sorted.

In this paper, we apply the above evaluation algorithm for synchronized supply chain based on entropy weight and TOPSIS to implement a case.

2 Case Study

In order to compare different methods, this paper selects the basic data of supply chain performance evaluation from certain power equipment limited company in Shanghai [7].

Table 1. The basic data of supply chain performance evaluation

| sample | Indicator attribute values | | | | | | | | | | | | | | | |
|--------|----------------------------|-------|-------|-------|-------|-------|-----|----|-------|-----|-------|-------|-------|-------|-----|-----|
| | a1 | a2 | a3 | a4 | a5 | a6 | a7 | a8 | a9 | a10 | a11 | a12 | a13 | a14 | a15 | a16 |
| X1 | 0 | 0.853 | 0.067 | 0 | 0.625 | 1 | 0.5 | 1 | 0.535 | 1 | 0.511 | 0.609 | 0.969 | 1 | 1 | 1 |
| X2 | 0.96 | 1 | 0.071 | 0.849 | 0 | 0.259 | 0.5 | 1 | 0.359 | 1 | 0.561 | 0.7 | 0.7 | 0.14 | 1 | 1 |
| X3 | 0.94 | 0.853 | 0 | 0.689 | 0.25 | 0.331 | 1 | 1 | 0.529 | 1 | 0.538 | 0.698 | 0.735 | 0.093 | 1 | 1 |
| X4 | 0.967 | 0.603 | 0 | 0.501 | 0.75 | 0.389 | 0.5 | 1 | 1 | 1 | 0.553 | 0.803 | 0.78 | 0.083 | 1 | 1 |
| X5 | 0.925 | 0.873 | 0.175 | 0.613 | 0.25 | 0.35 | 1 | 1 | 0.41 | 1 | 0.475 | 0.683 | 0.703 | 0.096 | 1 | 1 |
| X6 | 0.846 | 0.631 | 0.18 | 0.709 | 0.5 | 0.901 | 0.5 | 1 | 0.771 | 0 | 1 | 1 | 0.43 | 0.205 | 1 | 1 |
| X7 | 0.654 | 0.539 | 0.065 | 0.891 | 0.375 | 0.528 | 0.5 | 1 | 0.32 | 1 | 0.527 | 0.413 | 1 | 0.07 | 1 | 1 |
| X8 | 0.991 | 0.561 | 0.481 | 0.981 | 0.375 | 0.711 | 0.5 | 1 | 0.309 | 1 | 0.611 | 0.221 | 0.325 | 0.021 | 1 | 1 |
| X9 | 0.983 | 0.763 | 0.105 | 0.815 | 0.375 | 0.115 | 0 | 1 | 0.051 | 0 | 0.244 | 0.223 | 0.261 | 0.005 | 1 | 1 |
| X10 | 0.99 | 0.735 | 1 | 0.97 | 0 | 0.093 | 0.5 | 1 | 0.048 | 1 | 0 | 0 | 0.329 | 0 | 1 | 1 |
| X11 | 0.981 | 0.907 | 0.413 | 1 | 0.25 | 0 | 0.5 | 1 | 0.191 | 0 | 0.248 | 0.223 | 0.491 | 0.021 | 1 | 1 |
| X12 | 1 | 0.781 | 0.071 | 0.963 | 0.191 | 0.151 | 0.5 | 1 | 0.191 | 0 | 0.239 | 0.227 | 0.485 | 0.003 | 1 | 1 |
| X13 | 0.991 | 0.093 | 0.071 | 0.963 | 0.311 | 0 | 0 | 1 | 0 | 1 | 0.211 | 0.225 | 0 | 0.002 | 1 | 1 |
| X14 | 0.919 | 0.939 | 0.344 | 0.813 | 0.88 | 0.061 | 0 | 1 | 1 | 1 | 0 | 0 | 0.095 | 0.003 | 1 | 1 |
| X15 | 0.995 | 0.41 | 0.995 | 0.961 | 0.061 | 0.109 | 0.5 | 1 | 0.411 | 0 | 0.247 | 0.301 | 0.271 | 0.003 | 1 | 1 |
| X16 | 0.997 | 0 | 0.997 | 0.971 | 0.25 | 0.217 | 0 | 1 | 0.221 | 1 | 0 | 0.225 | 0.327 | 0.003 | 1 | 1 |
| X17 | 0.993 | 0.411 | 0.993 | 0.957 | 0.25 | 0.061 | 0 | 1 | 1 | 1 | 0 | 0.22 | 0.325 | 0.005 | 1 | 1 |
| X18 | 1 | 0.339 | 1 | 0.948 | 0.269 | 0.07 | 0 | 1 | 1 | 1 | 0 | 0.211 | 0.331 | 0.004 | 1 | 1 |

As Table 1 shows, there are total of 18 samples, 16 indicators, the indicators correspond to indicator system of the second part. The indicator data has been standardized in [0, 1] interval, and transferred into the efficiency data (the bigger is the better).

Specific evaluation step for performance of each sample indicators is showed as follows:

Step 1: determine the weight of performance indicators for the supply chain by using entropy method.

I . Use of Table 1 to determine the decision-making program evaluation matrix M (because the sample is too large, indicators are too many, the data can not be displayed in a page, if the process of calculation is needed, you can contact with the author)

$$M = \begin{bmatrix} 0 & 0.8530 & 0.0670 & \dots & 1.0000 & 1.0000 \\ 0.9600 & 1.0000 & 0.0710 & \dots & 1.0000 & 1.0000 \\ 0.9400 & 0.8530 & 0 & \dots & 1.0000 & 1.0000 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0.9970 & 0 & 0.9970 & \dots & 1.0000 & 1.0000 \\ 0.9930 & 0.4110 & 0.9930 & \dots & 1.0000 & 1.0000 \\ 1.0000 & 0.3390 & 1.0000 & \dots & 1.0000 & 1.0000 \end{bmatrix}$$

Entropy method requires that the data should not include 0, but the data is normalized to [0,1] in this paper, and there are some data that is 0, so the original data



must be transformed before using it. The essence of entropy method is to use data difference to determine the weight, so we use the way which the original data maps from [0,1] to [1,2], it is not to change the size of differences within indicator data, but also makes use of entropy method.

Specific form of implementation is as follows:

$$M1 = M + ones(18,16) = \begin{bmatrix} 1.0000 & 1.8530 & 1.0670 & \dots & 2.0000 & 2.0000 \\ 1.9600 & 2.0000 & 1.0710 & \dots & 2.0000 & 2.0000 \\ 1.9400 & 1.8530 & 1.0000 & \dots & 2.0000 & 2.0000 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1.9970 & 1.0000 & 1.9970 & \dots & 2.0000 & 2.0000 \\ 1.9930 & 1.4110 & 1.9930 & \dots & 2.0000 & 2.0000 \\ 2.0000 & 1.3390 & 2.0000 & \dots & 2.0000 & 2.0000 \end{bmatrix}$$

II. According to the formula (2) in literature [1], matrix Pij which is attribute contribution of each program can be calculated by using matrix M1, and Pij is 18 × 16 matrix because of a total of 18 evaluation program, 16 properties. Matrix as follows:

$$P_{ij} = \begin{bmatrix} 0.029298 & 0.063262 & 0.042632 & \dots & 0.055556 & 0.055556 \\ 0.057424 & 0.06828 & 0.042792 & \dots & 0.055556 & 0.055556 \\ 0.056838 & 0.063262 & 0.039955 & \dots & 0.055556 & 0.055556 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0.058508 & 0.03414 & 0.079791 & \dots & 0.055556 & 0.055556 \\ 0.058391 & 0.048172 & 0.079631 & \dots & 0.055556 & 0.055556 \\ 0.058596 & 0.045714 & 0.079911 & \dots & 0.055556 & 0.055556 \end{bmatrix}$$

III. Calculate the entropy (Ej) of programs on attribute, the consistency level (dj) of contribution, the weight (wj) by using the formula (3), (4), (5) in literature [1], the result is shown in Table 2.

Table 2. Ej and dj and wj of attribute

| attribute | E _j | d _j | w _j | attribute | E _j | d _j | w _j |
|-----------|----------------|----------------|----------------|-----------|----------------|----------------|----------------|
| a1 | 0.99696 | 0.003045 | 0.031763 | a9 | 0.99086 | 0.009136 | 0.095311 |
| a2 | 0.99462 | 0.005379 | 0.05612 | a10 | 0.98695 | 0.013055 | 0.1362 |
| a3 | 0.98651 | 0.013485 | 0.14069 | a11 | 0.99278 | 0.00722 | 0.075323 |
| a4 | 0.99654 | 0.003461 | 0.036102 | a12 | 0.99318 | 0.006821 | 0.071163 |
| a5 | 0.99503 | 0.004973 | 0.051876 | a13 | 0.99409 | 0.005907 | 0.06163 |
| a6 | 0.99159 | 0.008407 | 0.087709 | a14 | 0.99392 | 0.00608 | 0.063429 |
| a7 | 0.99112 | 0.008884 | 0.092686 | a15 | 1 | 0 | 0 |
| a8 | 1 | 0 | 0 | a16 | 1 | 0 | 0 |

Step Two: To determine program priorities level by using TOPSIS

I . To construct normalized decision matrix R



$$R = M = \begin{bmatrix} 0 & 0.8530 & 0.0670 & \dots & 1.0000 & 1.0000 \\ 0.9600 & 1.0000 & 0.0710 & \dots & 1.0000 & 1.0000 \\ 0.9400 & 0.8530 & 0 & \dots & 1.0000 & 1.0000 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0.9970 & 0 & 0.9970 & \dots & 1.0000 & 1.0000 \\ 0.9930 & 0.4110 & 0.9930 & \dots & 1.0000 & 1.0000 \\ 1.0000 & 0.3390 & 1.0000 & \dots & 1.0000 & 1.0000 \end{bmatrix}$$

The original data M has been standardized, so R is equal to M.

II. To construct the weighted normalized matrix

$$V = \begin{bmatrix} 0 & 0.04787 & 0.0094261 & \dots & 0 & 0 \\ 0.030493 & 0.05612 & 0.0099888 & \dots & 0 & 0 \\ 0.029857 & 0.04787 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0.031668 & 0 & 0.14027 & \dots & 0 & 0 \\ 0.031541 & 0.023065 & 0.1397 & \dots & 0 & 0 \\ 0.031763 & 0.019025 & 0.14069 & \dots & 0 & 0 \end{bmatrix}$$

III. To determine the ideal solution and negative ideal solution

The risk evaluation of knowledge management is that the lower risk is the better, so attribute value is loss type. In this case, the ideal solution is the min value in each column; the negative ideal solution is the max value in each column. It can be expressed sepecifically as follows:

$$A^* = [0.031763 \quad 0.05612 \quad 0.14069 \quad 0.036102 \quad 0.045651 \quad 0.087709 \quad 0.092686 \quad 0.095311 \quad 0.1362 \quad 0.075323 \quad 0.071163 \quad 0.06163 \quad 0.063429 \quad 0 \quad 0]$$

$$A^- = [0 \quad 0 \quad 0]$$

IV. To determine the distance between each program and the ideal solution or negative ideal solution

According to the formula (8) in literature [1], to calculate the distance between each program and the ideal solution or negative ideal solution, including of Si* and Si-, the results are as Table 3.

Table 3. S* and S-

| Sample | S* | S- | Sample | S* | S- |
|--------|---------|---------|--------|---------|---------|
| X1 | 0.16134 | 0.21301 | X10 | 0.18838 | 0.21191 |
| X2 | 0.18485 | 0.1836 | X11 | 0.23099 | 0.11139 |
| X3 | 0.17823 | 0.20172 | X12 | 0.24904 | 0.09203 |
| X4 | 0.17486 | 0.20546 | X13 | 0.24555 | 0.14713 |
| X5 | 0.16401 | 0.19924 | X14 | 0.20446 | 0.19135 |

Table 3. (Continued)

| | | | | | |
|----|---------|---------|-----|---------|---------|
| X6 | 0.19821 | 0.17131 | X15 | 0.21112 | 0.16491 |
| X7 | 0.18262 | 0.18131 | X16 | 0.19358 | 0.20517 |
| X8 | 0.15255 | 0.19083 | X17 | 0.1789 | 0.2252 |
| X9 | 0.26546 | 0.07242 | X18 | 0.17929 | 0.22549 |

V. To calculate the relative proximity degree

Then, according to formula (9) in literature [1] to calculate the relative proximity C_i , the results are as Table 4.

Table 4. The relative proximity of the solution

| | | | | | |
|-------|--------------------|-------|--------------------|-------|--------------------|
| C_i | relative proximity | C_i | relative proximity | C_i | relative proximity |
| C1 | 0.56901 | C7 | 0.4982 | C13 | 0.37468 |
| C2 | 0.49831 | C8 | 0.55574 | C14 | 0.48344 |
| C3 | 0.53091 | C9 | 0.21434 | C15 | 0.43856 |
| C4 | 0.54023 | C10 | 0.52939 | C16 | 0.51453 |
| C5 | 0.54849 | C11 | 0.32533 | C17 | 0.55728 |
| C6 | 0.4636 | C12 | 0.26983 | C18 | 0.55706 |

VI. According to the size of relative proximity C_i , programs are sorted, the ranking results are as Table 5.

Table 5. The sorting results of synchronized supply chain performance evaluation

| | | | | | |
|----------|------|----------|------|----------|------|
| solution | sort | solution | sort | solution | sort |
| X1 | 1 | X7 | 11 | X13 | 15 |
| X2 | 10 | X8 | 4 | X14 | 12 |
| X3 | 7 | X9 | 18 | X15 | 14 |
| X4 | 6 | X10 | 8 | X16 | 9 |
| X5 | 5 | X11 | 16 | X17 | 2 |
| X6 | 13 | X12 | 17 | X18 | 3 |

From Table 5, we can see that synchronized supply chain performance of the sample X1 is best. It is the selected solution, and synchronized supply chain performance of the sample X9 is the worst.

This evaluation results are basically the same with those of literature [7]. But compared with BP neural network method in literature [3] and rough set and BP neural network in literature [6], entropy+TOPSIS method has the advantages of wide applicability, simple operation and no need of the priori knowledge; but BP neural Network need a training sample set.

So, entropy and TOPSIS is more effective and simple for the evaluation of synchronized performance of supply chain.

Entropy and TOPSIS also have some limitations. Firstly, this method does not consider the evaluation problem's characteristics of itself, but purely starts from the data to look for sorting results. For example, entropy method determines the weight only from changes intensity of the attribute data, without considering the importance of

the attribute itself; Secondly, this method only has the sort capability, but can not give a mark to a sample. So, this method can only solve the issue which a competitive solution is selected from multiple programs; Thirdly, when determining the weight using entropy method, we try to use the original data. Because of limitation of data acquisition, we use the normalization data to calculate weights, the final sorting results is not entirely consistent with the original literature.

3 Conclusion

Synchronized supply chain is a new hotspot of supply chain. This paper evaluates its performance in an objective and easy operative way, and provides an idea for practical application, which has the following two contributions:

First, an indicator system of synchronized supply chain performance evaluation is established and its advantages compared with other indicators system are presented.

Second, the entropy and TOPSIS is used to evaluate performance of synchronized supply chain and a specific case is implemented, and the advantages and disadvantages compared with other methods are pointed out in detail.

References

1. Li, Z., Wang, X., Zheng, N.: An Evaluation Model for Synchronized Supply Chain Based on Entropy Weight and TOPSIS I. Accepted by MSCS 2011 (2011)
2. David Anderson, Hau Lee. Synchronized Supply Chain: The New Frontier. ASCET 1996,6(1)
3. Shi, C., Chen, J.: The performance prediction of supply chain collaborative management based on heuristic attribute reduction and neural network. *Science and Technology Management Research* 29(3), 283–286 (2009)
4. Hong, W., Liu, J.: A Survey on Supply Chain Performance Measurement. *Commercial Research* (8), 18–21 (2006)
5. Kaplan, R.S., Norton, D.P.: The Balanced Scorecard-Measures That Drive Performance. *Harvard Business Review* (1), 71–79 (1992)
6. Ma, S., Li, H., Lin, Y.: An Application Research of Supply Chain Performance Measurement Using Balanced Scorecard. *Industrial Engineering and Management* (4), 5–9 (2002)
7. Ye, C., Ma, H., Li, D., Liu, Y.: Application Research of Backpropagation Neural Network in Supply Chain Management Performance Index. *Industrial Engineering and Management* (5), 35–37 (2005)
8. Beamon, B.M.: Measuring Supply Chain Performance. *International Journal of Operations & Production Management* 19(3/4), 275–292 (1999)
9. Yang, B., Zhang, K.: *Multi-objective Decision Analysis Theory, Method and Application*. Donghua University Press, Shang Hai

Design of an Online Parameters Configuration Based on C8051F USB Processors Family

Zhu Lei^{1,2}, Zhi Wenxia¹, and Zhang Gaoqiang¹

¹ Electronics and Information College, Xi'an Polytechnic University, Xi'an 710048, China

² National Lab of Radar Signal Processing, Xidian University, Xi'an, 710071, China
zhu1987802@sina.com

Abstract. The embedded systems with the online parameter configuration not only facilitate the debugging and the upgrading of the system, but also improve the adaptability of the system for different applications, so it has become an essential feature of the many embedded application systems. For the C8051F USB processors series as the core of the embedded system, a new program of online parameters configuration has been proposed. The program will interconnect the host computer and the embedded system through the USB bus, use the LabVIEW online configuration GUI software of PC to complete the online updates for the system parameters in the processor on-chip FLASH memory, and just change the parameter settings without impacting the original program of the processor. The online parameters configuration embedded system designed using the program has four characteristics: can switch automatically in the two working states which are normal work mode and online configuration mode without external intervention; can facilitate the realization of the online parameters configuration through the USB bus and the host computer GUI software, take full advantage of two types of hardware resources: one is the processor coming with on-chip USB interface, the other is flash memory, without external other expansion devices, simple structure and low cost, both offline parameter configuration, and online parameters can also be configured, with more flexible working ways and adaptable. The Ethernet communication system designed base on the program with the function of online parameters configuration has been successfully applied to the armed police information systems.

Keywords: online configuration, C8051F, USB, FLASH, GUI, LabVIEW.

1 Introduction

In practice, people always expect to buy the embedded applications which can adapt to different environments and application requirements, at the same time it's easily to upgrade and update. On the other hand, embedded application systems developers hope the designed systems have good scalability, and can facilitate the completion of the system parameter changes and feature upgrades in the case of without changing the original hardware of system, to bring the convenience for the later maintenance and upgrade of system. To achieve the above two requirements, a more effective solution is

to allow embedded application system with the function of online parameters configuration, so the system can be convenient online parameters configuration, which can be configured different system parameters according to different user needs, while also facilitate the update and the upgrades as well as the maintenance of the system.

Currently, there are several more common online parameters configuration of embedded system. One option is to expand external the power-down non-volatile memory for the embedded system, like the EEPROM, FLASH and other, and store the system parameters in the memory, if need modify the system parameters, you first need to connect the embedded system with the programmer, then program parameter data in the memory, and ultimately complete the parameters configuration. Lack of this program is that it needs to expand the external memory, and often the operations that writing data to this memory are not convenient. Another solution is in embedded systems allow in advance reserve programming interface (for example, JTAG programming debug interface commonly used in DSP and C8051F processor) for the programmable processor, when need parameters configuration, directly re-download/program the new procedures code to complete the parameters configuration. The disadvantage of this approach is need to re-program all of the processor code to update the parameters, also need reserve programming interface in advance in embedded systems, these interfaces are often not the common interfaces.

For the drawbacks of the two online parameters configuration in embedded system, this paper proposes a new online parameters configuration using C8051F USB series processors as the core of the embedded system. The program doesn't need external expansion memory, through the USB bus interconnecting the host computer with the processor with a USB interface, with the LabVIEW parameter configuration GUI software in the PC, to realize the re-configuration for the parameters stored in on-chip FLASH of C8051F processor system, and only change the parameters without any impact on the original program of the processors. In addition, the embedded system of the online parameters configuration designed by the program has many fine attributes: can switch automatically in the two working states which are normal work mode and online configuration mode without external intervention, with the smart property; can facilitate the realization of the online parameters configuration through the USB bus and the host computer GUI software, with the convenient connections and simple operation attributes; take full advantage of two types of hardware resources: one is the processor coming with on-chip USB interface ,the other is flash memory, without external other expansion devices, with the simple structure and low cost attributes; both offline parameter configuration, and online parameters can also be configured, with a more flexible working way .

The following will in detail the design ideas, hardware design and software design of the program, in addition, when introduce the program, the Ethernet communication module with online parameters configuration designed based on the text-based program as an example to illustrate.

2 Design Ideas of Online Parameters Configuration

The paper proposes the program of online parameters configuration which is aimed to C8051F USB series processors as the core of the embedded system, and take full

advantage of the inherent on-chip peripheral hardware resources of the processor, both the FLASH and USB. Meanwhile, C8051F processor of the embedded system has two modes, namely: the normal operating mode (run user program) and the online parameter configuration mode (operating parameter configuration program), and C8051F processor automatically select the processor mode according to whether change of the USB bus. For the embedded applications system with the function of online parameters configuration designed based on the program, its steps of the online parameters configuration are as follows: First, interconnect C8051F processor of the embedded system and PC through the USB bus; Second, complete the new parameter settings of the embedded systems in the LabVIEW parameters configuration GUI software of PC-side, and send the parameters setted to the C8051F processor through the USB bus; Third, the C8051F processor in the online parameter configuration mode immediately write the new parameters in the specified address area of the on-chip FLASH of the processor after receiving the system parameter data send by the USB interface of PC, then start processor software reset after the completion of parameter storage, thus completing the online parameters configuration and making the processor into the processor's normal Mode after updating system parameters.

2.1 Introduction

The program diagram of the online parameters configuration shown in Fig.1. View From the point of hardware, the program doesn't need expand other devices, only need design USB interface circuit for C8051F processor of the embedded system. As C8051F USB processors family with on-chip USB interface, the design of USB interface circuit is simple and convenient. View from the point of software, the software of the program, including online parameters configuration program of C8051F processor and LabVIEW parameters configuration GUI program of PC. When the embedded system interconnect to PC through USB bus, C8051F processor automatically enter in the online parameters configuration mode, now, both C8051F processor online parameters configuration application program and LabVIEW parameters configuration GUI program on PC need their own the USB driver program, to achieve communication between C8051F processor and USB of PC. In the preparation of the two USB driver program, the paper gets help from USB driver library functions (USBXpress DLL) and USB application program interface (USBXpress API) supported by the Silicon company, the specific details refer to references [1]. In addition, the function that the embedded system automatically selects the operating mode based on the actual situation is implemented by the joint of the USB interrupt service program of the C8051F processor and the software reset program.

2.2 Hardware Design of Program

Based on the proposed program, expand the embedded system which can't online parameter configuration to become the system with this feature, the hardware design only needs set interface circuit for on-chip USB peripheral in the C8051 processor of embedded systems, it doesn't need the other devices of external expansion, thus simplifying the hardware structure, improving integration, reducing costs.

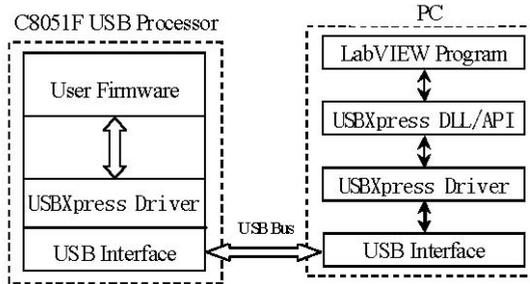


Fig. 1. Program diagram of online parameter configuration

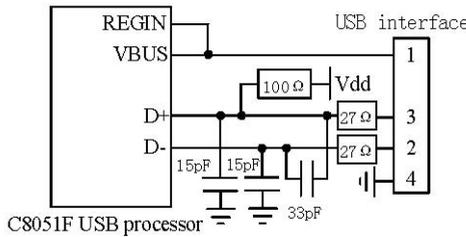


Fig. 2. Hardware schematic diagram

Expanding the USB interface circuit for embedded system shown in Fig.2. When interconnecting on-chip USB peripherals of C8051F processor and USB interface circuit, the differential data port D + and D- need use resistors and capacitors to limiting, filtering and the pull-up protection. Meanwhile, +5V power pin 1 of USB interface circuit need connect to the voltage regulator +5V input pin REGIN and USB VBUS voltage detection pin of C8051F processor. Thus, when the power of the embedded system is off, though C8051F processor can't require +3.3V power for its operation from the system, C8051F processor can still receive USB interface of PC voltage +5V through on-chip voltage regulator and convert it into +3.3V, for its own need to operation, thereby achieving the offline parameter configuration of the embedded systems. Furthermore, regardless of whether the power is supplied of the embedded system, once the system interconnects to USB interface of PC, C8051F processor can automatically detect the voltage change of the USB interface through VBUS pin, and automatically switches the module from never work or normal work mode to the online parameter configuration mode, to achieve the online parameters configuration of the embedded system.

2.3 Software Design of the Program

Based on the proposed program, expand the embedded system which can't online parameter configuration to become the system with this feature, it is needed to design two types programs in the software: first, need to design the LabVIEW of the PC-side parameters configuration GUI procedures; Second, need to the design the online parameters configuration procedures for the C8051F processor of the embedded

system, while the address allocation of system parameters in the user's original application procedures has certain requirements. Here, the specific designs of the two types of procedures of the program are separately introduced.

Design of LabVIEW Parameter Configuration GUI Procedures. The software design of LabVIEW of PC parameter configuration GUI mainly completes two major functions, one is to design the graphical human-machine interface; the other is to design the USB communication program based on LabVIEW, thus complete to send the system parameters entered by the user through the PC USB interface to the C8051F processor of the embedded system. The following will combine the Ethernet communication module with online parameters configuration to indicate the software design of LabVIEW of PC, parameter configuration GUI.

Figure 3 is the front panel of the GUI programs designed based on LabVIEW, mainly including the following input parameters controls: IP address of the source, IP address of the destination, subnet mask, IP address of gateway, the size of data buffer, number of socket, model Choice of server / customer service machine, in addition including the on button and indicator of the system, the send button of the Ethernet module parameters and the two send state output control.



Fig. 3. The front panel of Ethernet module parameter configuration based on LabVIEW

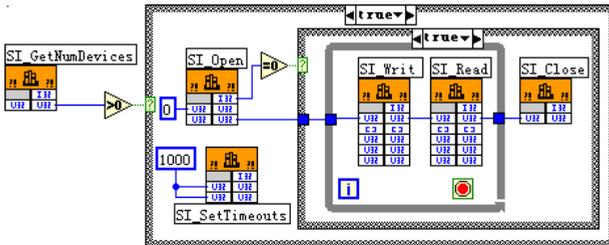


Fig. 4. The back panel of brief example of USB communication based on LabVIEW

Based on the LabVIEW front panel of the designed program, LabVIEW rear panel program need complete communication driver program of PC, as USB Master and the calls, connections and settings between the controls with the sub-VI. The LabVIEW rear panel program completely starts around USB communication process, the establish of USB communication use USBXpress software of the Silabs company providing 10 USB host API-driven functions (included in SiUSBXp.dll in the installation directory USBXpress), which SI_GetNumDevices(), SI_Open(), SI_Close(), SI_Read(),

SI_Write(), SI_SetTimeouts() function are necessary when created the USB communications in LabVIEW. Figure 4 is a brief example of LabVIEW rear panel USB communication program, the specific function, parameter configuration, and LabVIEW call method of USB host API function can refer to references [2, 3].

Program Design of Online Parameters Configuration of C8051F Processor.

Embedded systems has two working mode: the normal operating mode and online parameters configuration ,therefore, embedded systems C8051F processor program should also include the normal operating mode and online parameters configuration mode ,the two main branches of the program flow. The former is application program written by the user which is responsible for applications function of the embedded system, the latter is responsible for the parameter configuration of the embedded system, to complete online debugging and program upgrades of system. In addition, to complete the mutual automatic conversion of the program flow of the two branches, so we introduce soft reset of the processor and the USB VBUS level interrupt service routine in particular.

The main program flow of C8051F processor shown in Figure 5, C8051F need to complete the self-initialization after power-on, and then by checking the VBUS pin whether there is +5 V voltage provided by USB host, to determine which branch program flow to be entered. If there are +5 V voltages of VBUS pin detected, then the processor must enter the branch program flow of online parameters configuration, or need to enter branch program flow of the normal operating mode. When completed online parameters configuration of C8051F processor, using the C statement "RSTSRC = RSTSRC | 0x10;" by the fifth position 1 of the register RSTSRC start processor software reset, so the system parameters of the re-configuration can be successful Loaded after the processor reset. When C8051F processors work in normal operating mode, this time if users interconnect PC and embedded system through USB connection, the VBUS pin of C8051F processor conducts +5 V voltage of USB interface of PC, and then have led VBUS pin of C8051F processors level interrupt, and VBUS level interrupt service routine of C8051F processor also use statement "RSTSRC = RSTSRC | 0x10;" start processor soft reset, so the processor enter the online parameters configuration flow after reset.

The branch program flow of online parameters configuration of C8051F processor shown in Figure 6, the figure appears to multiple functions which are USB devices API-driven functions provided USBXpress (in the USBX_F34X.LIB under the installation directory USBXpress). The details of USB Device API function can refer to reference[1], there is one point need to point out, despite FLASH memory of the C8051F processor write byte operation, but must first erase the entire which the bytes in (512 bytes of a sector). Thus, in programming, all the parameters of online configuration of the system should target to a particular sector of the FLASH memory. If all parameters less than a sector, in order to prevent the compiler of the developed software (for example, Keil C51 software) allocate the rest bytes space of the sector to the other program code, so needs to use a code variable to locate under the remaining storage space. Thus, when write the system parameters data Of C8051F processor received from PC USB interface to the corresponding sector, can avoid damage of the FLASH write operation on the other original non-parametric program code of the FLASH.

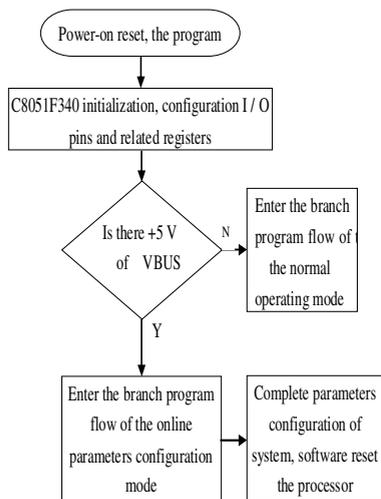


Fig. 5. Main program flow of C8051F processor para- processor

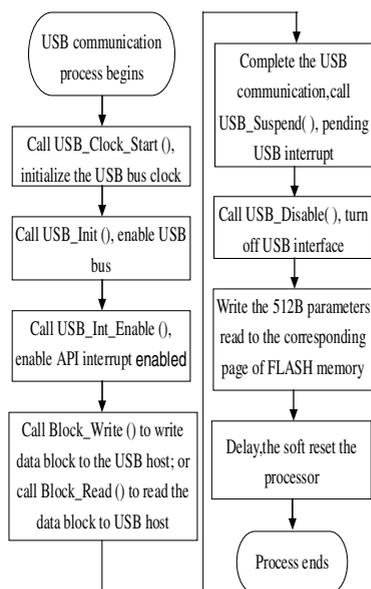


Fig. 6. Branch program flow of online meters configuration of C8051F

3 Conclusion

The paper proposes the program of online parameters configuration which is aimed to C8051F USB series processors as the core of the embedded system. The program doesn't need expand the external memory, that the USB bus and PC, LabVIEW parameter configuration GUI software can realize the online parameters configured of embedded system, and just change the parameter settings without any impact on the original program of processor. At the same time, the embedded system of the online parameters configuration designed by the program has many fine attributes: can switch automatically in the two working states which are normal work mode and online configuration mode without external intervention, with the smart property; can facilitate the realization of the online parameters configuration through the USB bus and the host computer GUI software, with the convenient connections and simple operation attributes; take full advantage of two types of hardware resources: one is the processor coming with on-chip USB interface, the other is flash memory, without external other expansion devices, with the simple structure and low cost attributes; both offline parameter configuration, and online parameters can also be configured, with a more flexible working way and adaptability. The Ethernet communication system designed base on the program with the function of online parameters configuration has been successfully applied to the armed police information systems.

Acknowledgement. This paper is supported by Indicative Planning Program of China National Textile and Apparel Council (2009089), Scientific Research Program of Shanxi Province Department of Education (09JK474), and The Basic Research Fund of Xi'an Polytechnic University (11JK0983).

References

1. USBXpress programmer's guide (Rev.1.8 3/07), <http://www.silabs.com>
2. Lei, Z., Dong, L.: The Realization of USB Communications Between C8051F340 And Labview Programs Based On API. *Microcontrollers & Embedded Systems* 12, 74–76 (2008)
3. Lei, Z.: The design of virtual signal generator with USB interface based on LabVIEW and DDS. *Science Technology and Engineering* 8, 1851–1855 (2008)

Study on Theory of Model Test in Model Credibility Evaluation

Hao Li, Shuo Tang, and Xiaodong Yan

College of Astronautics, Northwestern Polytechnical University,
Xi'an 710072, China
hao_sun@yeah.net

Abstract. In order to assure the credibility of simulation models used in virtual test, Model Test is one of most important techniques. In this paper, the theory of Model Test in Virtual Test is researched. Firstly, the definition of Test Case is given and techniques of producing test case are researched and presented. Secondly, Evaluation Index System of Model Test and classification of Evaluation Reference are introduced. Thirdly, the Process Model of Model Test is proposed. Finally, researches mentioned above have been applied into the model test of one simulation model of aerodynamic. The practical results showed that all the studies above have important directive to simulation model test in virtual test and effectively assure the credibility of simulation model.

Keywords: virtual test, credibility evaluation, model test, test case.

1 Introduction

Having high credibility is the prerequisite of virtual test's substituting for real test to provide right test, analysis and evaluation on product performance and achieving reducing costs and shortening develop cycle. The credibility of model is the key factor to assure the credibility of virtual test, for models of system and subsystem are the nucleus of virtual test. Any model without validation is worthless; any simulation without evaluation is worthless[1]. Therefore, the credibility of model is always the core of research on simulation credibility, from the beginning of attention on simulation credibility in 1960's to VV&A now.

As one of techniques of model verification and validation, model test comes from Software Test, and it is involved in VV&A[2] and T&E. Based on a large number of documents concerning this subject and requests on project, the paper researches on the theory of Model Test in Virtual Test which aims at assuring the credibility of model. As is known, Modeling and Simulation (M&S) follows the process 'three steps' of 'Physical Model - Concept Model - Computer Model', model test is mainly carried out along transformations among three models. The transformation between concept model and computer model isn't taken into account in this paper, in other words, the credibility of computer model is equivalent to that of physical model. Aim at building computer model with high credibility in virtual test, through analysis and evaluation of the similarity between computer model and physical model.

2 Test Case

Test Case is the abstract description of test purposes and input data. Test case is the programmatic file in test activities[3]. The quality of Test Case decides whether the goal of model test can be achieved and the completeness of model test. It is denoted by this vector:

$$\text{Test Case} : \{CC; \langle DV \rangle, \langle V \rangle, TT, ER\} \quad (1)$$

Where, *CC* (Credibility Character) is one aspect of model credibility characteristics, which is intended to be analyzed and evaluated. It offers evaluation aim for test; $\langle DV \rangle$ (Design Variable) is input variables or model parameters chosen to design Test Case. It include N_v variables; $\langle V \rangle$ is the value vector of $\langle DV \rangle$; *TT* is test case producing technique adopted to assign $\langle DV \rangle$ to get $\langle V \rangle$; *ER* (Evaluation Reference) is the most credible reference available used as evaluation reference.

Based on credibility characteristics of input data type of model, Test Case is categorized into Typical Test Case, Bound Test Case, Normal Test Case, Trend Test Case, and Robust Case.

Typical Test Case: Some clear result can be obtained with some initial settings or hypothesis for most models. This way, corresponding test case can be produced to evaluate specific credibility characteristic based on that initial settings or hypothesis. This case is usually used at early stage of credibility evaluation, and need a competent understanding of object model.

Bound Test Case: Most models usually show specific characteristic with some specific data as input, which is at the bound of input space. From this point, Bound Test Case can be produced to evaluate that specific credibility characteristic.

Normal Test Case: The values of design variables of this kind of case meet every input requirement of object model, in other words, the values of $\langle V \rangle$ is within input space of design variables. For the count of input variable combinations is too large to make a complete test, we need to adopt the techniques of design of experiment to produce test case to evaluate model credibility at limited times.

Trend Test Case: The result of single test shows statistic characteristic of object model. However, the trend characteristic of one parameter can offer a better credibility reflection for some models. This way, we can design test case with the value of one design variable changing at one trend.

Robust Test Case: The value of one or several design variables of this kind of test case is without its input space, containing incorrectness data sometimes. Robust test case is used in credibility evaluation through evaluating the robustness of object model.

3 Test Case Producing Techniques

How to assign $\langle DV \rangle$ rightly and effectively decides the quality of test case directly. We usually adopt techniques of design of experiment, including Random Assign, Control Variate Method, Full-Factor Design, and Uniform Design. Of course, we can

assign $\langle DV \rangle$ with pointed values directly, which are regarded. Here, we call this way Direct Assign.

(1) *Direct Assign Method*

Based on the credibility characteristic of object model and expert experience, we can assign $\langle DV \rangle$ directly with specific values which is regardful by evaluation worker. Direct assign is usually used when the credibility characteristic of object model under some specific input.

(2) *Random Assign Method*

As is known, we can't evaluate the credibility at every parameter combination within their input space $[V_{i\min}, V_{i\max}]$, for the amount of combinations is tremendous. Therefore, we can evaluate credibility of object model randomly but fully in limited times, if every variable of $[DV]$ is assigned randomly within its input space.

(3) *Control Variate Method*

In trend test, the tide of one characteristic parameter along with one design variable is tested. We adopt Control Variate Method to assign $\langle DV \rangle$ as follows:

- ① Pick DV_i within $\langle DV \rangle$ as trend variable;
- ② Choose test times N_t ;
- ③ The values of DV_i are assigned as follows:

$$v_{ij} = V_{i\min} + (V_{i\max} - V_{i\min}) \times j / N_t \tag{2}$$

Here, j is test serial number;

- ④ The values of $\langle DV \rangle$ are assigned as follows:

$$V_{N_t \times N_v} = \begin{pmatrix} v_1 & v_{i1} & v_{N_v} \\ v_1 & v_{ij} & v_{N_v} \\ v_1 & v_{iN_t} & v_{N_v} \end{pmatrix} \tag{3}$$

Trend test case is produced as 4 steps above.

(4) *Full-Factor Design*

It means all variables combinations considering the numbers of data level are considered, called Cross Test too. If the number of input variable is N_v and the number of data level of every variable is $q_i (i=1,2,\dots,N_v)$, the test number

$N_t = \prod_{i=1}^{N_v} q_i$. It is adopted in bound test and normal test, because it can test model characteristics fully. However, N_t is too large for us to use it, when N_v or q_i is large. For example, N_t is 625 when q_i is 4 and N_v is 5.

(5) *Uniform Test*

In order to make $\langle V \rangle$ distribute in input space fully in limited test times, uniform test is adopted. Here it is carried out by two steps: produce uniform design table and assign $\langle DV \rangle$ based the design table.



The techniques of producing uniform design table include God Lattice Pint Method (GLPM)[4], Latin Squares[5], and so on. For example, the design table $U_{10}(10^3)$ of 3 factors and 10 test times is presented in Table.1, produced by GLPM.

Table 1. $U_{10}(10^3)$ produced by GLPM

| No. | Factor 1 | Factor 2 | Factor 3 |
|-----|----------|----------|----------|
| 1 | 1 | 8 | 9 |
| 2 | 2 | 5 | 7 |
| 3 | 3 | 2 | 5 |
| 4 | 4 | 10 | 6 |
| 5 | 5 | 7 | 1 |
| 6 | 6 | 4 | 10 |
| 7 | 7 | 1 | 8 |
| 8 | 8 | 9 | 6 |
| 9 | 9 | 6 | 4 |
| 10 | 10 | 3 | 2 |

If the number of design variables is N_v and test time is N_t , design table $U_{N_t}(N_t^{N_v})$ is needed. Based on $U_{N_t}(N_t^{N_v})$, matrix $\mathbf{U} = (u_{ij})_{N_t \times N_v}$ is obtained and $\langle V \rangle$ is assigned as follows:

$$v_{ij} = V_{i\min} + (u_{ij} / N_t) \times (V_{i\max} - V_{i\min}) \tag{4}$$

In this way, $\langle V \rangle$ is obtained.

4 Work Flow of Model Test

Model test of simulation in virtual test is a very complex job. A normal workflow is needed to offer guidance for engineering application. Fig.1 presents the flow chart of model test.

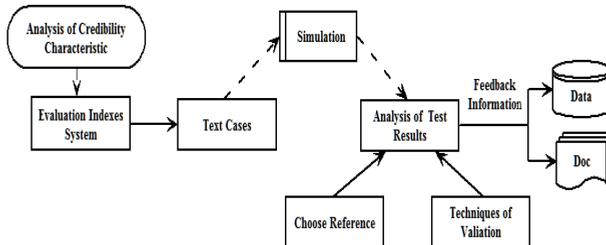


Fig. 1. Flow chart of model test

① *Analysis of credibility characteristic of model*

Any model has its credibility characteristics in essence of physics. The credibility characteristics of object model are needed to be analyzed fully at this stage. Pick some clear characteristics as valuation aim of model test. Besides, appropriate forms of indexes expression are needed to be chosen to denote the picked credibility characteristics. At last, integrate all indexes picked above and design evaluation indexes system for model test. In a word, the evaluation aim is need to be decided at this stage.

② *Design test cases*

Based on the analysis of credibility characteristics and designing evaluation indexes system, the next work is to design effective test cases. The work of designing one test case includes picking one credibility characteristic (*CC*) as aim for this case, choosing sound design variable ($\langle DV \rangle$), picking appropriate test technique (*TT*) to assign $\langle DV \rangle$ to get the values of design variables ($\langle V \rangle$), and collecting the most reliable evaluation reference (*ER*). In a word, several test cases $\{CC; \langle DV \rangle, \langle V \rangle, TT, ER\}$, which can evaluate the credibility of object model fully, are designed at this stage.

③ *Simulate and collect data*

According to the inputs offered by test cases designed in stage②, carry out simulations. After every simulation, collect data needed in analysis of simulation results in next stage, and save them in stated format.

④ *Analysis of simulation results*

At this stage, the main work is analyzing simulation results and evaluating the credibility of object model according to the evaluation reference collected in stage②. The validation techniques of VV&A are used at this stage.

The work of analysis is divided into two steps: evaluation of indexes at the bottom of indexes system and the integrated evaluation of model based all indexes evaluation.

⑤ *Feedback of analysis results*

Getting evaluation results of object model is not the ending of model test, the more important job is to make the evaluation results feed back into M&S. If errors in modeling are discovered, object model can be modified in time. If there is no error, the developers and users of model can know about the credibility of model better and have a every confidence in the model.

5 Design of Toolbox for Model Test

In order to offer software support for techniques and process manage, toolbox for model test is designed and actualized in this paper, based on the basic theory of model test studied above.

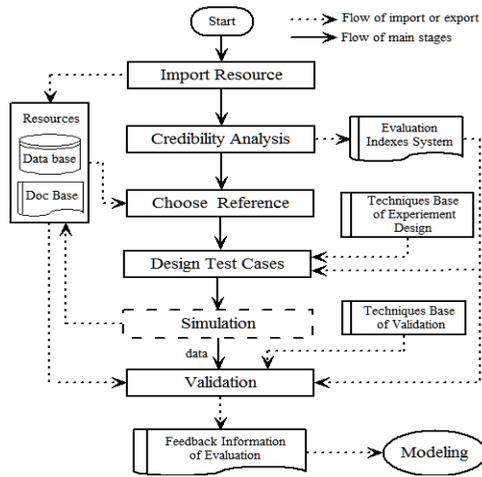


Fig. 2. Operation flow of toolbox

(1) Units of toolbox

The toolbox for model test is made up of 4 units as follows:

Resource base: It is the storage that store various resource used or exported in model test and evaluation. The resources include model information, simulation results, and evaluation results.

Techniques base of experiment design: The base can offer all techniques of experiment mentioned in section 2. Besides, the contents of every test case designed can be saved in configuration file. The configuration of test case records all information of test case.

Techniques base of validation: There are parts in this base. One is the techniques of mature techniques of validation, such as TIC, Correlation Method, Spectrum Analysis Method, etc. The other is the techniques of techniques of information integration, such as Expert Mark Method, AHP, Fuzzy evaluation, etc.

Visualization: The curve of parameter data can offer intuitional evaluation of credibility for object model.

(2)The flow of operation of toolbox

The flow of operation of toolbox is designed based on the workflow of model test mentioned above. Fig.2 presents the operation flow.

6 Application Cases

The studies above on theory of model test can offer guidance for project application, and the actualiz-a tion of toolbox can promote the efficiency and veracity of model test. They have been used in the model test of one simulation model of aerodynamic. The input and output of object model is as follows,



Table 2. Input and output of object model

| | Variable | Input Space | | Variable |
|-------|----------|--------------|--------|----------|
| Input | Ma | [4 , 25] | Output | C_L |
| | $alpha$ | [0° , 20°] | | C_D |
| | h | [0 , 100000] | | |

Here, Ma is Mach number, $alpha$ is attack angle, h is the height, C_L is lift coefficient, and C_D is drag coefficient.

Now, adopt trend test and normal test to evaluate the credibility of object model as follows:

(1) *Trend test*

Based on analysis of domain expert, the change trend of C_L/C_D (lift-drag ratio) changes with $alpha$ can reflect the credibility well. So, design trend test case using Control-Factor method as follows:

CC is the change trend of C_L/C_D ;

$\langle DV \rangle$ is $\langle Ma, alpha, h \rangle$;

TT is Control-Factor method;

ER is the specialty knowledge of domain expert;

$$\langle V \rangle \text{ is } \left\{ \begin{matrix} 10 & 2 & 70000 \\ 10 & 4 & 70000 \\ \vdots & \vdots & 70000 \\ 10 & 20 & 70000 \end{matrix} \right\}_{10 \times 3} .$$

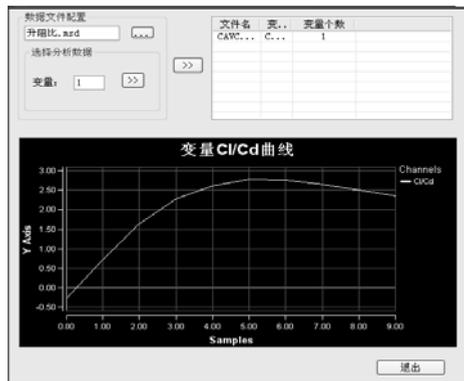


Fig. 3. Change trend of presented in toolbox



Fig. 4. The toolbox interface of normal test

(2) Normal test

Firstly, design $U_{10}(10^3)$ adopting GLPM, presented in Table.1. Then assign $\langle Ma, \alpha, h \rangle$ according to equation 4. Fig.4 presents the actualization of them in toolbox.

The simulation data of C_D and C_L is got after simulations, denoted by X_1 and X_2 . Y_1 and Y_2 are the real experiment data of C_D and C_L , got under the same input. TIC is adopted to check the consistency of X_1 and Y_1 , X_2 and Y_2 , for they are all small samples.

The result of TIC is that the consistency of X_1 and Y_1 is 0.97 and that of X_2 and Y_2 is 0.95, which presents that the credibility of aerodynamic model is high, and the model can be used. Fig5 presents interface of consistency check of toolbox.



Fig. 5. The toolbox interface of validation



7 Summary

Based on the requirements of project application, the basic theory of model test is studied in this paper, including the definition of test case, techniques of experiment design, evaluation indexed system, workflow of model test, and actualization of toolbox. All this have been applied in the model test of one simulation model of aerodynamic. The result shows that the studies on model test have a good application prospect in engineering.

References

1. Tang, J., Cha, Y., Li, G.: An Overview of the Research on VV&A in Simulation. *Computer Simulation* 23(11), 82–85 (2006)
2. US DOD Defense Modeling and Simulation Office (DMSO). Verification, Validation and Accreditation Recommended Practice Guides(RPG) [EB/OL], <http://www.vva.msc0.mil>
3. Sun, Y.: Research of Verification, validation and Accreditation Technology in Modeling and Simulation. Dissertation for Ph.D of Nanjing University of Science & Technology, Nanjing (October 2005)
4. Wang, Y., Fang, K.T.: A note on uniform distribution and experimental design. *Chinese Science Bulletin* 26, 485–489 (1981)
5. Fang, K.T., Shiu, W.C., Pan, J.X.: Uniform designs based on Latin squares. *Statistica Science* (9), 905–912 (1999)

The Learning Evaluation of Learning Enterprises Based on BP Neural Network

Yong Zhang

College of Control Engineering, Chengdu University of Information Technology, NO.24,
Block 1 Xuefu Road, Chengdu 610225 P.R. China
School of Public Administration, Sichuan University, Chengdu, 610064, P.R. China
scuzhangyong@126.com, h-y18@live.cn

Abstract. The Learning Company is a modern new enterprise with the characteristics of a learning organization, the strength of learning has decisive importance on its survival and development. For the shortcomings of current evaluation method to enterprise learning, this paper put forward an evaluation method of learning ability based on BP neural network. First, establish a Learning ability evaluation index system based on learning organization, and designs BP neural network model, the use of MATLAB neural network toolbox for network learning and sample Test to verify the model is reasonable and effective.

Keywords: Enterprise learning, evaluation index system, BP neural network.

1 Introduction

The impact of learning ability on enterprise survival and development. In order to deal with the changing internal and external environment, enterprise should obtain and apply new knowledge and technology, and improve their own behavior and optimize organizational structure. The dynamic organizational capability that is embodied in this process is called enterprise learning ability. For modern enterprise, the learning ability is not only the premise and guarantee of learning speed and quality, but also the determinant on the difference of organization learning efficiency, which is the fundamental factor affecting enterprise competitive advantage. Through the study, enterprise can not only adapt to the changing environment, but also can trigger innovation, and promote the enterprise development and competitiveness. At the same time, effective learning is the basis of competitiveness of enterprises. Learning ability is also a kind of core ability which other enterprise cannot imitate and copy. Its shape and enhancement have huge significance for the existence and development of enterprises of learning style.

Enterprise as a huge system needs continuous learning to improve the ability of survival and development. The lack of learning ability cannot make the enterprise rapidly strain in environmental change, which severely damages its survival and development. With the development of knowledge economy, the learning organization has become a trend of the enterprise development. As a positive motivation of

enterprise technological innovation, learning has become the enterprise core competitive ability and the important way to gain competitive advantage in the fierce competition in the market. Now it has drawn more and more attention of scholars and entrepreneurs.

Defects of domestic research on enterprise learning ability. At present, many domestic scholars such as Rui Mingjie, Chen Guoquan, Wu Jiabao etc. has studied the learning ability of enterprises. The evaluation method mainly includes analytic hierarchy process (AHP), fuzzy comprehensive evaluation method and grey clustering method. Its deficiency mainly reflects in: (1) the selection of evaluation index is not comprehensive and objective; (2) the weights of evaluation are most determined by researcher's experience. Because of lack of scientific subjective theory basis, the precision and reliability of the evaluation of the results are questioned frequently. Therefore, this article comes up with a new method which is based on BP neural network. It is with the consideration of the deficiency about current evaluation methods of enterprise learning ability and combined with the unique advantages of artificial neural network. That is the enterprise learning capability evaluation method.

As a mature and widely applied feed-forward neural Network, the BP neural Network (Back-a Propagation) can realize the function of linear and nonlinear mapping arbitrary, avoid man-made identify weight, and reduce the randomness in evaluation process and the uncertainty and fuzziness of evaluation people. Through the systematical training and learning, we can make the system error reach the demand of precision, improve the reliability of the evaluation and make evaluation results more effective and objective[1].

Enterprise learning ability is influenced by many factors, so that the ability of learning evaluation is related to various qualitative and quantitative factors. The factors of influence are not isolated. They are connected with each other and restrained to each other. So they form into a complex and highly nonlinear system mutual. As a kind of effective tool solving nonlinear system problems, this paper attempts to design an information processing model of learning enterprise by using BP neural network and use it to give a comprehensive evaluation on learning ability of learning enterprise.

2 The Construction of Evaluation Index System of Learning Ability in Learning Enterprises

The study of enterprise learning ability can not only make the manager know more about the type and characteristics of various learning ability, but also provide the reference for enterprises to improve their learning ability and cultivate core competence.

At present, many scholars at home and abroad have taken a large amount of research on learning ability evaluation index system. Among all evaluation index systems established by the scholars, there are different emphasis and many things in common. Goh, and Richard first define organization learning ability and developed a set of organization learning ability assessment tools, including five aspects: the clarity of the purpose and mission, leadership commitment and authorization, experiment and reward, knowledge transfer and team work with groups solve problems [2]. Through the analysis, Lahteen maki etc. develop a measurement and evaluation index system of organization learning ability which include learning subject, learning context and

learning process the three dimensions[3]. Thereafter, the domestic scholar Wu Jiabao through the application layer induces the promoting factors of organization learning to evaluation organization learning ability. He combined with others research results such as Goh, increased "employee education and training" and "organizational culture" the two dimensions and developed a evaluation index system include seven assessment dimensions and 35 specific index evaluation [4]. Based on organizational learning theory and empirical research of Sinkula, Baker, Noordewier, Senge etc., Li Yi combined with Kert Scale's five scales, designed eleven questions to value the enterprise learning ability, including: clear vision, Vision consensus, Vision of participation, Questioned common assumption, opening think, think mode changes, grow in communication, needed knowledge gained smoothly, Effective sharing experience and tastes and effective knowledge transfer the ten aspects [5]. Except this, Wang Xiaobin, based on the anglicize of promoting factors, also has established six indicators evaluation index system. They are ranked from important rate from high to low like this, organizing maturity, organizational efficiency, organizational vision, the organization leadership, organization human resources, and organizational culture.

This paper is based on the five indicators which learning organization must possess (individual beyond, improve mental model, building Shared vision, team learning and systematic thinking) and with reference of literature research at home and abroad. According to the principles of index system building including scientism, systematism, orientation, comparability and the dependence of each indicators, this paper establishes a learning capability evaluation index system as shown in Fig. 1.

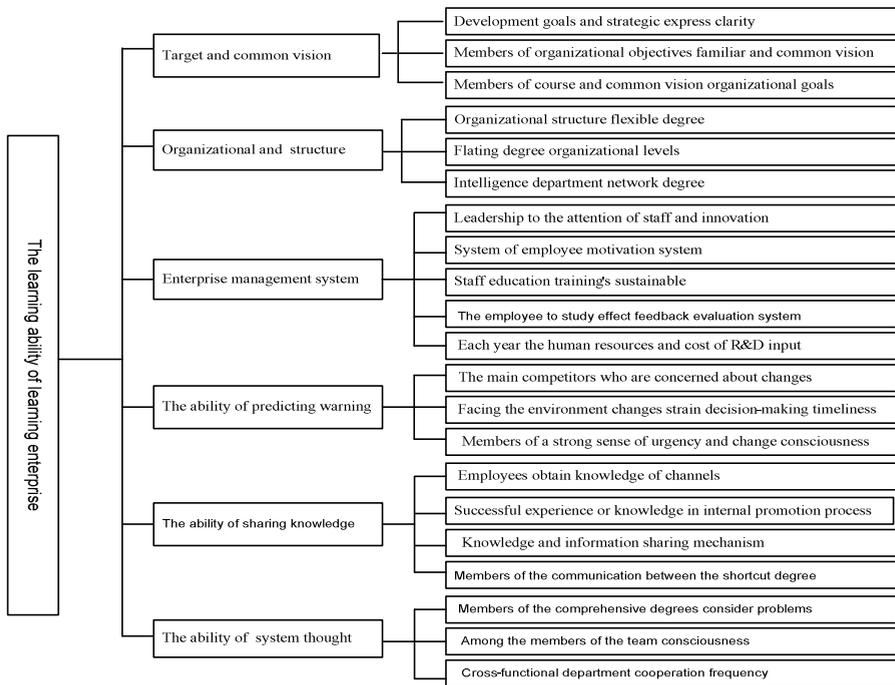


Fig. 1. Learning enterprises learning capability evaluation index system

3 Learning Ability Evaluation Model of Learning Enterprise Based on BP Neural Network

The BP neural network evaluation model of learning enterprise’s learning ability.

The process of BP neural network model includes: firstly establishing neural network structure, namely determining inputs node, the output node number, the number of hidden layers and layers of knot points each. Then the BP network structure will be adjusted and the initial network parameters will be established. The neural network training studies begins. When the network after learning becomes stable and the output as specified error or iterative training requirements reaches to the maximum number of regulations, the training ended; Using test data set to test the trained neural network, if it reaches the required accuracy, it means establishing a BP neural network model.

Kohnogrov theoretical proof: any given continuous function $\Phi = X \sim Y, X, Y \in (R^n \in [0, 1] n$, then can accurately Φ by a three-story network realization. Therefore, only need to determine inputs layer, hidden and output layer node is finished building number of neurons in the task of BP neural network structure.

1) Determine the input layer node number i : according to established evaluation index system of learning enterprise’s learning ability, take the lowest level index number as input layer neurons number. In this paper it is 21. So the number of nodes in the input of the network layer is 21.

2) Determine the hidden nodes number j : the number of neurons select in hidden layer relates to the BP network precision and the learning efficiency, We can use empirical formula to calculate, $j = \sqrt{n + m} + a$, $j = \sqrt{\log n}$ or $j = \sqrt{nm}$. Among them, m represents the output node, n represents input node number, a is the constant between 1 and 10. In this paper, we combine with the theoretical analysis and experience and Select hidden neurons number of 14.

3) Determine output layer node number t : the evaluation of learning ability is a process which is from qualitative to quantitative to qualitative again. Through BP network model, we can change qualitative into quantitative output, then synthesize evaluation sets and output, at last make qualitative evaluation of the learning enterprise’s learning ability. So, we set the output layer as 1, $t=1$.s represents the comprehensive evaluation value. We state $S \in [0, 10]$, the evaluation sets is like this: poor, weak, modest, good , exceptional. It shows as the table 1.

Table 1. Learning ability level of learning enterprise

| the comprehensive evaluation value of learning enterprise’s learning ability | $0 \leq S < 2$ | $2 \leq S < 4$ | $4 \leq S < 6$ | $6 \leq S < 8$ | $8 \leq S \leq 10$ |
|--|----------------|----------------|----------------|----------------|--------------------|
| learning ability of learning enterprise | poor | weak | modest | good | exceptional |

The BP neural network evaluation model. According to the basic principle of neural network, we design BP neural network evaluation model of learning enterprise in Fig. 2.



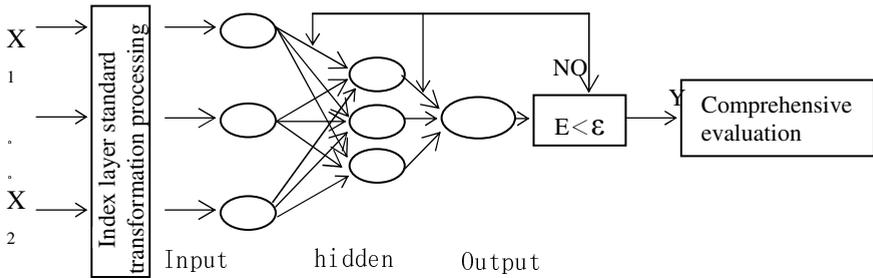


Fig. 2. The BP neural network evaluation model of learning ability

4 The Evaluation Procedure of BP Neural Network Model of Enterprise Learning Ability

A BP neural network model is formed based on the model structure design. The procedure of the evaluation of the learning enterprise's learning ability is as follows:

1) First, BP neural network model structure of the evaluation of the learning enterprise's learning ability should be figured out, which includes the neural network, the number of piles, input, implication and the number of the node neuron of output of three piles. According to the calculation of the preceding text, the BP neural network of this text has adapted as to the evaluation of the learning enterprise learning ability, the number of the network layers is three; the number of the input layer's neuron is 21; the number of the implication's is 14; and the number of the output layer's is 1. And the grade of the evaluation of the enterprise learning ability is rated in five levels: weaker, weak, general, strong and stronger.

2) BP neural network studying model means what is described in the above target system. It collects the values $\{X_{pi}\}$ of the study samples from different companies, then conducts a standardized process to the values, and finally determines the weight W_n of various layers of the neuron by using MATLAB for network training.

5 Case Study

Studying on 40 Sichuan enterprises's learning ability evaluation, we selected 30 enterprises as the training sample. In order to strengthen the generalization of the performance, 30 enterprises which we selected contains the private enterprise, state-owned enterprises, and foreign enterprises. Other 10 enterprises are the test samples.

We make a programming using MATLAB 7.0 [6] and start a model from the enterprise's date and give an evaluation.

BP neural network model includes 21 input layers, 14 implication layers, and one output layer according to the already established the model. The output of output layer changes from 1 to 10. And the transfer function, between the input layer and the implication layer, and between the implication layer and the output layer, adopts the logarithmic function and logsig function.

Then, we can get the network-learning started after inputting studying sample assessed accurately by experts. Set error target is 10^{-4} . If the Network error sum of squares and MSE meet the condition of target error, then we can conclude that the evaluating models for learning ability of enterprises have been established based on BP Neural Network, and the training of learning samples ends.

In the process of standardizing the evaluation index, we should pay attention to the different treatment to the qualitative index and quantitative index. Firstly, we must conduct standardization and homogeneous treat to the quantitative index for the different units of measurement and the methods of treat are as follows:

a) In the case of the bigger the goal is, the better appraisal,

$$F_j = (X_j - X_{j\min}) / (X_{j\max} - X_{j\min})$$
.

b) In the case of the smaller the goal is, the better appraisal,

$$F_j = 1 - (X_j - X_{j\min}) / (X_{j\max} - X_{j\min})$$
.

In a) and b), F_j is the standardized value to the target value, X_j ; X_{\min} is the minimum of the j -th index which predetermined, X_{\max} is the maximum of the j -th index which predetermined, j is the quantity of evaluation index.

Secondly, the qualitative index is a description to the level or grade reached for the phenomenon rather than an expression of the standardized statistical index. To facilitate the measurement, it is common to make use of fuzzy statistical method, which rank these qualitative index into five standard: the value is 0.9, 0.7, 0.5, 0.3, 0.1 for exceptional, good, modest, poor, weak, respectively. Let the experts involved in the evaluation adopt the above evaluation standard, and compute the mean value, which is the evaluation value of this index.

After 1583 steps training, we meet the accuracy requirement. The training error is shown in figure 3.

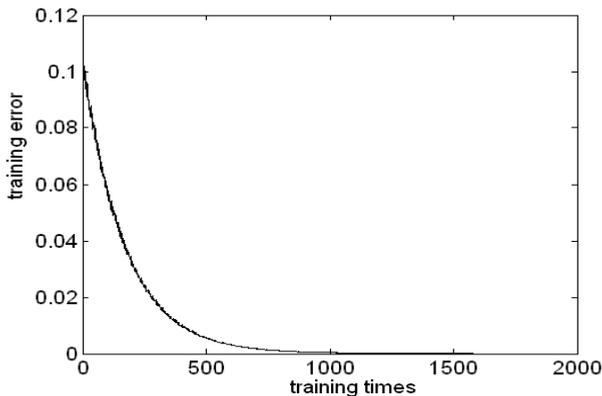


Fig. 3. Network training error drop curve

We test the 10 samples, and calculated their simulation results (output) and grades separately. To measure the effectiveness of the method, the evaluation of each enterprise expert value is listed in the table 2.

Table 2. The result of the testing samples and error list

| Test numbe | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|--------------------|-------|------|--------|-------|-------|---------|--------|-------|-------------|-------|-------|
| Model | index | 1.21 | 5.33 | 2.62 | 7.14 | 4.21 | 5.97 | 3.45 | 8.16 | 6.72 | 3.32 |
| output | grade | poor | modest | weak | good | general | modest | weak | exceptional | good | weak |
| Expert | index | 1.29 | 5.13 | 2.71 | 7.72 | 4.34 | 6.26 | 3.18 | 8.57 | 6.44 | 3.51 |
| evaluation | grade | poor | modest | weak | good | general | good | weak | exceptional | good | weak |
| value | | | | | | | | | | | |
| Relative error (%) | | 6.2% | 3.9% | 3.32% | 7.51% | 3.00% | 4.63% | 8.49% | 4.78% | 4.35% | 5.41% |

Simulation results show that, in 10 the test samples, one enterprise's learning ability is weak ($0 \leq S < 2$), 3 of them learning ability are poor ($2 \leq S < 4$), 3 of them learning ability are modest ($4 \leq S < 6$), 2 of them learning ability are good ($6 \leq S < 8$) and 1 of them learning ability is exceptional ($8 \leq S < 10$). meanwhile, from table 1 we can see, there is certain deviation between the results of using neutral network and expert evaluation. But in a certain extent, it proved the Validity and practicability of the model. Because of the fewer samples above the study, the BP neural network finally determined is not very stable and each training results will have different amplitude change. However, the calculated measure results have already can evaluate enterprise learning ability accurately. It also proves the feasibility of BP neural network model in enterprise learning ability evaluation.

6 Conclusion

The study and assessment to a learning enterprise's learning ability help the management to understand the level of development of the enterprise's learning ability at present. They also offer service and reference to the managers and the authorities, which is significant to an learning enterprise's existence and development. It is established an evaluation index system to enterprise's learning ability based on learning organization, in addition, an evaluation model has been designed out by the use of BP Neural Network method, which offer an scientific evaluation method to the enterprise's learning ability in this paper. BP Neural Network is the most widely applied artificial Neural Network in every field, and it is best at processing the problem of high-dimensional nonlinear mapping whose law is hidden below numerous messy data. There is no need for it to establish a mathematical model, however it summarize the knowledge from the data through network training and restore the knowledge in the neurons in the terms of several weights and threshold which constitute the network knowledge, then, use this kind of knowledge to evaluate and predict the result of similar factors. But it also has some shortcomings, for instance, the quantity and quality of the selected learning sample affect the learning performance of Neural network model to a large extent, the selection of network layer and number of neurons in the hidden layer will also influence the learning ability and learning efficiency of the whole network and the like. However, the theoretical and applied research of this essay is still stand at the

initial stage. I believe , along with the deep study and application of the Neural network model in the field of the evaluation of the enterprise's learning ability, BP network model based on the MATLAB Neural network tool box will become an effect method to evaluate the enterprise's learning ability.

References

1. Han, L.: The theoretical design and application of artificial Neural network. Chemical Industry Publisher, Beijing (2002)
2. Goh, S., Richard, G.: Benchmarking the learning capability of organizations. *European Management Journal* 15(5), 575–583 (1997)
3. Lahteenmaki, S., et al.: Critical aspects of organizational learning research and proposals for its measurement. *British Journal of Management* 12, 113–129 (2001)
4. Wu, J.-B.: The Measurement of Organizational Learnings Capability. *Chinese Journal of Management Science* 11(4), 73–78 (2003)
5. Li, Y.: Research on organizational learning method. Fudan university PhD thesis, ShangHai (2002)
6. Xu, D., Wu, Z.: Analysis and Design of MATLAB System—BP neural network. Xidian University Press, XiAn (2002)

An Approximate Algorithm to Solve the Location-Selection of Wireless Network Problem

Xiao-xu Lu and Wei Zhang

Department of Mathematics and Physics of Zhengzhou Institute of Aeronautical Industry Management, 450015, Zhengzhou, China
xiao2828@yeah.net

Abstract. The minimum Steiner tree problem has wide application background, such as transportation system, communication network, pipeline design and VISL, etc. It is unfortunately that the computational complexity of the problem is NP-hard. People are common to find some special problems to consider. Since the complexity of the Steiner tree problem, the almost of papers are relate to the object of small data problem, i.e., the number of involved objects is small. Those conclusions are useful to the theoretical research from which some algorithms are originated. For the practical problems, there are large number of objects are need to be considered. In this paper, we introduce an approximate algorithm and give an example with seven objects to consider the location-selection of wireless network problem.

Keywords: restricted Steiner tree, computational complexity, location-selection, wireless network.

1 Introduction

On building communicating network problem, people should be considered in the construction of some cities (site) how lay cable, to make these cities become a connected networks and at the same time to consider how to minimize the cost of the total attachment in traffic. It has the similar problem in transmission system. Since the 1950s, information science has the vigorous development, such that promoting the shortest connection of this kind of problem, formed the research field of a very active combination optimization's topic. There are too much models on this area, roughly divided into two categories: the attachment problem in the plane and the network connection problems. For the former, it is considered the connection is the straight line of the plane and the definition of distance is Euclidean distance; For the latter the connection is considered the shortest path in the network, where the distance of two vertices is equals the distance of the shortest path between the vertices.

On location-selection of wireless network, the consumer always considers whether the network is smooth or not. In the common case, the signal of the network bases on the distance between the signal launch and consumer. If the distance is closer, then the signal is stronger. We want to build more launchers but the cost is too expensive. Then we want to find the suitable place to set the launcher such that the sum distance between

the launcher and consumers is minimal. The problem is equitable to the famous Steiner tree problem. The computer complexity is NP-hard. In this paper, we consider some special cases of the Steiner tree problem. If we want to build a launcher, then we have to consider the situation of the launcher, such as the geographical environment, the artificial construction etc. Without loss of generality, we demand the launcher in the fixed line L . We assume the launcher must be placed on the fixed line L and the consumers are on the same side of the line L , i.e., the restricted Steiner tree problem in [8] and [9].

In this paper we consider the connection line problems in the plane. In the plane of the shortest connection problem is divided into two types: the first kind only consider the connection among some fixed vertices and all attachment is constructed the spanning tree of these vertices. Therefore, this problem is called the spanning tree problem. The second is allowed to add some attachment vertices such that the distance of all attachment edges is constituted the shorted network; this is called the shortest Steiner smallest tree problem.

The minimal spanning tree problem is originated from the comprehensive work of Kruskal and Prim. In the plane of the minimum spanning tree problem can be stated as follows: Given a set N which has some fixed vertices and $|N|=n$ in a plane, where the distance of each line (or edge) $e=(u, v)$ is defined the Euclidean distance between the two vertices u and v . Then we can construct a weight graph G according to the model: $V(G) = N \cup P$, $V(G)=w(e) = d(u, v)$, where P is the Steiner vertex. Now for a spanning trees of G , namely a connection graph with the vertex set is N to make the sum weight of edges minimize. Kruskal and Prim has established by the famous Greedy algorithm, even applied to mastoid optimization problem. By using the algorithm, the spanning tree problem can be briefly described as follows:

- (1) for the graph G , we can sort the distance of the edges;
- (2) we choose a smallest edge to construct an edge set;
- (3) choose one the one the shortest edge e which is not selected, if e and selected edges does not constitute a circle, then we choose the edge to the edge set, otherwise, we abandon the edge;
- (4) the algorithm does not end until we choose $n-1$ edges to the edge set, such that we have a smallest spanning tree T of the graph G .

The algorithm time bound of the algorithm is $O(n^2 \log n)$, now there exists more effective polynomial algorithm. Minimum Steiner tree problem is originated from three-point sitting Fermat problem, where there are three villages whose position fitly constitute a triangle, and assumed the ground can be used anywhere for all road's construction. The problem is how to design a graph G such that the villages are connected and the distance of roads reached the shortest.

The answer to this question is very simple. We can choose one point inside the triangle such that the point attachment angle to the three points is 120° . This is an

optimal plan (if the maximum angle among the triangle is at least 120^0 , then the optimal plan is the two shorter edges.)

There are many papers to consider the spanning tree problems. In [6], Bing Yao etc considered some properties: a spanning tree with minimum leaves such that it contains a longest path, and the spanning tree with the maximum leaves about the diameter of the graph; In [7], Ming-yi Yue gives some new and simple proofs for the cases $n=3, 4$ and 5 , and a proof of the Steiner Ratio Conjecture. Some researchers consider some application on Steiner tree problem, such as Hui-min Jin, in [8], considered intelligent optimization algorithms for Euclidean Steiner minimum tree problem, and there are two intelligent optimization algorithms: the simulated annealing algorithm and the ant algorithm. In the paper, we give an approximate method to solve the large number of units, which the distance between the units and the recycle water station is minimum. In [11] we introduce an approximate algorithm to solve the location-selection of recycle water station problem, where first find one line L by the Least Square Method, then we choose one suitable vertex on the L such that the distance between the vertex and the units is minimum.

2 Approximate LLQ-Algorithm

The problem is proposed by the environmental protection. The model is there are n consumers, the plane coordinates (x_i, y_i) , $i = 1, 2, \dots, n$, it is needed to build one station to connect the consumers such that

$$\min_{(x,y)} f = \sum_{i=1}^n \sqrt{(x - x_i)^2 + (y - y_i)^2}.$$

The distance between the units and the station is Euclidean distance. The problem is to select a suitable location such that the sum distance between the units and the station is minimal. It is easy to see that the problem is equitable to the minimum Steiner tree problem. Since the computer complexity is Np -complete, it is difficult to find the accurate Steiner vertex to minimize the distance. Therefore, there are lots of papers are considered the small data Steiner tree problem such that $|N| = 2, 3, 4, 5$ and some special cases such as the restricted condition, like the Steiner vertex on the line, some curve etc. For the practical problem, like we proposed the location-selection of wireless network problem, maybe there are many consumers to connect the station, the previous achievements cannot to solve the model. We construct an approximate algorithm to solve the model.

Least square method is a kind of optimal technique. It can determine the unknown data and make deviation between these values of data and the practical data minimal. The method can also be used to curving fitting. There are n units, denoted the plane coordinates (x_i, y_i) , $i = 1, 2, \dots, n$. We use the Least square method to find a

quadratic curve L. Nextly, we find a vertex P on the line L such that the distance between the vertex and all of units is minimum. Therefore, P is an approximate Steiner vertex of the problem.

Approximate Algorithm

(1) for units (x_i, y_i) , $i = 1, 2, \dots, n$, we use the least square method to find the quadratic curve $L: y = ax^2 + bx + c$, this is can be directly attained by the matlab computation;

(2) we find one vertex $P \in L$ such that the distance between P and the all of consumers is minimum, this is conditional extreme value problem, and we can use the Lagrange multiplier method to find the vertex.

In order to simple purpose, we say the approximate algorithm by LLQ-algorithm.

3 Application of the LLQ-Algorithm

We use the approximate algorithm to solve an example. There are seven units and the according coordinates are $x=[17,19,21,23,25,27,29]$, $y=[20,25,26,30,26,20,19]$.

Step 1 find the fixed line L using the Least square method by matlab

The matlab procedure is

```
poly- curving fitting
```

```
%%% input data
```

```
x0=17:2:29;
```

```
x0=[x0];
```

```
y0=[20 25 26 30 26 20 19];
```

```
%%% coefficient fitting p
```

```
[p,s]=polyfit(x0,y0,2);
```

```
p
```

```
%%%%%%%%%y incredible radius delta
```

```
[y,delta]=polyconf(p,x0,s);
```

```
y
```

```
%%%%%%%%% output image
```

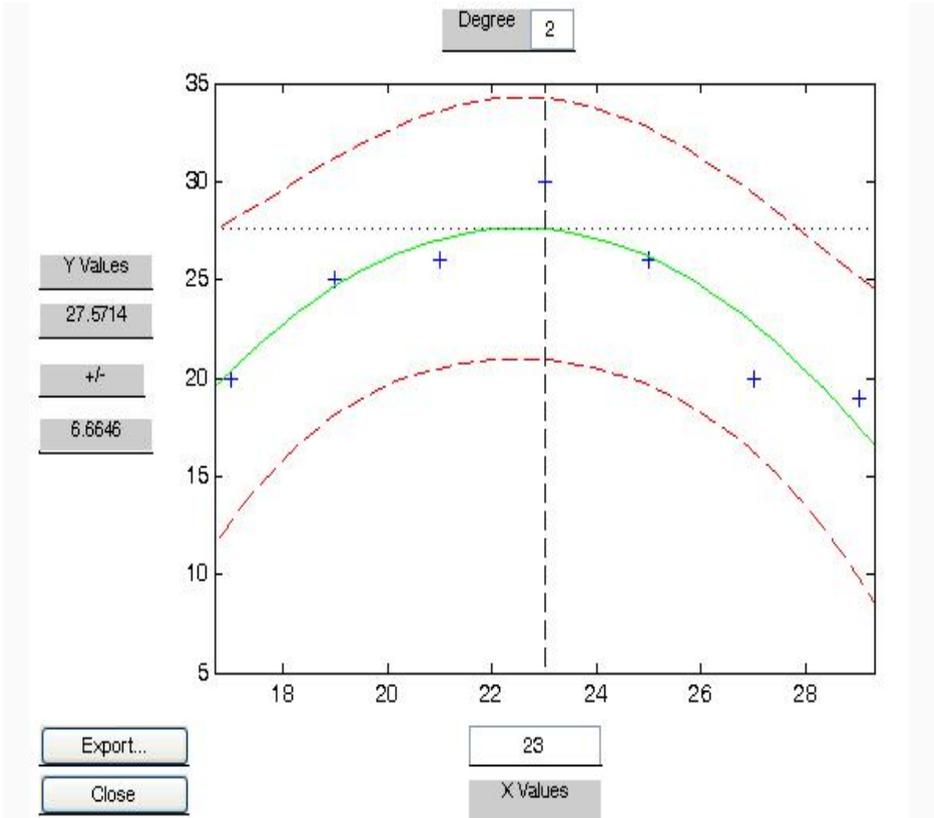
```
polytool(x0,y0,2)
```

```
p =
```

```
-0.2411  10.8571  -94.6161
```

```
y =
```

```
20.2857  24.6429  27.0714  27.5714  26.1429  22.7857  17.5000
```



Therefore we have the function of quadratic curve L $y = -0.2411x^2 + 10.8571x - 94.6161$.

Step 2 find a suit vertex P on L such that the distance between P and vertices is minimum using the Lagrange multiplier method.

To simplified calculation, we consider the square of distance instead of the distance.

$$\min f = \sum_{i=1}^7 ((x - x_i)^2 + (y - y_i)^2), \text{ where } y = -0.2411x^2 + 10.8571x - 94.6161$$

By the similar computation we can have

$$x \approx 25.96, y \approx 25.32, \text{ and } \sum_{i=1}^7 \sqrt{(x - x_i)^2 + (y - y_i)^2} \approx 41.54,$$

and the distance between the each unit and the station is approximate 10.42, 6.97, 5.00, 5.54, 1.18, 5.42, 7.01, respectively.

4 Some Further Problems

Since the computation complexity of the famous Steiner tree problem is Np-complete we consider an approximate algorithm to solve it. We find a suitable vertex on a special

line L to approximate the Steiner vertex using LLQ-algorithm. Next, we can use the least square method to find some special curve, where locate the suitable vertex on the curve such that it approximates the Steiner vertex. Although the approximated vertex is not the accurate Steiner vertex, it provides a method to solve the location-selection of the Steiner tree problem. By comparing with the result of [11], the distance is less than the before, where the distance between the approximate vertex and the units is 44.5. It is natural to think that if we add the power of the function, the distance may be shorter, i.e., the approximate is trend to the Steiner vertex when the power of function is more greater.

Acknowledgment. The authors thank for financial supports from National Natural Sciences Foundation's Youth Science Fund (10901144). Vol. 83-87 (1992), p. 119.

References

1. Bondy, J.A., Murty, U.S.R.: Graph Theory with Applications. American Elsevier, New York (1976)
2. Melzak, Z.A.: On the problem of Steiner. J. Canad. Math. Bull. 4(2), 143 (1961)
3. Hwang, F.K., Richard, D.S.: Steiner tree problem. J. Network 22, 55 (1992)
4. Yue, M.: Minimal Network-The problem of Steiner Tree. Zhejiang Scientific and Technological Publishing House, Hangzhou (2001)
5. Yue, M., Cheng, C.: A Note on the Steiner Problem-An approach to find the minimal network with five given points. OR Transactions 14(1), 1 (2010)
6. Yao, B., Zhang, Z.-F., Wang, J.-F.: Some results on spanning trees. Acta Mathematica Application Sinica, English Series 26(4), 607 (2010)
7. Yue, M.-Y.: A Report on the Steiner Ratio Conjecture. OR Translations 4, 1 (2000)
8. Jin, H.-M., Ma, L., Wang, Z.-M.: Intelligent optimization algorithms for euclidean Steiner minimum tree problem. Computer Engineering 32, 201 (2006)
9. Lin, D., Du, L., Lu, X.: A Restricted Steiner tree Problem I. Accepted by ICIME 2011 (2011)
10. Zhang, Y., Lu, X., Lin, D.: A Restricted Steiner tree Problem II. Accepted by ICIME 2011 (2011)
11. Lu, X.-X., Yang, S.; Zheng, N.: An approximate algorithm to solve the location-selection of recycle water station problem (submitted)

A New Supervised Discriminant Locality Preserving Projections Algorithm

Jun Yu, Jintao Meng, and Xiao-xu Lu

Zheng Zhou Institute of Aeronautical Industry Management, He Nan 450015, China
jun1839@sina.cn

Abstract. When performing visualization and classification, people often confront the problem of dimensionality reduction. Recently, a new manifold learning algorithm named Locality Preserving Projections (LPP) that aims at finding an embedding that preserves local information has been proposed. But LPP is an unsupervised algorithm. Supervised locality preserving projection (SLPP) seeks to find the projection which efficiently preserves the local structure of data points embedded in high-dimensional data space. However, it has the over-learning problem and does not preserve the class and discriminative information of data which is also useful for data recognition. To solve the problem, a novel feature extraction method based on discriminant information, namely supervised discriminative locality preserving projections (SDLPP), is presented in this paper. The SDLPP takes into account class and discriminative information of data points. Different from the most existing supervised locality preserving projection methods, the SDLPP not only preserves both the local structure and diversity information of data, but also avoids the data over-fitting problem. We compare the proposed SDLPP with LPP and SLPP methods on different data sets. Experimental results suggest the efficiency of the proposed method.

Keywords: Dimensionality reduction, Supervised locality preserving projection, Discriminant information.

1 Introduction

When performing visualization and classification, people often confront the problem of dimensionality reduction. Dimensionality reduction is one of the important pre-processing steps in high-dimensional data analysis. Typically, a useful representation can make the latent structure in the data more explicit, and often reduces the dimensionality of the data so that further computational methods can be applied. For visualization, the goal of dimensionality reduction is to map a set of observations into a (two or three dimensional) space that preserves as much as possible the intrinsic structure. For classification, the goal is to map the input data into a feature space in which the members from different classes are clearly separated.

Many approaches have been proposed for dimensionality reduction, such as the well-known methods of principal component analysis (PCA) [2], and multi-dimensional scaling (MDS). In PCA, the main idea is to find the projection that

restores the largest possible variance in the original data. In MDS, efforts are taken to find the low dimensional embeddings that best preserve the pair wise distances between the original data points. The two methods are easy to implement. At the same time, their optimizations are well understood and efficient. Because of these advantages, they have been widely used in visualization and classification. PCA and MDS both extract the global feature of data. But sometime the local feature is very important. Recently, a novel method has been proposed to tackle the linear dimensionality reduction problem, namely LPP [1], LPP methods attempt to preserve as well as possible the local neighborhood of each object while trying to obtain highly linear embeddings.

Unfortunately, these methods have a common inherent limitation: they are all unsupervised methods. To solve the problem, some supervised feature extraction methods are proposed, such as Supervised LPP (SLPP) [4] and Local discriminating projection (LDP) [6]. SLPP and LDP seek to find the projection which efficiently preserves the local structure of data points embedded in high-dimensional data space. However, it has the over-learning problem and does not preserve the class and discriminative information of data which is also useful for data recognition. To solve the problem, a novel feature extraction method based on discriminative information, namely supervised discriminative locality preserving projections (SDLPP) is presented in this paper. The SDLPP takes into account class and discriminant information of data points. Different from the most existing supervised locality preserving projection methods, the SDLPP not only preserves both the local structure and diversity information of data, but also avoids the data over-fitting problem. We compare the proposed SDLPP with LPP and SLPP methods on different data sets. Experimental results suggest the efficiency of the proposed method. The remaining of the paper is structured as follows. In Section II, briefly reviews LPP and SLPP. In Section III, we present our SDLPP method. In Section IV, gives some experiments and comparisons, and Section V concludes the paper.

2 LPP and Supervised LPP

Locality Preserving Projection (LPP) [1] is a new algorithm for learning a locality preserving subspace. LPP seeks to preserve the intrinsic geometry of the data and local structure. The objective function of LPP is as follows:

$$\min \sum_{ij} (y_i - y_j)^2 S_{ij} \quad (1)$$

Where y_i is the one-dimensional representation of x_i and the matrix S is a similarity matrix. A possible way of defining S is as follows:

$$S_{ij} = \begin{cases} \exp\left(-\|x_i - x_j\|^2 / t\right), & \|x_i - x_j\|^2 < \varepsilon \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Where \mathcal{E} is sufficiently small, and $\mathcal{E} > 0$. Here, \mathcal{E} defines the radius of the local neighborhood. In other words, \mathcal{E} defines the “locality.” The objective function with our choice of symmetric weights $S_{ij} (S_{ij} = S_{ji})$, incurs a heavy penalty if neighboring points x_i and x_j are mapped far apart, if $(y_i - y_j)^2$ is large. Therefore, minimizing it is an attempt to ensure that, if x_i and x_j are “close”, then y_i and y_j are close as well. Following some simple algebraic steps, we see that

$$\frac{1}{2} \sum_{ij} (y_i - y_j)^2 S_{ij} = w^T X (D - S) X^T w = w^T X L X^T w \tag{3}$$

Where $X = [x_1, x_2, \dots, x_n]$, and D is a diagonal matrix; its entries are column (or row since S is symmetric) sums of S , $D_{ii} = \sum_j S_{ij}$. $L = D - S$ is the Laplacian matrix. If we impose a constraint as follows:

$$\begin{aligned} y^T D Y &= 1 \\ \Rightarrow w^T X D X^T w &= 1 \end{aligned} \tag{4}$$

Finally, the minimization problem reduces to finding:

$$\begin{aligned} \arg \min \quad & w^T X L X^T w \\ & w^T X D X^T w = 1 \end{aligned} \tag{5}$$

The transformation vector w that minimizes the objective function is given by the minimum eigenvalue solution to the generalized eigenvalue problem:

$$X L X^T w = \lambda X D X^T w \tag{6}$$

SLPP[4][5] is a supervised learning method. Its idea comes from Linear Discriminant Analysis (LDA) [3][7]. LDA is a supervised learning algorithm. LDA searches for the project axes on which the data points of different classes are far from each other while requiring data points of the same class to be close to each other. Unlike PCA which encodes information in an orthogonal linear space, LDA encodes discriminating information in a linearly separable space using bases that are not necessarily orthogonal. It is generally believed that algorithms based on LDA are superior to those based on PCA. However, some recent work shows that, when the training data set is small, PCA can outperform LDA, and also that PCA is less sensitive to different training data sets. While PCA seeks directions that are efficient for representation, Linear Discriminant Analysis seeks directions that are efficient for discrimination. Suppose we have a set of n d-dimensional samples x_1, x_2, \dots, x_n , belonging to l classes of faces. The objective function is as follows:



$$\max_w \frac{w^T S_B w}{w^T S_W w} \tag{7}$$

$$S_B = \sum_{i=1}^l n_i (m^{(i)} - m)(m^{(i)} - m)^T \tag{8}$$

$$S_W = \sum_{i=1}^l \left(\sum_{j=1}^{n_i} (x_j^{(i)} - m^{(i)})(x_j^{(i)} - m^{(i)})^T \right) \tag{9}$$

Where m is the total sample mean vector, n_i is the number of samples in the i th class, $m^{(i)}$ is the average vector of the i th class, and $x_j^{(i)}$ is the j th sample in the i th class. We call S_W the within-class scatter matrix and S_B the between-class scatter matrix. He obtains SLPP by connecting LPP and LDA [4][5]. Its objective function is as follows:

$$\min_w \frac{w^T XLX^T w}{w^T XDX^T w} \tag{10}$$

Where $w^T XLX^T w$ is local scatter matrix of the samples. $L = D - S$ is Laplacianface matrix. SLPP seek to find the projection which efficiently preserves the local structure of data points embedded in high-dimensional data space. However, it has the over-learning problem and does not preserve the class and discriminative information of data which is also useful for data recognition.

3 Supervised Discriminative Locality Preserving Projection (SDLPP)

To solve the problem, Here, we extend the cost function given by the LPP technique by proposing a class-dependent approach called supervise Discriminant LPP. We preserve the same constraints on basis as for LPP and we only introduce two more constraints on the coefficients:

1. $S_B = \sum_{i=1}^l n_i (m^{(i)} - m)(m^{(i)} - m)^T \rightarrow \max S_B$ denotes the between-class

scatter matrix and defines the scatter of the class mean around the global mean m . Each cluster formed by the samples that belong to the same class must be as far as possible from the other clusters. Therefore, S_B should be as large as possible.



$$2. S_W = \sum_{i=1}^l (\sum_{j=1}^{n_i} (x_j^{(i)} - m^{(i)})(x_j^{(i)} - m^{(i)})^T) \rightarrow \min S_W \text{ represents the}$$

within-class scatter matrix and defines the scatter of the class samples around their mean. The dispersion of samples that belong to the same class around their corresponding mean should be as small as possible.

We modify the divergence by adding these two more constraints. The new cost function is expressed as:

$$\min \sum_{ij} (y_i - y_j)^2 S_{ij} + \gamma w^T S_B w - \delta w^T S_W w \tag{11}$$

When $\gamma = \delta = 0$, one obtains unsupervised LPP. When $\gamma = 0, \delta = 1$, the result is the fully supervised LPP. Obviously, LPP and SLPP just are specialty of SDLPP. Like LPP, cost function of SDLPP is solved by Computing the eigenvectors and eigenvalues for the generalized eigenvector problem:

$$X(L + \xi I)X^T w = \lambda XDX^T w \tag{12}$$

$\xi = \gamma - \delta$, Where γ and δ are constants. While enhancing the class separability by the minimization of S_W and the maximization of S_B . The SDLPP takes into account class and discriminative information of data points. Different from the most existing supervised locality preserving projection methods, the SDLPP not only preserves both the local structure and diversity information of data, but also avoids the data over-fitting problem.

4 Experiments and Results

In this section, we present a set of experiments to evaluate our proposed algorithms, compared with traditional PCA, LPP and SLPP. PCA and LPP are unsupervised dimensionality methods, SLPP and SDLPP are supervised dimensionality techniques. Where γ and δ are constants.

Intrinsically two dimensional manifold embedded nonlinearly in three dimensional space, as a "swiss-roll", Points nearby in the embedding space may be far apart on the underlying manifold. (Fig.1 shows "swiss-roll" images). In this experiment, we use three PCA, SDLPP to reduce the three dimensionality data to two dimensionalities Fig.1 shows the result of data visualizing. SDLPP preserves the local structure and diversity information of data.

In this experiment, The ORL face database are used, The ORL face database composed of 400 images of size 112x92. There are 40 persons, 10 images for each person. The images were taken at different times, varying lighting slightly, facial expressions (open /closed eyes, smiling/non-smiling) and facial details (glasses/no-glasses). All the images were taken against a dark homogeneous background. Each image was linearly stretched to the full range of pixel values of [0,255].



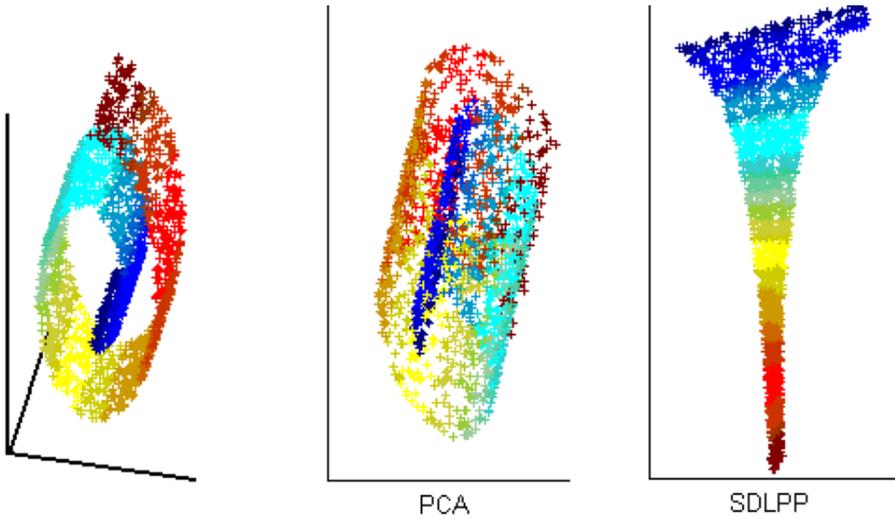


Fig. 1. Swiss-roll images is mapped into a 2-dimensional space. The left panel is a 3-D representation of the set of images. The middle panel shows the results of PCA. The right panel shows the results of SDLPP.

The first five images samples per class are used for training, while the remaining five images are used for testing. We chose a proper dimensionality of the four algorithms as the optimal projective directions then computed the average correct rate with different dimension. The nearest neighborhood classifier (1-NN) is used for classification. We compare the dimensionality purpose of PCA, LPP, SLPP and SDLPP respectively. Table 1 shows the accuracies averaged. Obviously, Experimental results demonstrate that SDLPP is not only good at dimensional reduction, but available to get better performance than conventional LPP and SLPP.

Table 1. The top recognition accuracies of four methods in ORL database (%) (The values of dimensions are same)

| PCA | LPP | SLPP | SDLPP |
|-------|-------|-------|-------|
| 71.44 | 77.80 | 82.60 | 85.92 |

5 Conclusion

In this paper, we presented a feature extraction algorithm by combing LDA and LPP to extract reliable and robust features for recognition. SDLPP is more general than LPP and SLPP. Experimental results on Swiss-roll dataset and the ORL database validated the effectiveness of the proposed algorithms. SDLPP can extract more useful features hide in the original data. There are several issues for future research of SDLPP. First, the selection of γ and δ parameters is crucial for the performances

of SDLPP. In this paper, we only consider to set constant parameter. We will investigate how to adaptively choose parameters in the future, compare SDLPP with other dimensionality reduction methods.

Acknowledgment. Financial supports from China He Nan Department of education Natural Science Foundation Program (NO.2011B110031, NO.2011A110021, NO.2010B520030) are highly appreciated. The helpful comments from some anonymous reviewers are also gratefully acknowledged.

References

1. He, X., Niyogi: Locality preserving projections. In: Proceedings of Advances In Neural Information Processing Systems, vol. 16, p. 153. MIT Press, Cambridge (2004)
2. Kirby, M., Sirovich, L.: Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Trans. Pattern Analysis and Machine Intelligence* 12(1), 103 (1990)
3. Swets, D.L., Weng, J.: Using discriminant eigenfeatures for image retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.* 18(8), 831 (1996)
4. He, X.F., Yan, S.C., Hu, Y., Niyogi, P., Zhang, H.: Face recognition using Laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(3), 328 (2005)
5. Cai, D., He, X., Han, J.: Document Clustering Using Locality Preserving Indexing. *IEEE Transactions on Knowledge and Data Engineering* 17(12), 1624 (2005)
6. Zhang, H., Sun, S., Jing, Z., Yang, J.Y.: Local structure based supervised feature extraction. *Pattern Recognition* 39(8), 1546 (2006)
7. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. *IEEE Trans. Pattern Analysis and Machine Intelligence* 19(7), 711 (1997)
8. Zhao, H., Sun, S.: Local structure based supervised feature extraction. *Pattern Recognition* 11(39), 1546 (2006)
9. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. *Science* 290(550), 2323 (2000)
10. Belkin, M., Niyogi, P.: Laplacian eigennlaps for dimensionality reduction and data representation. *Neural Computation* 15(6), 1373 (2003)

Theory and Application of Monte Carlo Method

Liu Xinhua, Liu Yongzhi, and Liu Hao

School of Mechanical and Electrical Engineering,
China University of Mining and Technology, XuZhou, China
Liuliu2011@hotmail.com

Abstract. The research progress and results of Monte Carlo method for the engineering technology fields are reviewed. Firstly, the basic ideas and principles of Monte Carlo method are briefly introduced. Secondly, the simulation process and procedure of Monte Carlo method are presented. Thirdly, the applications in several fields such as particle transport problems, inter-particle collision, mathematical, statistical mechanics calculation and other engineering applications are summarized. Finally, the direction of future development for Monte Carlo method was discussed.

Keywords: Monte Carlo method, Random problem, Numerical simulation.

1 Introduction

With the development of computer science and technology, so many problems and systems are more and more complicated. Many of these problems and systems are changed, random, high dimensional and nonlinear, and presents the characteristics of fuzzy and changed. It is difficult to build the mathematic model for this kind of problems and systems, except for some specific problems. Another the problem is calculated accurately that is impossible. Therefore, it is of great importance to study the method of solving random problem for solve complex system quickly.

The problem is solved by using Monte Carlo method's sampling method, which will become discrete values, and simplify the computational complexity greatly. This method was proposed by the von Neumann etc and was used to simulate the atomic fission processes. However, the basic ideas of Monte Carlo method were originated from Buffon's cast needle problems, which cast needle method to calculate the value of PI. But this method needs to do a lot of experiments. The method to obtain the accurate results, the greater workload must be done. So the application is not widespread. With the computer science and technology development, the computer could replace people do a great deal of randomized trials based on mathematical model. So this method has become an important method for research random problems. The Monte Carlo method are widely applied in physics, medicine, materials, machinery, management finance, insurance, biology, medicine, economic and etc other fields [1].

The basic idea of Monte Carlo, the step of simulation and the application of Monte Carlo were reviewed in this paper. The analysis compares these to study positive result, aim to promote the development that the method of theory and applied research.

2 The Basic Principle and Characteristics of Monte Carlo Method

The basic idea of Monte Carlo method is according to stay for random question or physical phenomena statistical law itself constructs a suitable probability model. According to the model makes lots of statistical testing, make some of its statistical parameter is the solving problem of solution. The basic process of Monte Carlo method involves structural or describe probability process and establishing various estimator expectations that are the problem of results. Then according to its probability model of calculation program are compiled [2].

Monte Carlo method has the following advantages: Firstly, it can be true reaction with random nature physical process, because the methods to solve practical problems directly from the problem itself are extracted, rather than from sample mathematical theory model. Secondly, its geometry condition the influence is weak; solving complex problems can only when the enough discrete sample points can work out approximate the exact solution. Thirdly, its algorithm convergence speed and problems of the dimension is irrelevant. Its convergence speed is constant and dimension is independence. The extraction sample size is independent of the dimension. The dimension high of problem only increase calculation but affect precision does not. So the Monte Carlo method to solve the 3d and high dimensional problem has advantages, as is not feasible with tradition methods. Fourthly, it can also calculate multiple problems or a problem of multiple states, and will not increase amount of calculation. Fifthly, its error can be calculated according to the error formula. Sixthly, its algorithm programming is simple and easy to running. Another Monte Carlo method compared with other numerical method has some shortcomings, such as the slow convergence speed and the error of Monte Carlo is a random and cannot obtain high-precision results [3,4].

3 The Steps of Monte Carlo method

The question is given by Monte Carlo simulation and the process of solving includes five steps as follow:

Step 1: Question analysis. The step has provided a preparation for the next step based accurate mathematical model by using a large number of data, the factors affecting the problem extracted and for each influential factor were analyzed, determined its weight coefficient.

Step 2: Mathematical modeling. The mathematical model of random questions is set up according to the step one determines the influencing factors and the weight coefficient. As for the problem does not belong to random problem, there need construct a man-made probability process and turn it into random problem.

Step 3: computer simulations. General engineering applications based on simulation Monte Carlo method are shown as follows: the system got the estimate of observable amount through the sampling simulation.

$$\langle I \rangle = \frac{\int_{\Omega} I(x) f(H(x)) dx}{\int_{\Omega} f(H(x)) dx}$$

$$\langle I \rangle = \frac{\sum_i^n I(x_i) p^{-1}(x_i) f(H(x_i))}{\sum_i^n p^{-1}(x_i) f(H(x_i))}$$

Where: $H(x)$ is a Hamiltonian of the system. $f(H(x))$ is distribution function. Monte Carlo method is the summation by selecting random state calculated the above formula. So there draw a large state will get the accurate results of simulation. If $P(x)$ is balanced distribution, so $\langle I \rangle$ is equivalent to average value.

Metropolis proposes a kind of algorithm to solve the above equation. The flow chart of Metropolis algorithm is shown as follows:

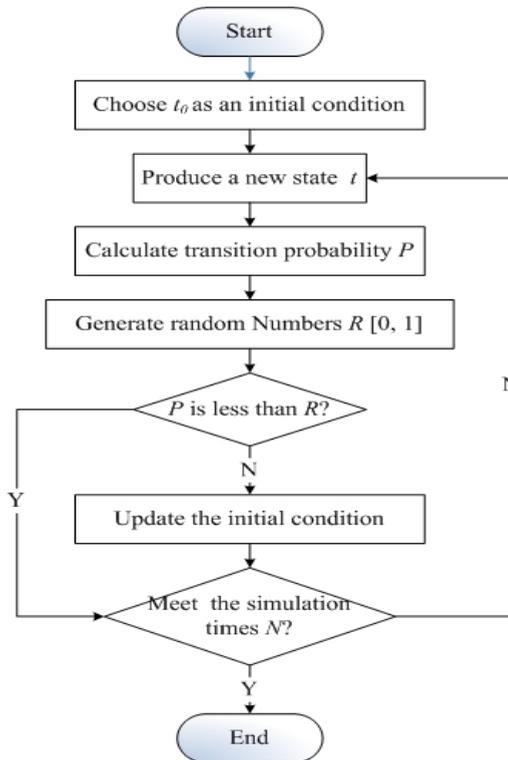


Fig. 1. The flow chart of Metropolis algorithm

Usually, as the key factors of Monte Carlo method, random number and sampling method decided whether the problems can be solved with high quality.

(1) Random number is random variables obtained value, which have three types of random number, such as True Random Number, Quasi Random Number, and Pseudo-random Numbers. But the Monte Carlo method is usually by using Pseudo-random Numbers. The selection of random number for simulation results has important influence. The two necessary characteristics of random number are uniformity and independence. But it often exist two problems are not independent and repetition. So the uniformity and independence must be examined. There must be selected a fit pseudo random recursive formula to solve these two problems, such as the following two kinds of methods are middle-square method and Congruence method. But Congruence method is better than the former method and can produce a good random numbers. Recursive formula is shown as follows:

$$\begin{cases} x_{n+1} \equiv \lambda x_n + C \pmod{M} \\ r_n = x_n / M \end{cases}$$

Where: λ is multiplier, c is increment M is modulus.

(2) Sampling method. The sampling method of Monte Carlo can be summarized as following: direct sampling method, transform sampling method, acceptance-rejection method sampling and Composite sampling method etc.

Step 4: The analysis of simulation result, which according to the simulation data of step 3.

Step 5: Finally, the improvement of the Monte Carlo algorithm.

4 The Application of Monte Carlo Method

The application of Monte Carlo method in the physics. (1) Particle transport problems. Monte Carlo method is considered to solve this kind of problem as one of the best numerical methods. Here, simulation of distribution sampling, particle state (space position, energy and movement direction), and the result analysis that needs to be resolved. So we can think of the particle transport problem is a 3d or high dimensional problem, which by using Monte Carlo method is simpler than other method. It also can simulate the each state parameter of particles [5-7]. One of the advantages of using the Monte Carlo method to solve particle transport problems is the fact that it does not rely on any theoretical limitations imposed by geometry or material composition, offering the possibility to obtain solutions for various combinations of materials and volume shapes [8]. In this respect, Deng Li who has done some research such as Monte Carlo method can simulate various complex geometric inside the system neutrons, photon, electronics, alpha particle, protons and coupling transport problems, and get the right result, they think MC method has become the preferred method of particle problem solving [9]. Such Monte-Carlo transport simulation can compute the X-ray photon emission given the incident electrons [12]. Xu Hai-yan also put forward some improvement Monte Carlo method Global Variance Reduction Method for Monte Carlo Particle Transport Problems [11].

(2) **Particle collisions.** Particle collisions can be considered as a nano or micro scale of problem. The methods of solving these problems were also suggested such as Monte Carlo method, molecular dynamics and Quantum mechanics. Here, the micro scale of particle collisions of three simulation methods is compared. For example, the Particle collisions are simulated by molecular dynamics, the precision of results is high, but much calculation is needed and it's complex. There is great gap between Quantum mechanics of model and the true model. The Monte Carlo method is not much accurate than the first method, but it can get myopia lifelike results after the enough simulation times. The Monte Carlo method for simulating those problems is effective. [12-14]. In this aspect, Chih-Hung Hsu with the direct simulation Monte-Carlo method for inter-particle collisions in turbulent flow [15] and S Menear, the size distribution of MRF particle has been simulated by using Monte Carlo method. He got the results of simulation is in line with the experimental results [16]. Yang Shi Qing had used to the Monte Carlo method simulates particles two-dimensional position and turns into chain process in additional magnetic field MRF and also be discussed in this field [17].

The application of Monte Carlo method in the mathematical and statistical mechanics. The mathematical some complex calculation process can be solved, such as some complex integral, multidimensional mathematical calculation, matrix calculation, the solution of differential equation and some linear equations. Some complex mechanics equations are calculated by using Monte Carlo in Statistical mechanics, such as Langevin equation and Planck equation [18].

The application of Monte Carlo method in economics. The stock market fluctuations and market price, the economic risk forecasting and other economic problems can be simulated. Such Phelim Boyle discusses some of the recent applications of the Monte Carlo method to security pricing problems, with emphasis on improvements in efficiency. The Monte Carlo approach has proved to be a valuable and flexible computational tool in modern finance [19]. MA Jun-hai research and explore the problem of pricing warrants by using Monte Carlo simulation methods and its improved technique, and set up Monte Carlo models pricing warrants. Draw MC method is to solve warrants for pricing an effective way [20]. Dong Li Juan using Monte Carlo method to calculate the stock index futures market VaR value and measure index futures market risk [21].

The application of Monte Carlo method in medicine. The Monte Carlo method can be calculated the radiation dose. Such as, the cancer and heart disease that it's the dosage and plan of drug can be simulated by using it [22,23].

The application of Monte Carlo method in other engineering field. (1) Location problems. The Monte Carlo method uses a great random sample to obtain the optimal position estimate, greatly improving the positioning accuracy of the system. Such as, Huang Ke Ping got through establishing the positioning of linear equations by using Monte Carlo simulation to achieve expected accuracy. The simulation experiment proves the method is feasible and effective. [24]. Shao Qing-liang researches show that the mobility of the nodes under coalmine may cause common localization algorithms inaccurate, in light of this, Monte Carlo localization algorithm is proposed. Simulation

results show that the location estimation error of this method was significantly lower than other methods in low-density anchor nodes environment. This method improved the accuracy of mobile nodes location algorithm [25].

(2) The problems of actuarial science. The mortality rate smoothing in life insurance, which is related to the mortality in premiums calculating can be simulated by Monte Carlo method. The results obtained in the article are compared with maximum likelihood and Bayesian methods. The result which is smoothed by Monte Carlo method combined with the conjugate gradient method is better than the other two methods and is more fitting the sample data. Thus proves the effectiveness of solving such problems by Monte Carlo method [26].

(3) Reliability problems. Predicting reliability products can avoid some security incidents and also relate to enterprise economy. The product of reliability during a certain period in the future was predicted by using the Monte Carlo method. The method effectively predicted that the product's reliability declines, and provide the basis for maintenance [27]. In addition Fan Hai-yan has used it calculated the bus lines running time of reliability [28].

5 Conclusions

In recent years, the Monte Carlo method as a kind of a new numerical method which rose with the development of computer technology. The method is not only used in physics and mathematics, but also more and more widely used in other other fields. Because some complex high dimensional problem is solve by using it that is simple and easy calculation. The method and other methods is compared, it has an obvious advantage. It as a new numerical calculation method compared with other classic numerical methods, also has its limitations that the problem of slow convergence and error randomness. But this does not hide in dealing with random problem, it has distinct advantage. In order to overcome these drawbacks, two methods are mainly employed: Quasi-Monte Carlo simulation and Dynamic Monte Carlo method. With the development of computer technology and the development of research, the Monte Carlo method will have more and more application and prospects.

Acknowledgements. The support of National Natural Science Foundation of China (No. 51005231), National Science Foundation for Post-doctoral Scientists of China (No. 20100471408) in carrying out this research is gratefully acknowledged.

References

1. Xu, Z.-J.: Monte-Carlo simulation. Shanghai Science and Technology Press, China (1985)
2. Ji, Q.-F.: Monte Carlo method and its application in hydraulics. Southeast University press, China (2004)
3. Li, Q.-C.: Monte Carlo Simulation for the Growth of Magnetron Sputtering Cu Films. Hefei University of Technology, China (2009)
4. Papadopoulos, C.E., Yeung, H.: Uncertainty estimation and Monte Carlo simulation method. Flow Measurement and Instrumentation 12, 291–298 (2001)

5. Zhu, X.-L.: Markov chain Monte Carlo method research reviewed. *Statistics and Decision* 21, 151–153 (2009)
6. Wang, Z.-L., Ma, X.-W., Yang, W.: Monte Carlo simulation of time-of-flight spectrum produced in ion-atom collisions. *Chinese Journal of Atomic and Molecular Physics* 20, 322–326 (2003)
7. Huang, C.-Q., Deng, S.-J.: Monte Carlo Simulation of 3-dimensional Ising Model. *Chinese Journal of Computational Physics* 26, 937–941 (2009)
8. Popescu, L.M.: A geometry modeling system for ray tracing or particle transport Monte Carlo simulation. *Computer Physics Communications* 150, 21–30 (2003)
9. Li, D., Gang, L.: A summarization on Monte Carlo simulation in particle transport. *Chinese Journal of Computational Physics* 27, 791–798 (2010)
10. Bowers, K.J., Devolder, B.G., Yin, L., Kwan, T.J.T.: A maximum likelihood method for linking particle-in-cell and Monte-Carlo transport simulations. *Computer Physics Communications* 64, 311–317 (2004)
11. Xu, H.-Y., Huang, Z.-F., Cai, S.-H.: Global Variance Reduction Method for Monte Carlo Particle Transport Problemes. *Chinese Journal of Computational Physics* 27, 722–731 (2010)
12. Tao, Y.-L., Zhu, J.-H., Wang, J.-G., Niu, S.-L., Fan, R.-Y.: Monte Carlo simulation of Gamma Ray transport in non-uniform atmosphere. *Chinese Journal of Computational Physics* 27, 740–744 (2010)
13. Qing, G.: The study and application of Kinetic Monte Carlo Method in multi-scale simulation. Chongqing University, China (2010)
14. Zhao, H.-B., Zhen, C.-G., Chen, Y.-M.: Multi-Monte Carlo method for inter-particle Collision. *Chinese Journal of Theoretical and Applied Mechanics* 37, 564–571 (2005)
15. Hsu, C.-H., Keh, C.-C.: A Lagrangian modeling approach with the direct simulation Monte-Carlo method for inter-particle collisions in turbulent flow. *Advanced Powder Technol.* 18, 395–426 (2007)
16. Menear, S., Bradbury, A., Chantrell, R.W.: A model of the properties of colloidal dispersions of weakly interacting fine ferromagnetic particles. *Journal of Magnetism and Magnetic Materials* 43, 166–176 (1984)
17. Yang, S.-Q., Gong, J., Zhang, W.L., Wang, R.-F., Wang, H.-C.: Monte Carlo simulations to influence of magnetic field strength on aggregate structure of ferromagnetic particles in Magneto rheological fluids. *Chinese Journal of Atomic and Molecular Physics* 15, 411–415 (1998)
18. Wu, H.-X., Liu, L.: Monte Carlo method and applications in solving special problems. *Journal of Taiyuan Normal University* 8, 76–79 (2009)
19. Boyle, P., Broadie, M., Glasserman, P.: Monte Carlo methods for security pricing. *Journal of Economic Dynamics and Contro.* 121, 1267–1321 (1997)
20. Ma, J.-H., Wang, Y.: Monte Carlo simulation methods and its improved technique for pricing warrants. *Journal of Industrial Engineering/Engineering Management* 24, 75–81 (2010)
21. Dong, L.-J.: The model of stock index future margin based on VaR of Monte Carlo simulation. Inner Mongolia University, China (2008)
22. Jiang, Y.-Q., Ma, J.: Study on transmission properties in gastric tissue with Monte Carlo simulation method. *Periodical of Ocean University of China* 39, 345–348 (2009)
23. Xu, S.-Y., Liu, B.-J., Li, Q.: Monte Carlo computation in the applied research of nuclear technology. *Nuclear Techniques* 30, 597–600 (2007)
24. Huang, K.-P., Ying, H., Huang, T., Zhang, Y.: Application of Monte Carlo algorithm to indoor 3D real-time location. *Metrology & Measurement Technique* 35, 1–2 (2008)

25. Shao, Q.-L., Zhang, Z.-C., Zhang, Z.-X.: On localization of coal miner based on Monte Carlo algorithm. *Computer Applications and Software* 26, 165–166 (2009)
26. Na, L.: The application of Monte Carlo method in actuarial. Chongqing University of Technology, China (2010)
27. Xu, Z.-G., Zhou, D.H.: Real-time reliability prediction method based on Markov chain Monte Carlo. *Journal of Mechanical Strength* 29, 765–768 (2007)
28. Fan, H.-Y., Yang, X.-G., Yan, L., Xia, X.-M.: Application of Monte Carlo method in the reliability calculation of bus travel time. *University of Shanghai for Science and Technology* 28, 59–62 (2006)

A New XML Database System Using Pagination Technique

Xiaopeng Wang

Henan University of Traditional Chinese Medicine, Zhengzhou China, 450008
wang_d1882@sohu.com

Abstract. This paper has proposed an approach to store, retrieve and query large XML documents efficiently by introducing pagination technology and a pagination-based index structure to a native XML database system – CanStoreX. The experimental results have shown that the pagination-based index structure can greatly improve the performance of the CanStoreX XQuery engine.

Keywords: XML database, pagination technique, index structure.

1 Introduction

Today, with the widespread of XML format data, many new challenges have been brought out to researchers in databases. It needs to be recognized that XML data is fundamentally different from the traditional relational data. Its unique characteristics make it better suited for specific types of data that are hard to manage with a relational paradigm. These include heterogeneous and sparse datasets. XML data is also symmetric in that data and meta-data are included in the same container.

XQuery language [3] is designed to query collections of XML data and has become a W3C standard for years. Currently, most existing XQuery engine implementations depend on DOM [1] or JDOM [4] to access XML documents. DOM and JDOM APIs, however, require the whole document to be loaded into the memory as a fully expressed tree. Memory limitations have thus become a major issue for XQuery engines using this strategy. For example, in Saxon [4] it will generally take around 1GB memory space to process a 200MB XML file. Therefore, an efficient approach is needed to store, retrieve, query and update large XML datasets.

2 Canstorex System Architecture

The overall architecture of CanStoreX native XML database system, as shown in Figure 1, consists of four layers from the bottom to the top: storage management, buffer management, CanStoreX DOM API and XQuery engine. Besides, a loading engine is integrated to import XML documents into CanStoreX storage, so that they can be accessed and queried by CanStoreX DOM API and XQuery engine. The functions of each component are explained in the following sections.

Storage and Buffer Management. In order to use CanStoreX, a user needs to configure CanStoreX storage on one’s local machine by providing a location and a size for the storage. Then the low-level device files will be created in the specified location and can be considered as the raw disk space reserved for CanStoreX storage. Storage manager supports the concept of a page as a unit of data in the storage and provides commands to allocate, deallocate, read or write a page. Thus, from the clients’ points of view, CanStoreX storage is just a sequence of fixed-size pages, which are managed by the storage manager as Figure 1 shows. Ideally, the size of a page is chosen to be a multiple of the disk block size, and pages are stored as adjacent disk blocks so that reading or writing a page can be done in minimal disk I/O operations. Currently, the size of a page in CanStoreX storage can vary from 1KB to 64KB.

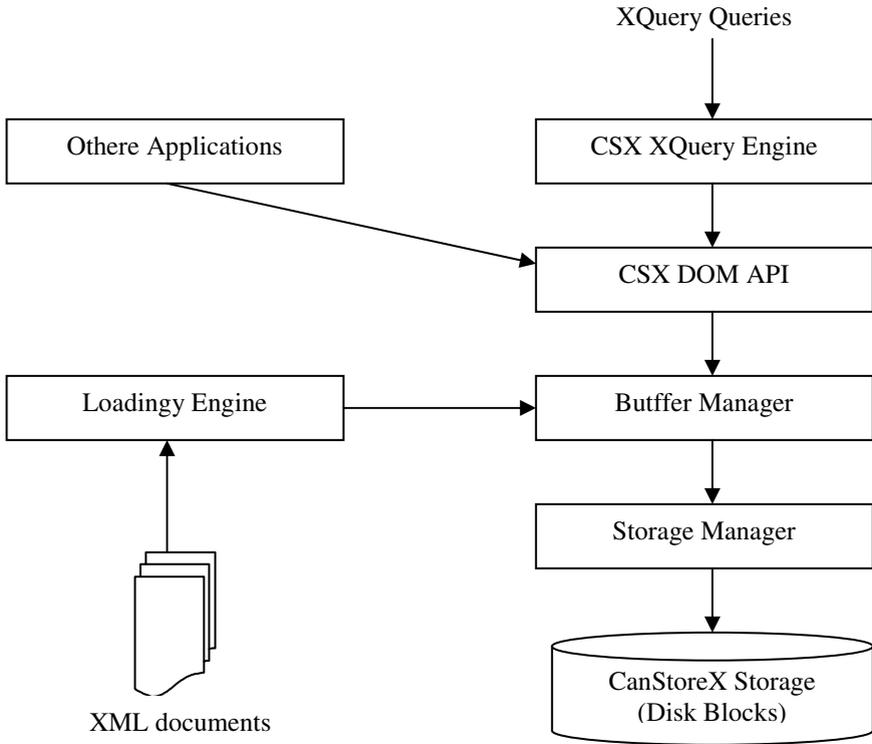


Fig. 1. CanStoreX system architecture

The role of the buffer manager is like a bridge between CanStoreX storage and its applications (such as CanStoreX DOM API). It is responsible for bringing pages from disk to memory and from memory back to disk. The buffer manager manages the available memory space as a collection of pages in its buffer pool. The size of the buffer manager, in terms of the number of pages, can be specified by users as they configure CanStoreX storage. Having its own buffer manager, CanStoreX can strictly control the size of the memory space it is using in all stages according to the users’ preferences,



and can have the freedom of managing pages in the buffer manager in finest granularity. When a page fault occurs, a decision needs to be made which existing page in the buffer pool is to be replaced to make space for the new page. This is implemented through replacement policies. A detailed discussion of different replacement policies in CanStoreX exceeds the scope of this paper.

Pagination Algorithm. The loading engine loads and maps an original XML document as a BXML document in CanStoreX storage using a pagination algorithm. The pagination algorithm partitions the original XML document into binary pages and a BXML file is used to represent the paginated XML document in CanStoreX storage.

After the pagination, the complete tree structure of the XML document is still kept intact in the BXML document. The only difference between the original tree structure and the tree structure of the BXML file is that some extra nodes are added to facilitate pagination, which are inserted by the pagination algorithm.

To demonstrate the idea of the pagination algorithm, a small XML example containing 8 nodes is shown in Figure 2. In order to make the example more interesting, we assume the sizes of the leaf nodes C, D, E and G are large enough so that we have to allocate a separate page for each of them, and the size of the root node A is also large so that it has to be put into one separate page. Then its paginated BXML representation in CanStoreX storage is shown in Figure 3. In Figure 3, a rootFony node is represented as a triangle and a childFony node is represented as a square with a number inside it that indicates the child page ID it points to.

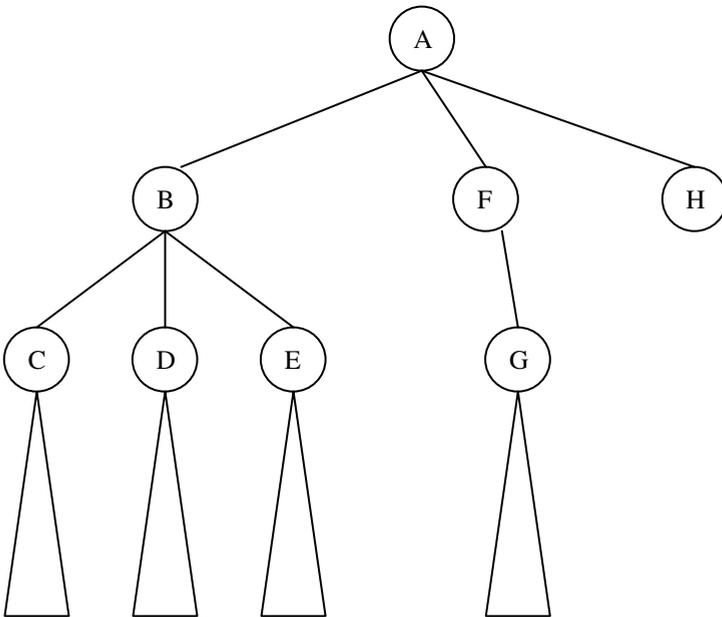


Fig. 2. Example of the tree structure of an original XML document

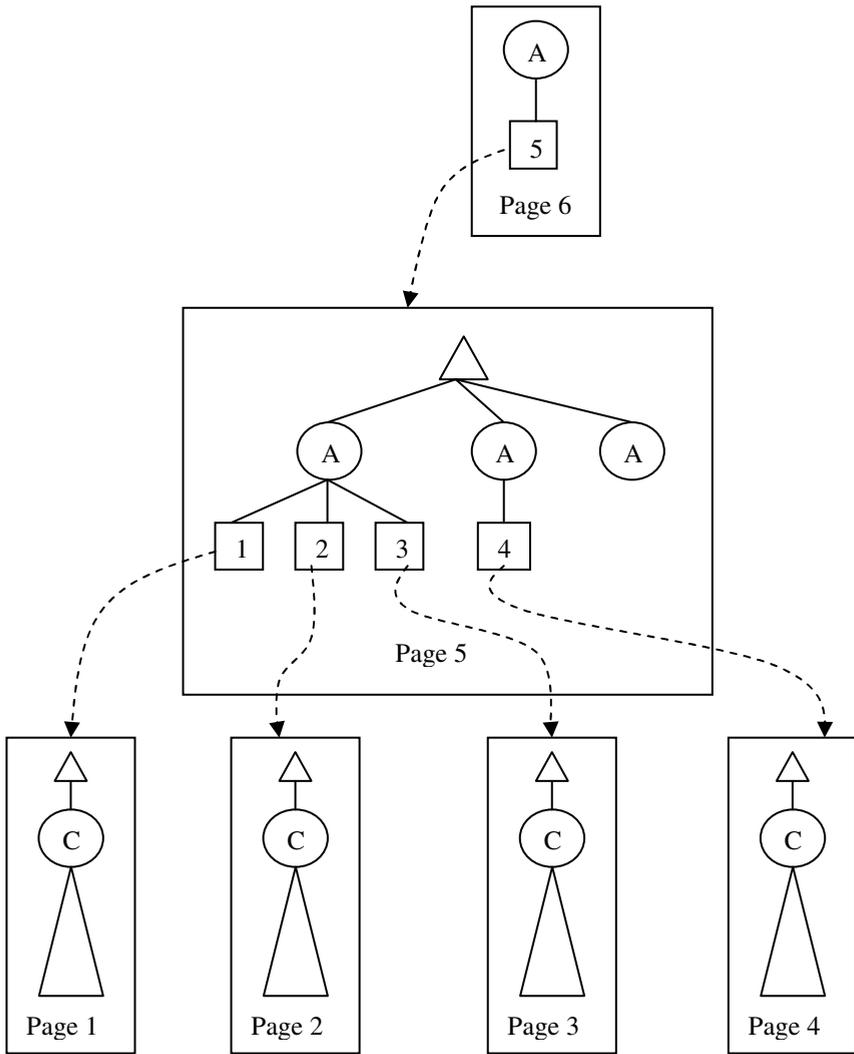


Fig. 3. Example of a paginated BXML document

The basic pagination algorithm in CanStoreX is implemented by Ma and Patanroi . SAX parser has been used to perform a depth-first traversal of the original XML tree and construct a BXML document in CanStoreX storage using a dynamic bottom-up strategy. Some heuristics have been done to improve the quality of the pagination algorithm following that. Because of the limited space of this paper, the details of the pagination algorithm and its optimizations are not discussed here.

3 Bxml Page Hierarchy and Index Structure

Since disk I/O operations are very expensive compared to in-memory computations, the number of disk page accesses always has dominant effect on the performance of a database query engine and needs to be optimized whenever possible. Because of the unique pagination technology utilized in CanStoreX, it is possible for CanStoreX XQuery engine to improve its performance by decreasing the number of disk page accesses.

Therefore, our focus in this section is to try to reduce the number of page accesses required by XQuery engine as much as possible. Several different strategies to embed information into our storage-facilitating nodes have been tested, until we find an effective way to build the index structure based on the page hierarchy of a BXML document in CanStoreX storage.

The Page Hierarchy of a BXML Document. As has been explained, each page in CanStoreX storage has its own tree structure inside the page starting from a rootFony node, thus each page can be considered as a self-contained XML node on its own right, and the tree structure consisting of these page nodes then forms the page hierarchy of the BXML document. The page hierarchy can be represented perfectly as an XML document because of its intrinsic XML tree structure, and is a good way to check how well the pagination algorithm has done in partitioning the original XML tree into pages.

For example, the page hierarchy of the BXML document shown in Figure 3 can be represented in XML in Figure 4.

```

<RootPage id="6">
  <Page id="5">
    <Page id="1"/>
    <Page id="2"/>
    <Page id="3"/>
    <Page id="4"/>
  </Page>
</RootPage>

```

Fig. 4. Example of the page hierarchy of a BXML document

The page hierarchy of a real BXML document could have many more nodes than those in the example of Figure 4. Each parent page node may have tens of or hundreds of child nodes, depending on the page size and the characteristic of the XML document. Nonetheless, compared with the original XML document, the page hierarchy tree structure is really simple and much smaller in size. Though simple as it seems, the page hierarchy contains very useful information to the XQuery engine. It serves as the skeleton of the original XML tree structure, and each node in the page hierarchy structure corresponds to a page in CanStoreX storage. If configured properly, the page hierarchy can help the XQuery engine save the number of page accesses greatly.

```

<RootPage id="6" xpath="//*">
  index:item[1,7];person;
  <Page id="5" xpath="/A/*">
    index:item[1,7];person;
    <Page id="1" xpath="/A/B/*">index:item[1,4];</Page>
    <Page id="2" xpath="/A/B/*">index:item[5,7];</Page>
    <Page id="3" xpath="/A/B/*"></Page>
    <Page id="4" xpath="/A/F/*">index:person;</Page>
  </Page>
</RootPage>

```

Fig. 5. Example of the index of a BXML document

4 Experimental Results

To test the effectiveness of the pagination-based index structure, we use XMark Benchmark [2] to generate XML documents of 4 different sizes: 1MB, 10MB, 100MB, 1GB, and load them into CanStoreX storage. An index is built for each of the files, with the XPath expression and the indexed element "item". Three queries are selected from the 20 XMark standard queries and are executed on each document with and without indexes.

The page size of CanStoreX storage is set to 16KB, and the size of the buffer pool is set to 100 pages. In order to avoid the influence of cached pages in the buffer pool, we empty the buffer pool at the end of each query.

In the experiments, the number of page accesses is used to evaluate the performance of CanStoreX XQuery engine, which is a reasonable measurement, since disk I/O operations have become the most expensive operations nowadays.

The three queries selected from XMark are as follows:

Q1: for \$b in document("xmark1G.xml")/regions
/africa/item[/@id="item166"]

return <E1> {\$b/name} </E1>;

Q2: for \$i in document("xmark1G.xml")/regions
/australia/item

return <item> {
 <N>{\$i/name/text()}</N>,
 <D>{\$i/description}</D>
} </item>;

Q3: for \$b in document("xmark1G.xml")

return <E> {count(\$b//item)} </E>;

The number of page accesses spent by each query with or without index on each of the four different files sizes is shown in Table 1.

As can be seen, for Query 1, the number of page accesses for an exact match is always a constant number with the use of index. This can be easily understood, since the XQuery engine can directly access the pages that contain the indexed element and whose id range covers the id value in the query, with the help of the index. The reason

we need 3 page accesses instead of 1 page access in this case, is that all the ancestor page nodes of the destination page node in the index may also be considered as possible candidates and will be searched by the XQuery engine.

For Query 2 and 3, it can be observed that the number of page accesses is almost reduced by half using the index. This is because in the XML documents generated by XMark generator, the size of all the “item” elements is almost half of the size of the whole XML document. Thus, in the BXML document in CanStoreX storage, nearly half of its pages contain “item” elements. (The “item” elements are stored consecutively in the original XML document.)

In all the above cases, the index structure in CanStoreX has reduced the number of page accesses used in XQuery engine dramatically. Actually, based on the paginated structure of a BXML file in CanStoreX storage, the data we have seen in Table 1 for the index-based XQuery engine is the best one can do in the current architecture. With the power of the index structure, the index-based XQuery engine can directly access the pages that contain the results of the query, and only spend minimal or no page accesses in searching other pages.

The size of the index is pretty small compared with the original XML document. According to our experiments, the size of the index file is proportional to the size of the original XML file. For a 100MB XML file, its index size is just 515KB, and for a 1GB XML document, the index size is around 5MB.

Table 1. Numbers of Page Accesses

| | | Query 1 | Query 2 | Query 3 |
|-------|------------|---------|---------|---------|
| 1MB | With index | 3 | 5 | 39 |
| | No index | 6 | 10 | 78 |
| 10MB | With index | 3 | 45 | 395 |
| | No index | 18 | 52 | 792 |
| 100MB | With index | 3 | 390 | 3854 |
| | No index | 110 | 786 | 7773 |
| 1GB | With index | 3 | 3902 | 38566 |
| | No index | 995 | 7819 | 77955 |

Therefore, in the current implementation, index files are stored as XML files outside CanStoreX storage and will be loaded into the memory completely to get best performance. In future, as the size of the XML document increases, we can consider importing the index file itself as a BXML file and accessing it through CanStoreX DOM API from the storage.

5 Conclusion

This paper has proposed an approach to store, retrieve and query large XML documents efficiently by introducing pagination technology and a pagination-based index structure to a native XML database system – CanStoreX. The experimental results have shown that the pagination-based index structure can greatly improve the performance of the CanStoreX XQuery engine.

It needs to be noticed that the current CanStoreX system is only a prototype system. To exhibit the real power of this pagination and indexing concept, some low-level issues need to be solved, such as the mapping between pages in CanStoreX storage and real disk pages, and a mechanism to access the raw disk pages efficiently. These will be considered in our future work.

References

1. Le Hors, A., Le Hegaret, P., Wood, L., Nicol, G., Robie, J., Champion, M., Byrne, S.: Document Object Model (DOM) level 3 core specification. Technical Report, W3C recommendation (2008)
2. Schmidt, A., Waas, F., Kersten, M.L., Florescu, D., Manolescu, I., Carey, M.J., Busse, R.: Assessing XML Data Management with XMark. In: Bressan, S., Chaudhri, A.B., Li Lee, M., Yu, J.X., Lacroix, Z. (eds.) CAiSE 2002 and VLDB 2002. LNCS, vol. 2590, pp. 144–145. Springer, Heidelberg (2003)
3. Chamberlin, D.: XQuery: a query language for XML. In: Proc. SIGMOD Conf., p. 682 (2003)
4. JDOM, <http://www.jdom.org>

Design and Implementation of Multimedia Format Converter Based on FFmpeg

Yun Cheng^{1,2}, Qingtang Liu^{1,*}, Chengling Zhao¹, Xiaoliang Zhu¹,
and Guoqing Zhang¹

¹ The Department of Information Technology
Central China Normal University, WuHan, Hubei, China, 430079

² Department of Educational Science and Technology
HuangGang Normal University, HuangGang, Hubei, China, 438000
liuengineering@TOM.COM

Abstract. FFmpeg is free software licensed under the LGPL or GPL. It is a complete, cross-platform solution to record, convert and stream audio and video. But it's complex and inefficient to use, for it doesn't have a GUI (Graphical User Interface). We develop a multimedia format converter with GUI based on FFmpeg, which is designed and implemented based on Qt framework, under Linux environment. It is more efficient and convenient for users than ffmpeg.exe. Its GUI is concise, and easy to use and update. Users can use it to convert the file from the source format to any other format they want without complex settings.

Keywords: FFmpeg, format conversion, GUI, converter, Qt.

1 Introduction

FFmpeg is a complete, cross-platform solution to record, convert and stream audio and video. It includes libavcodec - the leading audio/video codec library. FFmpeg is free software licensed under the LGPL or GPL depending on your choice of configuration options. FFmpeg is developed under Linux environment, but it can be compiled under most operating systems, including Windows. FFmpeg is supported by more than ten transfer protocols, such as HTTP, RTP, RTSP, TCP, UDP, MMS, etc, and also supported by more than forty audio formats and more than ninety video formats. The high ability of coding/decoding makes it widely used in multimedia area. [1]

2 The FFmpeg Project

The FFmpeg project provides three tools: ffmpeg, ffserver and ffplay. ffmpeg is a command line tool to convert multimedia files between formats. ffserver is a multimedia streaming server for live broadcasts. ffplay is a simple media player based on SDL(Simple DirectMedia Layer) and the FFmpeg libraries.[2]

* Corresponding author.

FFmpeg is developed under Linux, and is executed by compiling of GCC (GNU Compiler Collection) through the command line. It doesn't have a GUI (Graph User Interface), so it's complex for most windows users.

3 The Transcoding Progress of Ffmpeg

Format conversion progress of ffmpeg. Format conversion is a progress of decoding from source format, then encoding to the destination format, as described in the following procedure.

Register all the formats and codec in main function; find the suitable demuxer and muxer for the multimedia file.

The following work will be done by the function of av_transcode(it is the older version of av_encode). Av_transcode, as the core function of ffmpeg.c, completes most of the work of format conversion. It calls the function of output_packet to decode, encode, and write output files. The progress is as follows, shown in Fig.1.

- 1) Input streams initializing.
- 2) Output streams initializing.
- 3) Encoders and decoders initializing.
- 4) Set Meta data information from input file if required.
- 5) Write output files header.
- 6) Loop of handling each frame.
- 7) Write output files trailer.
- 8) Close each encoder and decoder.

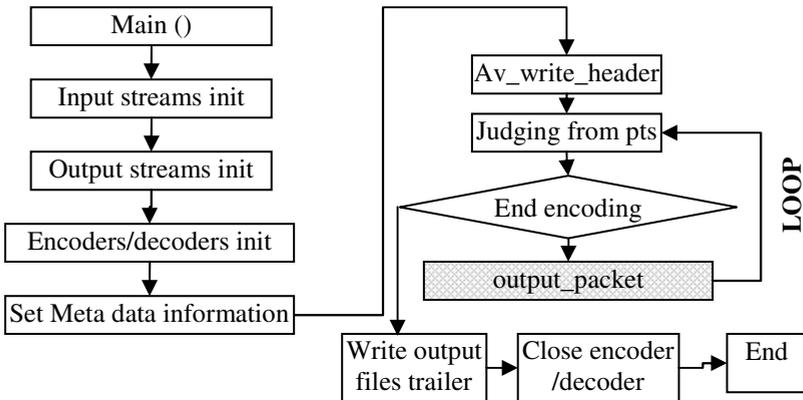


Fig. 1. Format conversion progress of ffmpeg

The details of format conversion by output_packet. Av_transcode completes the work of format conversion based on output_packet. We will explain the details that how the function works.

- 1) Read frame from input file and we will get a packet.



- 2) Call the function of `avcodec_decode_video` to decode the packet if needed, then we will get a frame data.
- 3) If no encoding needed, go to 6), output the packet directly; else, call the function of `do_video_out` to encoding.
- 4) If the video picture is raw picture, no encoding needed, pack the frame data to packet directly; else, encode new frame data, then pack the frame data to packet.
- 5) Init the packet.
- 6) Call the function of `write_frame` to output the packet.

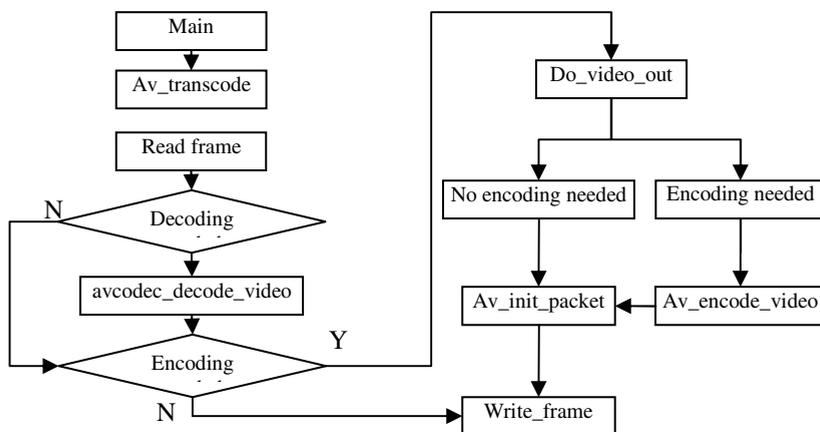


Fig. 2. The details that how the function of `output_packet` works

As we know, a video contains video pictures and audio, which are processed separately in `output_packet`. We have discussed the video pictures processing above, the processing of audio works in the same way.

4 The Design of GUI for the Converter

The user interface acts as a bridge between user and the software of FFmpeg. Passing parameters by user interface is more convenient and efficient.

Principle of design. The main characters of the GUI of the converter are as follows:

Its interface is concise and beautiful. The purpose we design the GUI is to make it more convenient and efficient for users, so the GUI should be concise and beautiful, and easy to use.

The command line parameters of FFmpeg are complex, we needn't to show all the parameters for most of them are not commonly used. So we choose the basic parameters that are commonly used to realize successfully convert the files from the source to the destination format normally.

Practical function. Users can convert the files from the source to any format they want without complex settings, and they can choose the parameters such as bit rate,

frame rate, resolution, etc. The converter should set the format and parameters automatically for the specific devices (Mobile, MP4).

Modular design. Use the modular design to make the codes more reusable. When updating the interface, only the portion of codes that actually changed needs to be rewritten.

Work independently with the source code of FFmpeg. Using functions to pass the parameters between interface and FFmpeg, and the interface should work independently with the source code of FFmpeg.

The layout design. The GUI consists of four parts, as shown in Fig.3. [3]

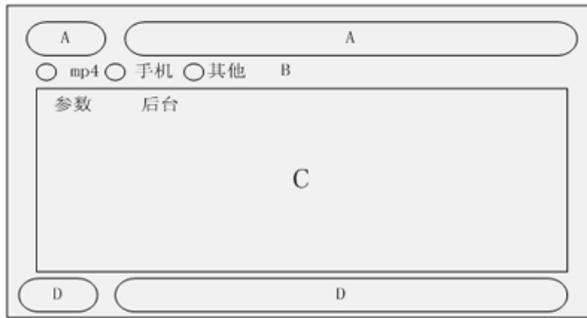


Fig. 3. Overall layout

Part A is designed to input files. Choose the input files from the store path by clicking the button.

Part B is designed to choose the destination format. Users can choose from “MP4”, “Mobile” or “Others”, corresponding with the format of “MP4”, “3GP” and “User-Defined” formats.

Part C is designed for parameter settings and message display. It will give an advice on parameter settings when we choose “MP4” or “Mobile” in part B, and you can choose these parameters by yourself. The other tab is designed to display the running messages of the FFmpeg procedure in Part C.

Part D is the output area. It will set the corresponding format automatically when we choose “MP4” or “Mobile” in part B. we should set the destination format by ourselves when we choose “Others” in part B.

5 The Technical Implementation of the Converter Based on Ffmpeg

Environment configuration. The converter is developed under Linux environment, which version is Ubuntu 5.10. And the version of FFmpeg is FFmpeg-0.6. The GUI is designed based on Qt SDK which version is qt-sdk-linux-x86-opensource-2010.05.

1) Linux environment. FFmpeg is developed under Linux, and is executed by compiling of GCC (GNU Compiler Collection) through the command line under

Linux environment. So, we should construct a Linux environment at first. Generally speaking, there are three ways to construct a Linux environment:

- Install the Linux system on your computer.
- Install a Virtual Machine (VM) under the windows environment, and then install the Linux system in the VM.
- Construct a similar Linux environment by Msys and Mingw.

MinGW is a contraction of "Minimalist GNU for Windows", is a minimalist development environment for native Microsoft Windows applications. It provides a complete Open Source programming tool set which is suitable for the development of native MS-Windows applications, and which do not depend on any 3rd-party C- Runtime DLLs. [4]

MSYS is a collection of GNU utilities such as bash, make, gawk and grep to allow building of applications and programs which depend on traditionally UNIX tools to be present. It is intended to supplement MinGW and the deficiencies of the cmd shell. [5]

2) Qt. Qt is a cross-platform application and UI framework. It includes a cross-platform class library, integrated development tools and a cross-platform IDE. Using Qt, you can write web-enabled applications once, but deploy them across many desktops and embedded operating systems without rewriting the source code. Qt is primarily developed and maintained by the intrepid developers at Qt Development Frameworks, a unit within Nokia. Qt provides a design tool for graphical development- QtDesigner. The QtDesigner works like Dreamweaver, it's efficient to design GUI by dragging controls. [6]

3) FFmpeg. We can use FFmpeg to convert files from FLV format to AVI/ASF/MPEG format or convert AVI/ASF/MPEG format to FLV format without any external libraries' support. In order to get more formatting support, we have to install external libraries before installing FFmpeg, such as Mp3, MP4, 3GP, etc [7], or it will not be supported by other formats.

We should compile FFmpeg to the format of .tar.g or tar.bz2 as installer, because the program of FFmpeg is source program. Install FFmpeg just like installing other softwares under the Linux environment.

The design of converter based on Qt designer. The next two steps show how to design the GUI of the converter: add basic objects, and then design the functions of the controls.

1) Add basic objects. Open Qt creator, using the templates of "Qt Gui Application" to create a new project, including four files: "mainwindow.ui", "mainwindow.h", "mainwindow.cpp" and "main.cpp". "mainwindow.ui" provides a Graphical interface that can be compiled and used for layout design by dragging controls. The needed external libraries and functions are defined in "mainwindow.h", and the functions are implemented in "mainwindow.cpp", "main.cpp" is created automatically by Qt.

Add the needed controls to the file of "mainwindow.ui", and adjust the position of the controls in the design view, as shown in Fig.4.

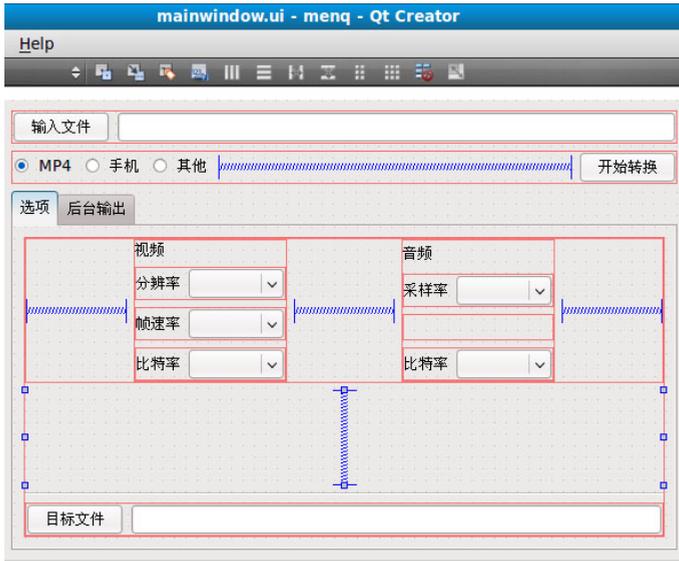


Fig. 4. The layout of the controls in Qt

2) Design the functions of the controls. The controls communicate with functions based on the mechanism of signals and slots in Qt. Controls emit signals when their statuses are changed. Slots are object member functions, they are used to receive the signals [8].

The signals are connected with the functions of slots by calling the function of “connect”, which is defined in QObject object. The function will be called as receiver when the signal is emitted by controls. The function of connect is defined as follows:

```
bool QObject::connect ( const QObject * sender, const char * signal, const
QObject * receiver, const char * member )
```

Enter the Signals/Slots edit mode, and create six Slots: executeCommand(), setSource(), setDestination(), setMP4(), setMobile() and setDefault (). Each Slot corresponding with the signals has six buttons: “Start”, “Input file”, “Output file”, “MP4”, “Mobile” and “Default”.

Add the properties and define the functions for each slot in the file of “mainwindow.h”, and then implement the functions in the file of “mainwindow.cpp”, such as the function of executeCommand() , setSource(), setDestination(), etc.[9]

The function of executeCommand() is implemented to receive the input command of FFmpeg, as shown in below:

```
void MainWindow::executeCommand()
{
    QStringList args;
    args<<“-i”;
    args<<ui->lineEdit->text();//Input file
    args<<“-s”;
```



```

args<<ui->comboResolution->currentText();//Set Resolution of video
args << "-r";
args << ui->comboFramerate->currentText();//Set frame rate of video
args << "-b";
args << ui->comboBitrate->currentText(); // Set bit rate of video
args << "-ar";
args << ui->comboSamplerate->currentText();//Set sampling rate of audio
args << "-ab";
args << ui->comboBitrate->currentText(); // Set bit rate of audio
args << ui->lineEdit_2->text(); //Output file
commandProcess.start("ffmpeg", args); //Execute FFmpeg
}

```

The function of setSource() is implemented to set the input file, as shown in below:

```

void MainWindow::setSource()
{
    QString file = QFileDialog::getOpenFileName(this, tr("选择输入文件"),
        QDesktopServices::storageLocation(QDesktopServices::MoviesLocation));
    ui->lineEdit->setText(file);
}

```

The function of setDestination() is implemented to set the output file, as shown in below:

```

void MainWindow::setDestination()
{
    QString file = QFileDialog::getSaveFileName (this, tr("保存输出文件"),
        QDesktopServices::storageLocation(QDesktopServices:: MoviesLocation));
    ui->lineEdit_2->setText(file);
}

```

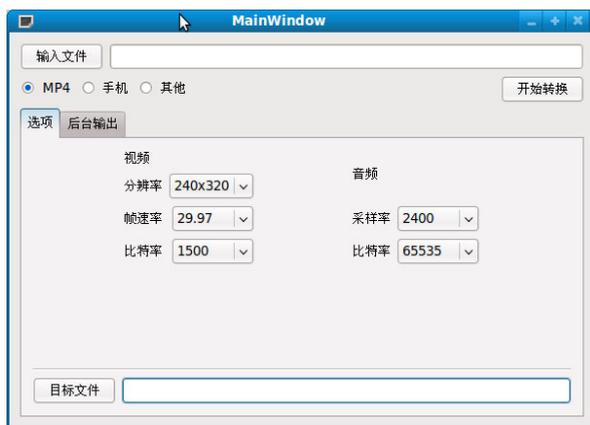


Fig. 5. The final interface of the converter

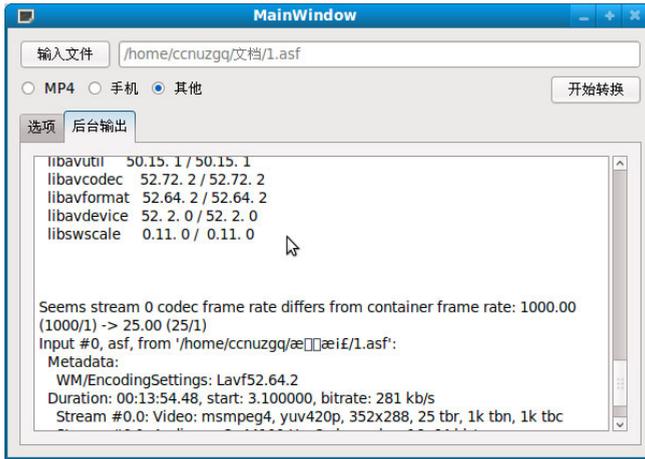


Fig. 6. Output messages

Implement all the defined functions in the same way. The final converter with GUI is shown in Fig.5. The interface is concise and beautiful, and works independently, and the files of FFmpeg will not be affected when converting files from the source to the destination format.

Software testing. Convert files from ASF format to WMV format as a testing. Choose 1.asf as input file, using default settings, input ccnu8.wmv as the output file, and then click the button of Start to convert. The converter works properly, and it will output the conversion messages and error messages in the output tab, as shown in Fig.6.

6 Conclusion

The multimedia format converter with GUI works properly, and is more efficient and convenient for users than ffmpeg.exe. The GUI is concise, easy to use and update. Users can convert the files from the source to any format they want without complex settings.

Acknowledgment. This research is supported by National Significant Science and Technology Special Project of “Core electronic devices, High-end general chips and Fundamental software” (CHB)—The Research and Application of Digital Education Smooth Transfer Oriented to Domestic Fundamental Software (2010ZX01045-001-005-3), National Key Technology R&D Program of “Integrated Services and Demonstration of Digital Music”(2009BAH51B03).

References

1. Project Description[EB/OL], <http://www.ffmpeg.org>
2. Zhang, G.: The design and implementation of video transcoding and protection system based on FFmpeg. Central China Normal University (May 2011)

3. Home of the MinGW, MSYS and mingwPORT Projects [EB/OL],
<http://www.mingw.org>
4. MSYS [EB/OL], <http://www.mingw.org/wiki/MSYS>
5. What is Qt[EB/OL], <http://qt.nokia.com>
6. Compiling and Application of ffmpeg [EB/OL],
http://www.360doc.com/content/10/0428/15/474846_25293076.shtml
7. Blanchette, J., Summerfield, M.: C++ GUI Programming with Qt4, 2nd edn. Publishing House of Electronics Industry, Beijing (2008)

Design and Implementation of Mediaplayer Based on FFmpeg

Yun Cheng^{1,2}, Qingtang Liu^{1,*}, Chengling Zhao¹, Xiaoliang Zhu¹, and Guoqing Zhang¹

¹ The Department of Information Technology
Central China Normal University, WuHan, Hubei, China, 430079

² Department of Educational Science and Technology
HuangGang Normal University, HuangGang, Hubei, China, 438000
liuengineering@TOM.COM

Abstract. FFmpeg is free software licensed under the LGPL or GPL. It is a complete, cross-platform solution to record, convert and stream audio and video. But it's complex and inefficient to use, for it doesn't have a GUI (Graphical User Interface) and can't be used directly under Windows environment. We develop a media player with GUI based on FFmpeg. The player is a MFC-based application which can be used under windows environment. It is more efficient and convenient for users than ffplay.exe. The player has the following advantages: (1) Less files, less size; (2) Easy to use, without installing; (3) High ability of coding/decoding; (4) Friendly user interfaces; (5) Easy to update.

Keywords: FFmpeg, ffplay, GUI; player, MFC.

1 Introduction

FFmpeg is complete, cross-platform solution to record, convert and stream audio and video, which includes libavcodec - the advanced audio/video codec library. FFmpeg is free software licensed under the LGPL or GPL depending on your choice of configuration options. FFmpeg is developed under Linux, but it can be compiled under most operating systems, including Windows. FFmpeg is supported by more than ten transfer protocols, such as HTTP, RTP, RTSP, TCP, UDP, MMS, etc, and also be supported by more then forty audio formats and more than ninety video formats. The high ability of coding/decoding makes it widely used in multimedia area. [1]

2 The FFmpeg Project

The FFmpeg project provides three tools: ffmpeg, ffserver and ffplay. Ffmpeg is a command line tool to convert multimedia files between formats. Ffserver is a multimedia streaming server for live broadcasts. Ffplay is a simple media player based on SDL (Simple DirectMedia Layer) and the FFmpeg libraries. [2]

* Corresponding author.

FFmpeg is developed under Linux, and it is operated based on command line, and is compiled and executed with the aid of GCC (GNU Compiler Collection). It doesn't have a GUI (Graph User Interface), so can't be used directly under Windows environment.

There are two ways to use the FFmpeg in the Windows environment: (1) Rewrite all code of ffmpeg; (2) To compile these C functions of FFmpeg to Dynamic Link Library (.dll) and Static Link Library (.lib), and then to develop application GUI and loading the DLL in Visual C++ environment. Obviously, the latter is more efficient. But we have to use 3rd-party tools to compile FFmpeg, the two most important are MSYS (Minimal SYStem) and MinGW (Minimalist GNU on Windows).

MinGW is a contraction of "Minimalist GNU for Windows", is a minimalist development environment for native Microsoft Windows applications. It provides a complete Open Source programming tool set which is suitable for the development of native MS-Windows applications, and which do not depend on any 3rd-party C-Runtime DLLs. [3]

MSYS is a collection of GNU utilities such as bash, make, gawk and grep to allow building of applications and programs which depend on traditionally UNIX tools to be present. It is intended to supplement MinGW and the deficiencies of the cmd shell. [4]

3 A Simple Media Player Based on Ffplay

The decoding process of ffplay. Ffplay is a simple media player based on SDL and the FFmpeg libraries. The Decoding process of ffplay is as follows [5], shown in Fig. 1:

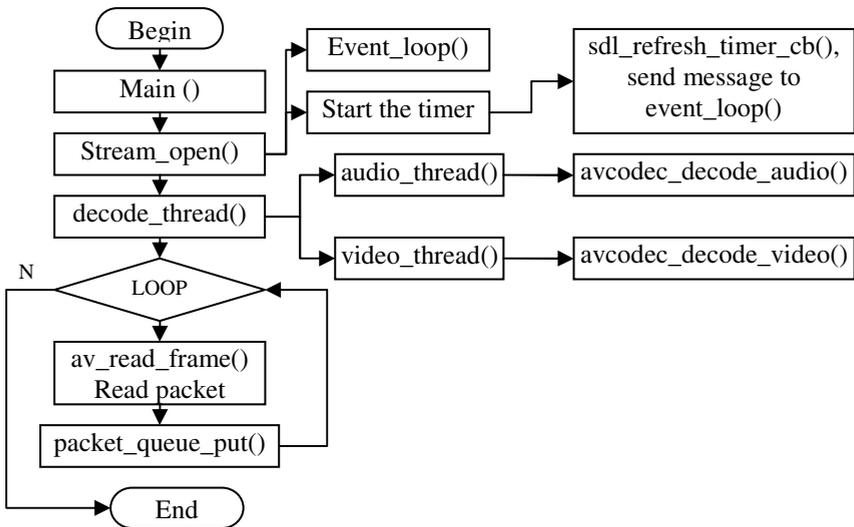


Fig. 1. The Decoding process of ffplay



- (1) Call the initialization functions in main function;
- (2) Then open a stream;
- (3) And call the decode_thread function to create video/audio/subtitle thread;
- (4) Start the timer as the notification function at the same time, and enter the event loop. The timer will call the function of sdl_refresh_timer_cb and send message of REFRESH_EVENT to event_loop function periodically;
- (5) Call the av_read_frame function to read a packet. If it is a video packet, then send it to video thread, call the avcodec_decode_video function to decode the packet; if it is a audio packet ,then send it to audio thread, call the avcodec_decode_audio function to decode the packet;
- (6) Judging whether to stop playing from time stamp;
- (7) If not, then Go to (5), else stop playing.

Compile FFmpeg under windows environment. We will get ffmpeg.exe, ffplay.exe, ffsrver.exe, Dynamic Link Library files (.dll), Static Link Library files (.lib) by compiling Ffmpeg under Windows environment. The Dynamic Link Library files include avcodec.dll、avformat.dll, avutil.dll, swscale.dll; the Static link library files include avcodec.lib, avformat.lib, avutil.lib, swscale.lib.[6]

Run the application of ffplay.exe to playback. Ffplay.exe is a simple windows application media player based on SDL and the FFmpeg libraries. So the player works depending on the DLL files, including avcodec.dll, avformat.dll, avutil.dll, swscale.dll. At the same time, the player displays the video images by using SDL, so it also depends on the DDL file of SDL-SDL.dll.

The player can't play video/audio directly, for there is no GUI of FFmpeg. We should use command line to call the windows application of ffplay.exe. The command syntax is as follows: ffplay.exe test.mpg or ffplay test.mpg.

Generally speaking, there are three ways to call the application of ffplay.exe under Windows environment:

- 1) Input command and execute it in the MSYS environment.
- 2) Run the windows application of cmd.exe, then input command and execute it in the DOS environment;
- 3) Use notepad to create a batch file under the windows environment (.bat), taking test.bat as an example, as shown in Fig. 3. Open the batch file by notepad. Input the command, save and close the file. Then open the batch file directly, windows will execute the command and run the application of ffplay.exe.



Fig. 2. Using the notepad to create the test.bat file

No matter what methods you decide to use, windows will execute the command and call the application of `ffplay.exe`. In this example, the `ffplay.exe` will play the video file of `test.mpg`, as shown in Fig. 3.

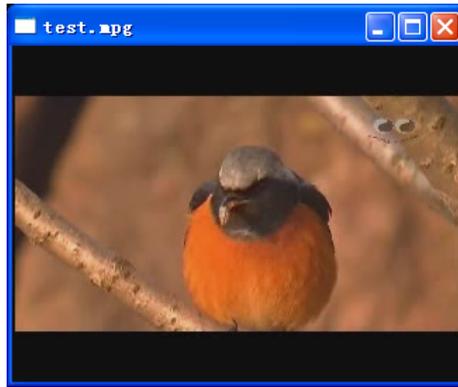


Fig. 3. The playing effect by directly calling `ffplay.exe`

4 The Design of GUI for the Player

We have to call the application of `ffpaly.exe` by executing command, but it's too complex and inefficient for most users because of no GUI, and it is much inconvenient to control the player. A GUI will improve the efficiency as well as the practicability of the player.

The layout design. The GUI of the player mainly consists of six parts: title bar, menu bar, display area of video images, progress bar, tool bar and status bar, as shown in Fig. 4.

The title bar is used to display the file name and the player name; the menu bar is used to provide various functional menus; the display area is used to display the video images, if the file is audio, the area is hidden. The progress bar is used to display playback progress. The tool bar is a control area, and it is used to provide common tools. The status bar is used to display the file information and the status of the player.

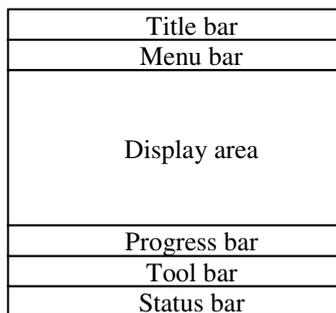


Fig. 4. The layout design of the player with GUI

Function analyses. A media player should provide some functions to control, such as file control, view operation, playing control and help information, etc.

The media player supports the commonly used function as follows: open file, close file, exit, play, pause, stop, previous, next, fast forward, rewind, scale, drag the progress, volume control, and so on, as shown in Fig.5.

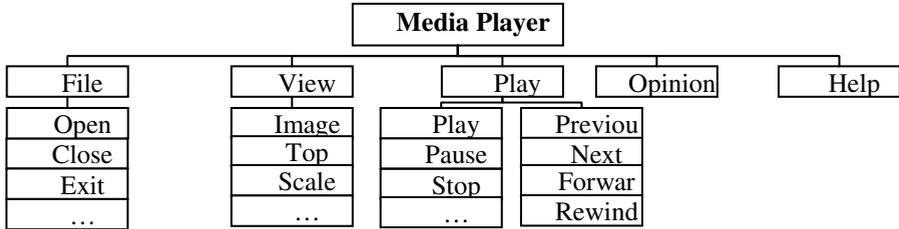


Fig. 5. Function and Architecture of the player

Design of the main functional module. The menu bar can be divided into five submenus including the following functional design: File, View, Play, Opinions and Help. File menu is designed about file operations, such as open, close, exit, etc. View menu is designed to control how the view displays, such as image sizes (half size/original size/double size/full screen), open or close the progress bar/tool bar/status bar, always display on top, etc. You can scale the image sizes as you want by dragging with the mouse. Play menu is deigned to provide play functions like Play, Pause, Stop, Previous, Next, Forward(10 Second/1 Minute/10 Minutes), Rewind(10 Second/1 Minute/10 Minutes), Volume control, etc. And we can go to anytime of the video. The Opinions menu gives users the ability to personalize the GUI. Help menu provides help information for users.



Fig. 6. The GUI of the player

The progress bar is used to display playback progress, and you can drag the slider to change the playing position.

The tool bar is a control area of common tools shown as graphic button, including play, pause, stop, previous, next, repeat, playlist and volume control, as shown in Fig.6. We can control volume between 0% and 100% by dragging the volume slider with the mouse.

The play list window appears when we click the play list button. We can manage the files from play list window, such as play, add files, delete files, load play list, etc, as shown in Fig.7.

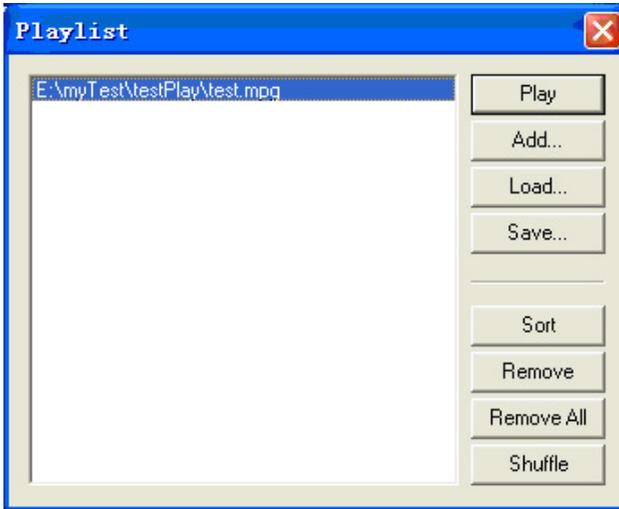


Fig. 7. Playlist Window

The status bar is used to display the file information and the status of the player. You can get the information such as author, title, bit rate, copyright, duration and the current time of the playing position, etc.

5 The Technical Implementation of the Player with GUI Based on Ffmpeg

Now, we have a GUI for the player, the next is to call the functions of FFmpeg, then we can use the player to control the playback

Technological solutions. We develop the player in visual studio 2005, and write the program with C++ so as to be fully compatible with the FFmpeg project. The player works based on SDL and the FFmpeg libraries. We can use the MFC (Microsoft Foundation Classes) to design a GUI for the player in visual studio 2005, and call the functions of SDL and FFmpeg to playback.

SDL is a cross-platform multimedia library designed to provide low level access to audio, keyboard, mouse, joystick, 3D hardware via OpenGL, and 2D video framebuffer. It is used by MPEG playback software, emulators, and many popular games. [7]

MFC is a library that wraps portions of the Windows API in C++ classes, including functionality that enables them to use a default application framework. Classes are defined for many of the handle-managed Windows objects and also for predefined windows and common controls. [8]

Call the functions of FFmpeg. Prepare the libraries for the program, including Dynamic Link Library files (avcodec.dll, avformat.dll, avutil.dll, swscale.dll, SDL.dll) and Static Link Library files (avcodec.lib, avformat.lib, avutil.lib, SDL.lib, swscale.lib, WinMM.lib). We can call some windows functions with WinMM.lib in visual studio 2005.

1) Copy the head files of FFmpeg to the program, or download the FFmpeg SDK from <http://www.ffmpeg.org>.

2) Set project properties. Add the static libraries to the project. Expand Project->Attribute-> Configuration attributes->Link->Input->Additional library, then write down the libraries: avcodec.lib, avformat.lib, avutil.lib, SDL.lib, swscale.lib, WinMM.lib, separate each file with a space. Add the directory of the static libraries. Expand Project->Attribute->Configuration attributes->Link-> General-> Additional library path, and then write down Additional library path. Add the directory of include files. Expand Project->Attribute->Configuration attributes->C/C++->General->Additional include files path, and then write down additional include files path.

Rewrite the programs of ffplay.c. Rewrite the programs of ffplay.c into the MFC-based player with GUI, and call the functions to playback. [9]

Compile and run the project. Compile and run the project when finished the programs, and we will get a player with GUI, as shown in Fig. 8.

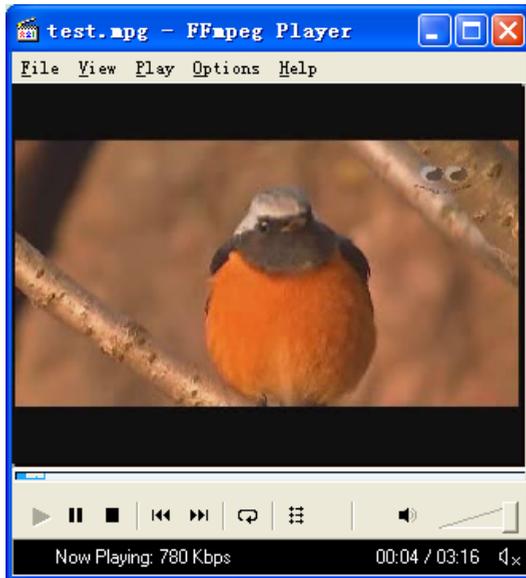


Fig. 8. The playing effect of mediaPlayer testing

6 Conclusion

The player with GUI is more efficient and convenient for users than ffplay.exe. The player has the following advantages:

(1) Less files, less size. We get the application of MPlayerGUI.exe as a player by compiling the project in Visual Studio 2005, and it works depending on the five additional DDL files (avcodec.dll, avformat.dll, avutil.dll, swscale.dll, SDL.dll). The total size of these six files is 7.94MB.

(2) Easy to use, without installing. The player will call the DLL files automatically when we run the application of PlayerGUI.exe to play video/audio. If we want to run it on other computers, what we need to do is just to copy the six files to another computer, and run it directly without installing.

(3) High ability of coding/decoding. The player retains high ability of coding/decoding of FFmpeg, and most of the video/audio formats are supported.

(4) Friendly user interfaces. The player is easy to be operated with a friendly GUI, so users can be easy to interactive with it.

(5) Easy to update. We compile the latest version of FFmpeg under windows environment. When the libraries of FFmpeg are updated, just replace the old DLL files with the new DLL files.

Acknowledgment. This research is supported by National Significant Science and Technology Special Project of “Core electronic devices, High-end general chips and Fundamental software” (CHB) ——The Research and Application of Digital Education Smooth Transfer Oriented to Domestic Fundamental Software (2010ZX01045-001-005-3), National Key Technology R&D Program of “Integrated Services and Demonstration of Digital Music ”(2009BAH51B03).

References

1. Project Description[EB/OL], <http://www.ffmpeg.org>
2. Home of the MinGW, MSYS and mingwPORT Projects [EB/OL], <http://www.mingw.org>
3. MSYS [EB/OL], <http://www.mingw.org/wiki/MSYS>
4. The principle of FFPLAY [EB/OL], <http://blog.csdn.net/shenbin1430/archive/2009/06/23/4291893.aspx>
5. Compiling FFmpeg under the windows environment [EB/OL], <http://hi.baidu.com/xulina809/blog/item/3e07bbd507a45bd6a144dfb1.html>
6. SDL[EB/OL], <http://www.libsdl.org>
7. Microsoft Foundation Class Library [EB/OL], http://en.wikipedia.org/wiki/Microsoft_Foundation_Class_Library
8. The working group of FFmpeg[EB/OL], <http://www.ffmpeg.com.cn/index.php>

The Research and Implementation of Incident Response Information System*

Zuyi Chen, Hua Yong, and Taixiang Zhao

Department of foundation, First Aeronautic Institute of the Air Force, Xinyang, China
chen_3341@163.com

Abstract. This paper puts forward a response decision-making algorithm of emergency response management system. And this algorithm which bases on all factors relate to it can judge which targets are immune to attacks and which events may be filtered out by response strategy, it also can do response benefit sort on all feasible response mode of security event that it needs to response and finally choose the best response mode from all. And the response suggestion it generates has pertinency. Besides, response decision-making can adjust according to the change of circumstance, meanwhile, it has a good expansibility to the response mode.

Keywords: Emergency response, Emergency response management system, response benefit, response decision-making.

1 Introduction

Emergency response is a new branch of information security filed, usually refers to the preparations an organization does for dealing with various security events and remedial measures and actions the organization adopts after the event happened in order to prevent or reduce the impacts on security of emergency response management system.

2 Emergency Response Decision-Making

Response decision-making and the execution of response action are two key technologies of event response, and response decision-making is the core, that's because response decision-making is the foundation of the execution of response action, it directs the execution of response action. Therefore designing a reasonable, efficient and feasible response decision-making algorithm is the most important part of event response technology.

Classificationn-based response decision-making. Fisch[1] classified the responses of events upon the detected timing and response target for the first time. But it also has

* Supported by natural Science Foundation of Henan Province in China under Grant No. 2011C520012.



its shortcoming, for example, it is not comprehensive and too coarse, besides, it can't be used in response decision-making.

Lindqvist [2] divided security events into three types on the basis of consequences, and according to the confidentiality of the system, events like Exposure can be further divided into the leakage of confidential information and unauthorized access to service.

Curtis [3] [4] proposed a response-oriented way of classification of security events on the basis of the study of Fisch. And this method bases on six dimensions, each one has specific granularities.

Cost-Sensitive response decision-making. Lee et al [5] thought that intrusion detection system should obtain maximum safety with the minimum cost. And on the basis of this thought they proposed an evaluation method about the effect of intrusion detection system. The most prominent progress of Lee's thought about cost sensitive response decision-making compared with classification-based response decision-making method is that it can carry out quantitative analysis of the response. But the imperfection of this quantitative method is that it is too rough, so, there still a lot needs to be improved: Such as, First, it doesn't know that different response to the same kind of attacks has different response cost. Secondly, the quantification of all kinds of cost is not reasonable. Thirdly, the insufficient analysis of response cost makes response decision-making algorithm too simple to be reasonable. Finally, this type of decision-making model neglects the negative influence caused by response behavior.

3 Classification-Based Benefits First Response Decision-Making

In view of the defects of the two thoughts about response decision-making mentioned above, this paper puts forward a classification-based benefits first response decision-making.

The model of classification-based benefits first response decision-making considers only costs associated with the response decision-making, and it divides the events into different types so that the quantification and evaluation of all kinds of cost can be specific to events of a kind. Classifying the events is critical for the quantification of cost. For events of a kind, in order to obtain maximum safety with the minimum cost, it's necessary to convert the main factors affecting the response into cost or quantitative values related to cost, and then establish a general model of calculating cost with most factors considered, and all response decision-making are based on the results of its calculations, then the response with lowest cost and best response effectiveness will be chosen.

The quantification of cost. In terms of the quantification of cost, the classification of events and the quantification of Damage Cost (DCost) the model of classification-based benefits first response decision-making adopts the methods of Lee, so does the calculation of Residual Damage Cost (RDCost). The Total Cost (TCost) including RDCost, Response Operation Cost (ROCost) and Response Negative cost (RNCost). So, the value of TCost is the sum of the three kind of cost, i.e.:

$$TCos = RDCost + ROCost + RNCost \quad (1)$$

The quantification of TCost. Damage Cost is the losses an event do to the target system under the situation of no response measures had be taken, and it is recorded as DCost.

The value of RDCost generally can be got through the way using Dcost multiplied by residual coefficient.If e is a security event of some kind, and r is some response mode, $\mu(e, r)$ is the residual coefficient of r for e , then:

$$RDCost(e,r)=\mu(e,r)*DCost(e) \quad (2)$$

Generally, ROCost can be measured by the loss of system availability, which in turn can be transformed into calculating on the basis of the Damage Cost caused by DOS attack. That's because the direct consequence of DOS attack is the loss of the system availability, then:

$$ROCost(r)=m*DCost(DOS),m\in[0,1] \quad (3)$$

RNCost is the loss the response mode caused to users while eliminating damage done by some event, then:

$$RNCost(r)=n*Dcost, n\in(0,1] \quad (4)$$

Then, the value of Tcost of e can be derived from the formula (1), (2), (3), (4).

$$\begin{aligned} TCost(e,r) &= RDCost(e,r) + ROCost(r) + RNCost(r) \\ &= \mu(e,r)*DCost(e) + m*DCost(DOS) + n*DCost(DOS) \\ &= \mu(e,r)*lethality(e)*criticality(target(e)) + m*lethality(DOS)*criticality(target(DOS)) + \\ &\quad n*lethality(DOS)*criticality(target(DOS)) \quad m\in[0,1], n\in(0,1] \end{aligned} \quad (5)$$

The qualification of the above costs are all discussed in a limited and open set of response modes.

Response decision-making algorithm. To improve the rationality and adaptability of response decision-making, The classification-based benefit first response decision-making algorithm takes the following factors into consideration. (1) The confidence of security events. By calculating the confidence of event alarming, response decision-making algorithm limits the negative impact caused by intrusion detection system.. (2) The type of security events. Different type of security requires different response method, and different type of events have different fatal attacks. (3) The importance of the target of the event. For different events, if the importance of the targets is different, then, the response mode should also be different. (4) The secure state of the target. The Intrusion detection systems used now are mostly rule-based, but it will not take the secure state of target into consideration while detecting. For example, an attack on netbios was detected, and the attack is specific to windows system, but the target operating system is unix, then the attack has no effect on the target, then there is no need to respond. This situation usually referred to as that the target has full immunity to this attack.. (5) Dcost. The quantification of the Dcost associates the seriousness of the event with a value. (6) ROCost. The quantification of the ROCost associates the cost needed to pay for the response to the event with a value. (7) The effect of response. Emergency response management system first sets out an initial response strategy and carries it out. But the initial one may be not the optimal one. So, through assessing and responding to the effects of all response strategies, and then storing all effects of a

response strategy in the information base, the system can call the response strategy which has the best effect and then carry it out when encounters security events of a kind. It is this idea that enables the system to enhance the effect of response gradually as the growth of running time and achieve the self-adaptive and self-learning of the emergency response system.

Combining the the theory about response strategy and practical effect of response, we can conclude that the response decision-making is a function of TCost and the effect of response. And then, Under the running of the system, response strategies with high cost and bad effect will be eliminated.

Definition 1: Setting $f(e,r)$ as the value of response benefit that the response mode r for event e and $effect(e,r)$ as the value of effect that the response mode r for event e , then, $f(e,r)$ should be proportional to $effect(e,r)$ and inversely proportional to TCost. Then, it can be defined that:

$$f(e,r) = effect(e,r) / TCost(e,r)$$

Definition 2: Setting the set of response modes supported by emergency response management system is R , then, $R = \{r_0, r_1, r_2, \dots, r_n\}$, r_0 stands for the response mode without any response (i.e. an empty response). For security event e , if there exists a response mode r^* meets for the following conditions:

$$f(e, r^*) = \underset{r \in R}{Max}(f(e, r))$$

Then, r^* was the best response mode for e the emergency response management system adopts.

The quantification of Dcost discussed earlier assuming that every event is a real event, but this is unrealistic, especially in cases that the events are caused by the intrusion detection system, that's because the intrusion detection system may misreport, and just because of this many intrusion detection systems take the confidence of the event alarming as a feature. Besides, the security level of network determines to some extent the implementation of response strategy is severe or mild. Therefore, in order to make the quantification of DCost more reasonable, the response decision-making algorithm takes the confidence of events and the security level of network into consideration while calculation the value of Dcost. The formula is showed as (6). And in this formula sererity stands for the quantitative value of the security level of current network.

$$DCost(e) = lethality(e) \times importance(target(e)) \times confidence(e) \times sererity \quad (6)$$

Assuming the security event needs to response is e , then, what the classification-based benefit first response decision-making algorithm should do after having e classified is find out the one with best response benefit from all response modes for e . The specific steps are as follows:

1. According to the type of e , the algorithm finds out the optional one from all response modes. Then, combining with empty response mode can form the set of response modes: $\{r_0, r_1, r_2, \dots, r_n\}$.



2. For each response mode r_i of e , according to the response strategies and formulas given above, we can separately calculate the value of $RDCost(e, r_i)$, $ROCost(r_i)$, $RNCost(r_i)$, and $TCost(e, r_i)$. Finally, we can conclude that:

$$TCost(e, r) = \sum_1^n TCost(e, r_i)$$

3. If the value of $Tcost(e,r_0)$ is the minimum, then, no response should be done on this attack. That's because the value of $TCost$ is greater than that of $Dcost$, there is no need to respond to the event, and the event will be filtered out by response strategies.

4. Then, for each response mode r_i , the algorithm can get its $f(e, r_i)$, and then, select the response mode that has the maximum $f(e, r_i)$.

5. According to the type of e , the algorithm can find the appropriate response strategies of the eradication phase and recovery phase from the list of response strategies.

6. Combining the response strategies generated above with information related to e , can produce the specific response recommendations on targets.

4 Algorithm Assessment

The analysis of the common examples of attack responses indicates that the classification-based benefit first response decision-making algorithm is reasonable and extensible.

Before putting the model of the algorithm into real-time use, we need to verify whether the system can work just as expected through simulation. Through simulation some parameter values can also be determined. After simulation, if the result is correct, then, the system can be put into normal use. To test the model, the following attacks from intrusion detection system will be used.

IRIS system received a report from ids about the attacks on the key servers, and Table 1 shows the details, Figure 1 shows the type distribution of the attacks.

Table 1. The list of attacks on the key servers in the March x, xxxx

| Source IP | Destination IP | Source Port | Socket Port | NO |
|---------------|----------------|-------------|-------------|-----------|
| 220.181.XX.XX | 202.112.XX.XX | 2048 | 19622 | 400000032 |
| 220.181.XX.XX | 202.112.XX.XX | 2048 | 19622 | 300000001 |
| 202.108.XX.XX | 202.112.XX.XX | 2475 | 80 | 400000092 |
| 202.108.XX.XX | 202.112.XX.XX | 3845 | 80 | 400000092 |
| 202.108.XX.XX | 202.112.XX.XX | 4845 | 80 | 400000092 |
| 202.108.XX.XX | 202.112.XX.XX | 2561 | 80 | 400000092 |
| 203.194.XX.XX | 202.112.XX.XX | 778 | 6257 | 300000063 |
| 202.108.XX.XX | 202.112.XX.XX | 1723 | 80 | 400000092 |
| 202.61.XX.XX | 202.112.XX.XX | 1024 | 42335 | 300000060 |



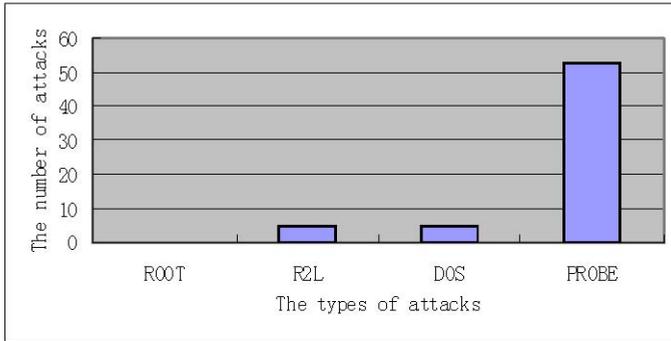


Fig. 1. The type distribution of the attacks

5 Conclusion

This paper studies the response decision-making which is one of the key issues of emergency response management system. Based on the study of wenke lee, some reasonable improvements on the method of qualifying the related costs has been done.

References

1. Fisch, E.A.: Intrusion Damage Control and Assessment: A Taxonomy and Implementation of Automated Response to Intrusive Behavior. PhD Dissertation, Texas A&M University, College Station, TX (1996)
2. Lindqvist, U., Jonsson, E.: How to systematically classify computer security intrusions. In: Proc. 1997, IEEE Symp. On Security and Privacy, Oakland, CA, May 4-7, pp. 154–163 (1997)
3. Carver, C.A., Hill, J.M.D., Pooch, U.W.: Limiting Uncertainty in Intrusion Response. In: Proceedings of the IEEE Workshop on Information Assurance and Security (2001)
4. Carver, C.A., Pooch, U.W.: An Intrusion Response Taxonomy and its Role in Automatic Intrusion Response. In: Proceedings of the 2000 IEEE Workshop on Information Assurance and Security, June 6-7, pp. 129–135. United States Military Academy, West Point (2000)
5. Lee, W., Fan, W., Miller, M., Stolfo, S., Zadok, E.: Toward Cost-Sensitive Modeling for Intrusion Detection and Response. *Journal of Computer Security* 10(1), 318–336 (2002)

The Application of Semantic Web in E-Learning Personalized Knowledge Query

Zuyi Chen, Liu Fang, and Taixiang Zhao

Department of foundation, First Aeronautic Institute of the Air Force, Xinyang, China
chen_3341@163.com

Abstract. Semantic Web represents the direction of the next generation of Internet technology, that's because it provides a new opportunity for the development of deep-level network applications. Applying Semantic Web technology in E-learning will help to break through the status of curriculum resource constructing for E-learning so as to highlight the advantages of new technology in resource constructing. Besides, it also help to deepen the using of curriculum resource. Meanwhile, it also provides an effective method for solving the problems of knowledge acquisition, knowledge reuse and knowledge sharing in E-learning. In addition, the application of Semantic Web in E-learning, will help to achieve the intelligent of E-learning system in a certain extent, and will effectively improve the learner's enthusiasm and initiative.

Keywords: Semantic Web technology, Ontology, E-learning.

1 Introduction

E-learning refers to a way of learning through Internet, it allows people to learn anytime, anywhere they like, thus it provides the possibility for lifelong learning. With the rapid development of information technology and network technology, more and more people from different fields began using E-learning as their perfect way of learning and make it an important way of achieving continued education or lifelong learning. So, All levels of China's education department invest a lot of human, material and financial resources in the construction of E-learning environment and resources, to make the environment of E-learning more and more perfect, to make all sorts of resource more and more abundant both in quantity and variety, besaids, the quality keeps improving.

However, under the existing web environment the information resources have the following characteristics: heterogeneous format, multiple meanings, lack of relations and the relations are not unified and so on. It is these characteristics that make it is hard to realize the interconnection and interoperability of information resources, and lead the low efficiency of knowledge sharing and knowledge reuse.

Meanwhile, the existing E-learning platform has deficiencies in providing personalized resources. Generally, E-learning platform provides different learners the same resources, but each learner has his own interest and characteristic, so, different learner may needs different resources, and the amount of interesting resources which is

limited needs to take a lot of time and energy to find, therefore, appropriate allocation of resources according to different learner characteristics become important.

Therefore, because of the above defects existing in E-learning system, a new way of organizing knowledge which can be adopted to realize the personalized knowledge query so as to meet the different needs of learners to knowledge and realize the sharing and reuse of resources is needed. The appearance of Semantic Web provides an effective way to this function.

2 Semantic Web and Its Key Technology

After Tim Berners-Lee[1]who presented that Semantic Web was the extension of the Web, the information of it were all had well-defined, and it can support man-machine coordinated work Semantic Web technology had received wide attention.

As a new emerging technology, Semantic Web technology try to make the information on the web can be understood by machines, and achieve automatic processing of Web information, so as to adapt to the rapid growth of web information, and provide better service to mankind.

With the development of Semantic Web, we have clearly know the seven-story structure of it which is proposed by Berners-Lee, so, we will not discuss it here.

For Semantic Web, we know that the key technologies of it are Extensible Markup Language(XML),Resource Description Model (RDF/RDF Schema) and Ontology. And they play an important role in the whole structure.

(1)XML which allows for custom tags and unifies the way of storing resources is a universal, structured description language, which unified the way to storage resources. XML provides the syntax for structured documents to achieve the separation of document structure and document forms. So, the same document for different purposes can have different forms.

(2)RDF which was proposed by W3C in February 1999, was used to describe resource and the relationship between resources. RDF has the following features: simple, easy to expand, open, easy to exchange and easy to integration.RDF gives a solution which uses metadata to solve the integration of web data, and it is the basis of encoding, exchange and reuse metadata through supporting the design of related mechanisms such as public norms of semantics, syntax and structure and so on.

(3) Ontology, as one important part of Semantic Web, is a research focus. Currently there is no uniform definition of ontology, and the generally accepted definition which is defined by Studer et al is that ontology is the explicit, formal specification of shared conceptual model.

Ontologies provide a common vocabulary of an area and define the meaning of the terms and the relations between them with different levels of formality. Knowledge in ontologies are formalized using five kind of components: classes, relations, functions, axioms and instances(Gruber,93).

(1) Concepts(classes) are used in a broad sense. A concept can be anything about which something is said and, could also be the description of a task, function, action, strategy, reasoning process, etc.

(2) Relations represent a type of interaction between concepts of the domain. They are formally defined as any subset of a product of n sets, that is: $R: C_1 * C_2 * \dots * C_n$. Examples of binary relations are: subclass-of and connected to.

(3) Functions are a special case of relations in which the n -th element of the relationship is unique for the $n-1$ preceding elements.

(4) Axioms are used to model sentences that are always true.

(5) Instances are used to represent elements.

3 The Key Technology to Realize Personalized Query in E-Learning

The technologies which are relate to Semantic Web are used in the E-learning system to realize the function of personalized query are: ontology, Semantic Annotation, Ontology mapping, Ontology integration and Jena Semantic Web framework.

3.1 Ontology

As we said before, ontologies provide a common vocabulary of an area, through defining the relationships among concepts strictly to ascertain the exact meaning of concepts and present a commonly recognized and can be shared knowledge so as to provide semantic supports to information query.

Ontologies have the potential to describe the semantics of data sources, there are many advantages of using ontologies to realize personalized knowledge query in E-learning system. Firstly, Ontologies provide a rich set of predefined vocabulary, and it can be used to describe the semantic information of data source exactly, and as the semantic model of data source which are independent of the data model. Secondly, knowledge that is presented by the form of ontology can meet the needs of user to semantic search enoughly. Thirdly, ontologies which based on description logic can support the reasoning of user queries.

So, based on the analysis above, we can conclude that to realize the function of personalized knowledge query the premise is to construct domain ontology. In terms of specific course of E-learning platform it means to construct course ontology. And for course ontology, we need to introduce the concept of knowledge points. Knowledge points which are locally complete can be used independently and are smaller units of course. So, knowledge points can be taken as concepts of ontology, so as to construct the hierarchy of concepts. Then after analysing, we believe that there mainly exist the following links among knowledge points: prior/next, consist of, relate to, cite and be cited.

3.2 Semantic Annotation

Ontologies all have certain commonality, usually used to represent knowledge of some specific area, but because the number of instances of a certain field is infinite and has a dynamic change, therefore, it makes sense to consider the instances only when relate the ontology to a specific application. Contacting the instances involved in practical applications and abstract concepts of ontology is the work semantic annotation should

do. Compared to the database, semantic annotation like add specific records to established database tables. From the perspective of the Semantic Web, semantic annotation is the process of releasing semantic information, that's to say, based on some certain ontology, the user adds semantic information for web pages.

Currently, there are many methods of semantic annotation, and all these methods can be classified into three categories, i.e. manual annotation, through concepts mapping between document type definition(DTD)and document modes, annotating through analysis of the semantic vocabulary.

Now, the main tools of semantic annotation are: SHOE Knowledge Annotator, Annotea, OntoMat-Annotizer, Annozilla and so on. In my study I use OntoMat-Annotizer[2], that's because it is a user-friendly interactive webpage annotation tool. We can use it to annotate web pages which contain the contents of courses of E-learning system that the learners should learn.

3.3 Ontology Mapping

Because of the existence of synonyms, different source of information using different terms when describing the same thing. So, it's necessary to introduce the technology: ontology mapping.

Ontology mapping is a method to solve the problem of knowledge sharing and reuse between different ontologies, which aims to identify the different ontology semantic relationship between entities, and to express its formal. Its purpose is to identify the semantic relationship between entities in different ontologies, and express it formally. Ehrig[3] defined ontology mapping as follows: ontology mapping refers to the two ontologies A and B,for each concept of A trying to find a same or similar concept in B. So is each concept in B.

3.4 Ontology Integration

When a task requires to interoperate multiple ontologies, then Then we need to introduce ontology integration to solve the problem.

Ontology integration means that when a task relate to ontology refers to more than one ontology, and the ontologies used are inconsistent in many ways, then, to have these heterogeneous ontologies can interact with each other, we need to set up and handle mapping between these ontologies, so as to achieve ontology alignment or ontology merging[4].

3.5 Jena Semantic Web Framework

Jena is a Java framework for building Semantic Web applications. It provides a programmatic environment for RDF, RDFS and OWL, SPARQL and includes a rule-based inference engine.

The Jena Framework includes[5]:

- A RDF API, The main functions of RDF API include creating, reading and writing RDF models, and manipulating and retrieving RDFmodels.
- Reading and writing RDF in RDF/XML, N3 and N-Triples

- An OWL API. An OWL API provides direct support to OWL, DAML+OIL and can operate the concepts(classess), attributes, instances and the relations between them. Besides, together with inference engine it also can do reasoning to ontology-based files to discover hidden information.
- In-memory and persistent storage.
- SPARQL query engine. SPARQL which is a query language for RDF, can be used to express queries across diverse data sources, whether the data is stored natively as RDF or viewed as RDF via middleware.

4 Conclusion

With the application of Semantic Web technologies become more mature, Semantic Web and related technology used in E-learning in the future will be very good. Of course, the Semantic Web technology to promote the development of E-learning still needs the participation of many researchers.

References

1. Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Scientific American 284(5), 34-43 (2001)
2. <http://annotation.semanticweb.org/ontomat/index.html>
3. Ehrig, M., Sure, Y.: OntologyMapping-Anintergrated Approach. In: Bussler, C.J., Davies, J., Fensel, D., Studer, R. (eds.) ESWS 2004. LNCS, vol. 5, pp. 10-12. Springer, Heidelberg (2004)
4. Meila, M., Xu, L.: Multiway cuts and spectral clustering. U.Washingt on Tech Report (2003)
5. <http://jena.sourceforge.net/documentation.html>

QoS-Based Web Service Selection Approach

Guofeng Chang

Department of Computer and Information Engineering, Xinxiang College, Xinxiang, China
GUOTEACH@tom.com

Abstract. In this paper, we proposed a QoS-based web service selection approach. The approach adopts genetic algorithm to select the most suitable web service with user's QoS requests. Experimental results show that our proposed approach is very effective in web service composition system.

Keywords: Quality of Service, Web service, Service selection.

1 Introduction

Web services encapsulate application functionalities and information sources, and promise to take cross-network application interactions one step further by enabling programmatic access to applications over the web [1-2]. For a web environment, multiple web services may provide similar functionalities with different non-functional property values. Then these web services will typically be grouped together in a single community. To differentiate the members of a community during service selection, their QoS values need to be considered and it is highly recommended to be taken into account during the web service selection [3-4].

Web services technology is a promising solution for addressing platform interoperability and compatibility problems faced by system integrators. However, their adoption rate has been very slow due to the lack of QoS during the web service selection [5-6]. web services technology has yet to face questions such as: "How I will know the web service will meet my performance requirements?" Until these questions have been solved, it is unrealistic to expect a business application in a Universal Description, Discovery and Integration (UDDI) registry. At present, the problem of QoS-based web service selection and composition has received a lot of attention during the last years [7-8].

In [9], the authors presented the research results of our group in delivering a framework for the deployment of adaptable Web service compositions. The publication infrastructure integrates existing heterogeneous repositories and makes them cooperate for service discovery. The deployment infrastructure supports BPEL-like compositions that can select services dynamically, and also adjust their behavior in response to detected changes and unforeseen events. The framework also provides a monitoring-based validation of running compositions. The various parts of the framework are exemplified on a common case study taken from the automotive domain.

In [10], the authors present an architecture to facilitate efficient evaluation and selection of 3rd party web services for service providers. Most service provider architectures have primarily focused on providing web service front ends to legacy systems, aggregating and delivering services via workflows. These architectures primarily considered static business contracts between the service provider and its (web service enabled) business partners. This approach makes these architectures inflexible to variations in business requirement, partners' performance and customer requirements. Their proposed architecture provides a flexible means for service providers to optimize business performance. Then based on the historical performance, extant context, and optimizing business rules, the appropriate service is selected and invoked to serve a customer request.

In [11], the author propose a QoS-aware fault tolerant middleware to attack this critical problem. Our middleware includes a user-collaborated QoS model, various fault tolerance strategies, and a context-aware algorithm in determining optimal fault tolerance strategy for both stateless and stateful Web services. The benefits of the proposed middleware are demonstrated by experiments, and the performance of the optimal fault tolerance strategy selection algorithm is investigated extensively.

In [12], The authors presented a novel Uncertain multi-attribute decision making-based semantic web service composition algorithm (UMC) to solve the above difficulties for the first time. UMC takes all possible QoS expression types (real number, interval and linguistic) into consideration and includes three main steps: defuzzifying linguistic data, normalizing the decision matrix and evaluating alternatives synthetically. Further, the strategies of uncertain preference conversion are also discussed in UMC

In [13], the author introduced an online monitoring approach for web service requirements. It includes a pattern-based specification of service constraints that correspond to service requirements, and a monitoring model that covers five kinds of system events relevant to client request, service response, application, resource, and management, and a monitoring framework in which different probes and agents collect events and data that are sensitive to requirements. The framework analyzes the collected information against the prespecified constraints, so as to evaluate the behavior and use of Web services.

2 Proposed Service Selection Approach

Genetic Algorithms (GAs) are adaptive heuristic search algorithm based on the evolutionary ideas of natural selection and genetics. As such they represent an intelligent exploitation of a random search used to solve optimization problems. Although randomised, GAs are by no means random, instead they exploit historical information to direct the search into the region of better performance within the search space. The basic techniques of the GAs are designed to simulate processes in natural systems necessary for evolution, specially those follow the principles first laid down by Charles Darwin of "survival of the fittest.". Since in nature, competition among individuals for scanty resources results in the fittest individuals dominating over the weaker ones [14].

GAs simulate the survival of the fittest among individuals over consecutive generation for solving a problem. Each generation consists of a population of character strings that are analogous to the chromosome that we see in our DNA. Each individual represents a point in a search space and a possible solution. The individuals in the population are then made to go through a process of evolution [15].

GAs are based on an analogy with the genetic structure and behavior of chromosomes within a population of individuals using the following foundations:

- Individuals in a population compete for resources and mates.
- Those individuals most successful in each 'competition' will produce more offspring than those individuals that perform poorly.
- Genes from 'good' individuals propagate throughout the population so that two good parents will sometimes produce offspring that are better than either parent.
- Thus each successive generation will become more suited to their environment.

A population of individuals are maintained within search space for a GA, each representing a possible solution to a given problem. Each individual is coded as a finite length vector of components, or variables, in terms of some alphabet, usually the binary alphabet [16]. To continue the genetic analogy these individuals are likened to chromosomes and the variables are analogous to genes. Thus a chromosome (solution) is composed of several genes (variables). A fitness score is assigned to each solution representing the abilities of an individual to 'compete'. The individual with the optimal (or generally near optimal) fitness score is sought. The GA aims to use selective 'breeding' of the solutions to produce 'offspring' better than the parents by combining information from the chromosomes.

GA is very suitable to be used to fulfill web service selection task for web service composition problem. To apply GA to web service selection, two important issues should be addressed: genetic encoding of web service selection and the definition of the fitness function.

An individual in the population of genetic algorithm represents a web service selection plan and it is encoded in an array of n integers $x_1, x_2 \dots x_n$, where n is the total number of web service classes in the workflow of the composite web service. In the gene encoding scheme, each genetic represents a web service class in the composite web service and a value of the gene represents a concrete web services of the web service class. If the gene is 1, then the web service is selected for service composition. If it is 0, then this service is discarded.

In this paper, in order to promote the probability to create different paths from the mutated path, the probability of mutation is for the chromosome instead of the locus. The concrete policy is as follows: before mutation operation of every chromosome, the probability of mutation is used to confirm whether the chromosome mutates or not. If mutation, the object path will be confirmed firstly whether it is the same as the current path expressed by the current chromosome. If difference, the object path will be selected from all available paths except the current one. If the object is itself, the new chromosome will be checked whether the new chromosome is the same as the old chromosome. The same chromosome will result in the mutation operation again. If the objects are different paths from the current path, a new chromosome will be

created on the basis of the object path. Obviously, it is not necessary to check whether the new and old chromosomes are the same.

3 Performance Evaluation

In this experiment, we evaluate the quality of the results obtained. We mainly compare the optimal degree of results. Fig. 1 and Fig. 2 show the results on the optimal degree. The results above show that the optimal degree of our proposed approach is larger than 90% regardless the number of web services in this experiment. This means that our approach is very effective in service selection process.

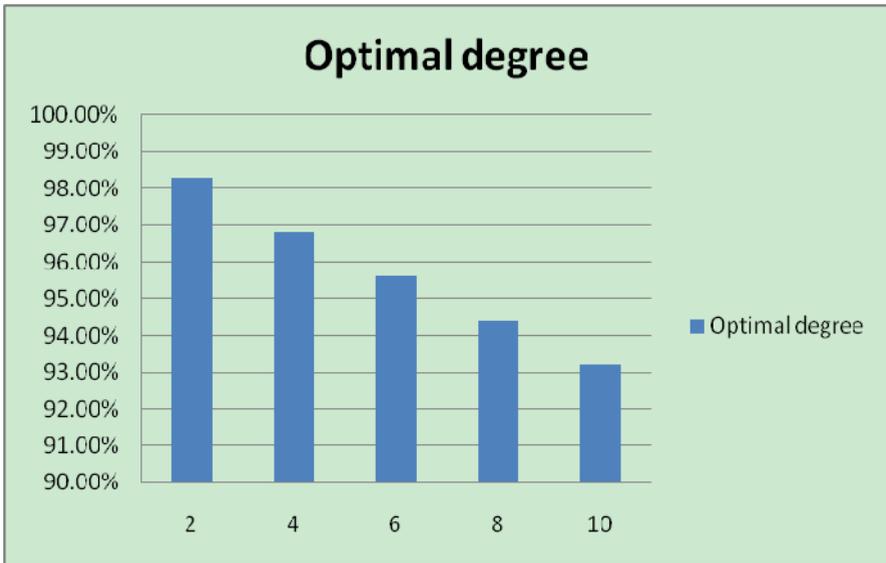


Fig. 1. The optimal degree with different service classes

For example, from Fig. 1, the optimal degree of our approach is larger than 95% on average. This means that our approach can perform the service selection with high optimal degree, i.e., service users can accurately find the most appropriate service regardless of the number of service classes. Although with the increasing number of service class, the optimal degree is decreasing, its optimal degree is still larger than 90%. Furthermore, from Table Fig. 2, the optimal degree of our approach is larger than 91% on average. This means that our approach can perform the service selection with high optimal degree, i.e., service users can accurately find the most appropriate service from many concrete services. From the experimental results, by means of the approach, service users can find the most web service from a great number of services according to their QoS requirements. Hence, our approach is effective for web service composition problem.

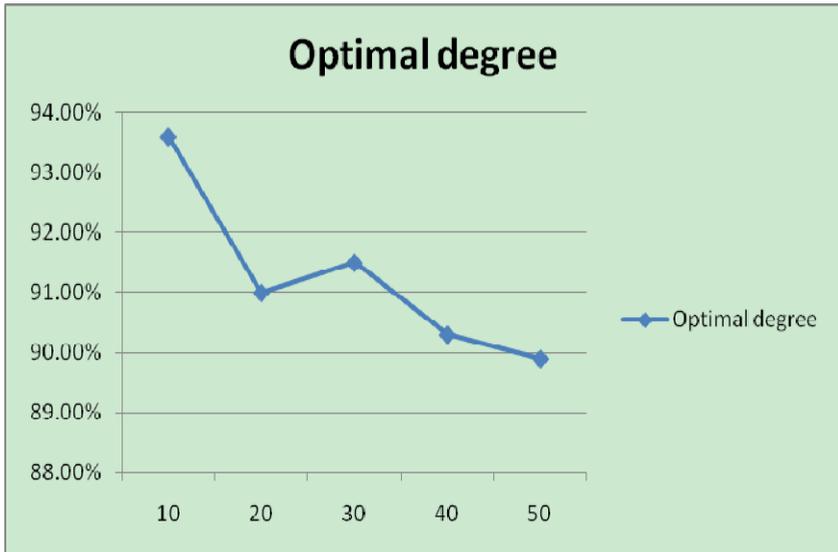


Fig. 2. The optimal degree with different service candidates

4 Summary

In this paper, we propose a QoS-aware web service selection approach for solving web service composition problem with a great number of web services. In our proposed approach, we use genetic algorithm to find the most suitable web service for service users. In order to evaluate our proposed approach, we conduct experiments to verify our approach. In the experiment, we adopt a real QoS dataset from world internet. Experimental results indicated that our proposed approach significantly improve the web service selection process in web service composition problem.

References

1. Wang, H.-C., Lee, C.-S., Ho, T.-H.: Combining subjective and objective QoS factors for personalized web service selection. *Expert Systems with Applications* 32, 571–584 (2007)
2. Hu, C.-H., Wu, M., Liu, G.-P., Xu, D.-Z.: Approach to constructing Web service workflow based on business spanning graph. *Ruan Jian Xue Bao/Journal of Software* 18, 1870–1882 (2007)
3. Qi, Y., Bouguettaya, A.: Computing Service Skyline from Uncertain QoWS. *IEEE Transactions on Services Computing* 3, 16–29 (2010)
4. Hwang, S.-Y., Lim, E.-P., Lee, C.-H., Chen, C.-H.: Dynamic Web service selection for reliable Web service composition. *IEEE Transactions on Services Computing* 1, 104–116 (2008)
5. Fan, X.-Q., Jiang, C.-J., Wang, J.-L., Pang, S.-C.: Random-QoS-aware reliable web service composition. *Ruan Jian Xue Bao/Journal of Software* 20, 546–556 (2009)

6. Shao, L.-S., Zhou, L., Zhao, J.-F., Xie, B., Mei, H.: Web service QoS prediction approach. *Ruan Jian Xue Bao/Journal of Software* 20, 2062–2073 (2009)
7. Liu, S.-L., Liu, Y.-X., Zhang, F., Tang, G.-F., Jing, N.: Dynamic web services selection algorithm with QoS global optimal in web services composition. *Ruan Jian Xue Bao/Journal of Software* 18, 646–656 (2007)
8. Lin, W.-L., Lo, C.-C., Chao, K.-M., Younas, M.: Consumer-centric QoS-aware selection of web services. *Journal of Computer and System Sciences* 74, 211–231 (2008)
9. Baresi, L., Di Nitto, E., Ghezzi, C., Guinea, S.: A framework for the deployment of adaptable web service compositions. *Service Oriented Computing and Applications* 1, 75–91 (2007)
10. Chakraborty, D., Jaiswal, S.K., Misra, A., Nanavati, A.A.: Middleware architecture for evaluation and selection of 3rd party web services for service providers, Orlando, FL, United states, pp. 647–654 (2005)
11. Zheng, Z., Lyu, M.R.: A QoS-aware fault tolerant middleware for dependable service composition. In: 2009 IEEE/IFIP International Conference on Dependable Systems and Networks, Lisbon, Portugal, pp. 239–248 (2009)
12. Li, Z., Yang, F., Su, S.: QoS-aware semantic web service composition with uncertainties. *Chinese Journal of Electronics* 17, 703–709 (2008)
13. Wang, Q., Shao, J., Deng, F., Liu, Y., Li, M., Han, J., Mei, H.: An online monitoring approach for Web service requirements. *IEEE Transactions on Services Computing* 2, 338–351 (2009)
14. Zhang, C., Su, S., Chen, J.: DiGA: Population diversity handling genetic algorithm for QoS-aware web services selection. *Computer Communications* 30, 1082–1090 (2007)
15. Zhang, C.-W., Su, S., Chen, J.-L.: Genetic algorithm on web services selection supporting QoS. *Jisuanji Xuebao/Chinese Journal of Computers* 29, 1029–1037 (2006)
16. Funk, C., Schultheis, A., Linnhoff-Popien, C., Mitic, J., Kuhmunch, C.: Adaptation of composite services in Pervasive Computing environments. In: IEEE International Conference on Pervasive Services, Istanbul, Turke, pp. 242–249 (2007)

Development Strategy of the Digital Library in University Based on SWOT Analysis

Lijuan Yang

Zhejiang Gongshang University, Hangzhou, China
l.y.now@qq.com

Abstract. Coming with the time of electronic information and network, the traditional mode of service function of library cannot meet far the diverse needs of readers of students. In this state, it has become an inevitable trend to establish a digital library with collection of network and electronic publications in the integrative multi-functional modern. In this paper, SWOT analysis on the strength, weakness, opportunity and threat of the digital library in universities in China, the development strategy of the digital library in universities is obtained. The results are useful for the digital library in universities in China to develop healthily.

Keywords: Digital library, SWOT analysis, development strategy, university.

1 Introduction

The development and application of the Internet, particularly the rise of the information highway in the global scope, led to the changes of literature carrier, they will impact inevitably traditional status of printed publications in the library. With the establishment of the digital library, readers can browse all the books in the Internet by using their own computer. Digital library or virtual library plays a huge advantage, is far better than the traditional library [1].

From the definitions of digital libraries by William Arms [2], we can know, "An informal definition of a digital library is a managed collection of information, with associated services, where the information is stored in digital formats and accessible over a network. A crucial part of this definition is that the information is managed. A stream of data sent to earth from a satellite is not a library. The same data, when organized systematically, becomes a digital library collection. Most people would not consider a database containing financial records of one company to be a digital library, but would accept a collection of such information from many companies as part of a library. Digital libraries contain diverse information for use by many different users. Digital libraries range in size from tiny to huge. They can use any type of computing equipment and any suitable software. The unifying theme is that information is organized on computers and available over a network, with procedures to select the material in the collections, to organize it, to make it available to users, and to archive it."

Digital library began to be studied the earliest in the United States, and then to be taken seriously in Britain, France, Japan and even the entire international community.

Many countries have spending heavily on their research and development of digital libraries. In recent years, Singapore also has invested heavily to the development of the related work of digital library, trying to establish the centre of Chinese information online.

The research of digital library in China started from tracking foreign relevant project. In bursting with "electronic library", "paperless society", "CD ROM" and other new theory and new things, the digital library began to formulate gradually in 1994. During the development of the digital libraries of universities in China, The digital library of Tsinghua University walked in the fore front of the construction of domestic digital libraries. By using the electronic resources of the library, it not only can provide online CD database retrieval, reports of new technology magazines, information query of degree thesis, stand-alone CD (abstract indexed database, full-text database), EI Village, OCLC First Search, electronic journals ordered by library, etc., and by using the electronic resources outside of the library, provide Science Online, Dialog, OCLC First Search, abstract indexed database, patent, standards, technical reports, electronic journals, electronic meetings, the web sites of important association /association, virtual library, Internet journals and so on many retrieval service.

According to their own conditions some university libraries in China develop tentatively some personalized information services. They provide the following service to main users: custom service, information delivery service, link services, bookmarks functional service and so on [3].

In short, with huge storage capacity, fast speed, saving long time, low cost, convenient exchange and so on, the digital library will be a great promotion on the construction of domestic various university libraries. However, a digital library is still a kind of new concepts and new technology; there exist many unsolved puzzles in technology. Especially the construction of the digital libraries of the university is a large-scale systematic engineering. It contains many factors, such as the person's factor, technological factors, and economic factors. And its construction needs the joint efforts and collaboration of many libraries, and the whole society even. Therefore, for the healthy development of the digital library of the university, it is needed to apply SWOT analysis, consider comprehensively the internal advantages and disadvantages, and the external opportunities and threats, and choose reasonably development strategies.

2 SWOT Analysis on the Development of the Digital Library in Universities

2.1 Strength

Compared with the traditional library, the digital library creates a new environment of information exchange. Using new information processing technology of literature resources digitization, the digital library can form a rich content of digital information resources including books, periodicals, maps, manuscript, audio-visual, etc. For the huge amounts of information resources today the digital library provides the splurge on the orderly organization, complete preservation and efficient utilization.

Network environment provides the basis on sharing resource and joint construction among libraries. Compared with the traditional library, through the network a digital libraries link into mesh structure including users, libraries and information resources, improve greatly the opening and sharing degree of libraries, and can achieve the purpose sharing existing collection. Compared with the traditional library, the breadth and depth sharing information is unmatched, the digital library reaches the sharing resources in the true sense.

The libraries of university pay attention to interaction with users universally. According to actual situation, the libraries of university in China set up different forms of virtual reference service. This shows that the university libraries through the virtual reference service as an important service, pay attention to develop the interaction between university library and the users, achieve the interaction with the users better.

2.2 Weakness

The change of service mode. Compared with the traditional library, digital library in the management mode, operation mode, collection mode and service pattern will change qualitatively.

Digital library requires more money generally, is also long cycle of construction. The construction of the digital library is a huge task, and requires a lot of funds. Due to many professional features shared in various universities, it will cause the repetition and waste work if every university made electronic material by itself.

The universities in China are lack of joint virtual reference service. The cooperating consultation system achieved remarkable results in USA. And the most typical is the cooperation virtual reference consulting service launched jointly by the U.S. national library and OCLC. The Question Point service is the most popular cooperating consultation system at present. It provides an interacting tool in variety ways with users using real-time reference and E-mail. Among them the real-time consultation provides trustworthy one-on-one assistance by professional librarians on the real-time. Due to development later, the key of the virtual reference consulting system in the university libraries in China is being built; their joint virtual reference in universities is still in the early stage.

The level of interaction in libraries of university is low. The virtual reference service of university library in China has a feature that and the face-to-face communication haven't developed maturely with the reference librarians or experts, just limited to simple contact list with the reference librarians, only in the interactive level of information distribution. A few university libraries set an appointment with the subject librarian in China; the ratio is too low.

Traditional librarians are not competent for new work requirements. Traditional librarians age structure are too large widespread, the knowledge structure are low, the technology content of operation process are not high, cause them to accept the ability of modern electronic equipment relatively weak, therefore, they cannot fully qualified for work needs in electronic information age.

2.3 Opportunity

Digital library compensates for the defects of traditional library. Limited to development scale, site, environment, service and other factors, traditional university library primarily in printed books and literature already cannot satisfy with the requirements of the student readers; and borrowing business are gradually atrophic. With the popularization of computer and network utilization, people are more and more interest in online data access, search for information, reading articles. With network and information the digital library will be inexhaustible.

The information technology taking network as the center has laid a technology foundation for the development of digital library. Along with the development of computer technology, communication technology, electronic storage technology and network technology, especially the opening of the Internet, digital library arises at the historic moment, and become the future developing direction of the library.

In recent years, various universities are successively built campus network, and realize the mutual connect with the Internet. In order to adapt to the requirement of library in network era, many universities put the agenda on the construction of digital library. Some universities first step, has gained some experience.

2.4 Threat

Insufficient capital investment, lead to development fund shortage. The establishment of the digital library needs a lot of the funds to buy computer equipment, installation network facilities, etc. Although much money is input to purchase books and maintenance in university every year, but relative to the entire development of digital library, is only part of a few. It is a big challenge to invest once to reach the designated position [4].

Problems of standardization and compatibility of resource construction. The standardization and compatibility are the keys to realize the coordinated organization of universities, interworking of networks, resources sharing and management orderly. The standardization and standardization of resources construction are the premise and basic guarantee to realize of electronic resources sharing, and will lay the foundation for the national digital resources sharing in the future.

The safety questions of digital library. At present the network security is more and more serious problem, digital library also is the problem of network security. The security and reliability of digital library is the fundamental guarantee of running smoothly. Therefore an important work in the plan of digital library should ensure safety on digital library. As part of network the safe problems of the digital library displays mainly in the operating system security, database security and the web server security, etc. Because digital library is based on the public Internet, except based on public Internet safety and reliability of the service, it is needed for the digital library to provide safe and reliable support itself.

The intellectual property rights of digital library. Due to the development of computer technology and the powerful function of network transmission, in digital library there was a tremendous change of the emergence and spread, preservation and

obtaining information, the author's intellectual property interests are a serious threat. Meanwhile, the regulations and measures in these respects by the law have formed certain restrictions to the continued development of digital library. How to coordinate the contradiction between the development of digital library and intellectual property also becomes a problem to be solved.

3 Development Strategies of the Digital Library in Universities

According to earlier SWOT analysis on development of the digital library in universities, SWOT matrix can be appropriate, as shown in Table 1. In this matrix, we can make the development strategies and the specific measures of the digital library in universities [5-10].

Overall planning and accurate positioning on the construction of the digital library. It is the key for a university to plan the construction of digital library. The planning implementation of digital library should in line with the elaborate organization, result-oriented spirit, overall planning, rational layout, separate construction, sharing, with the fastest speed and the smallest input, and achieve the best effect. First, the construction of the digital library of the university involving wide range and many problems, only strengthen the leadership and unify management, so that make its construction to be able to develop smoothly. Secondly, in the planning process it is necessary for the digital library to locate accurately, to make digital library vitality, utilization and output. When orientating must solve the electronic direction and key of the library, make a characteristic and avoid the similar.

Constantly develop using new network information resources. The construction of network information resource is the core content of the construction of digital library. Through digital processing various vector information resources can storage in CD or hard disks and other media, and form a collection system of digital information resources. Therefore, only strengthen to develop and use the network information resources and technology, it will be possible to ascend the service quality and technology level of digital library of the university.

Digital library should change the service mode. First, digital library should change the service mode from the simple type service mode to varied type service modes. The library should take the user as the center and provide what users need, such as agent retrieval, agent inquires, online catalog inquiry, online project information service, etc. The digital library should transform from the literature material collectors and providers into information product manufacturer, developer and provider.

Secondly, digital library should transform from shallow level service mode to deep level service mode. Information society needs the deep processing, library processes information from the literature unit into the knowledge unit mainly in the deep sense. Library service work transforms also from borrow and return service to the multi-level information consultant service. Information service personnel have from simple labor steering intellectual labor.

Table 1. SWOT matrix of development of the digital library in universities

| | | | |
|-----------------|---|--|--|
| | | Strength(S) | Weakness (W) |
| | | <ul style="list-style-type: none"> ● Create a new environment of information exchange ● Network environment provides the basis on sharing resource and joint construction among libraries ● Pay attention to interaction with users universally | <ul style="list-style-type: none"> ● The change of service mode ● Require more money generally, long cycle of construction ● Lack of joint virtual reference service ● Low level of interaction in libraries of university ● Traditional librarians are not competent for new work requirements |
| Opportunity (O) | <ul style="list-style-type: none"> ● Compensate for the defects of traditional library ● The information technology taking network as the center has laid a technology foundation ● Various universities are successively built campus network | SO Strategy (Growth Strategy) | WO Strategy (reverse-type strategy) |
| | | <ul style="list-style-type: none"> ● Overall planning and accurate positioning ● Constantly develop using new network information resources | <ul style="list-style-type: none"> ● Adhere to the principle of digital resources sharing ● Establish and perfect the incentive mechanism of librarian ● Complete the talent cultivation and reserve work |
| Threat (T) | <ul style="list-style-type: none"> ● Insufficient capital investment, lead to development fund shortage ● Problems of standardization and compatibility of resource construction. ● The safety questions of digital library ● The intellectual property rights of digital library | ST Strategy (diversification strategy) | WT Strategies (defensive strategy) |
| | | <ul style="list-style-type: none"> ● Change the service mode ● Note the standardization and compatibility of resource construction | <ul style="list-style-type: none"> ● Ensure the safety and reliability of digital library ● Perfect the policy and regulations of digital intellectual property |

Adhere to the principle of digital resources sharing. During the planning of digital library Universities should adhere to the basic principle of digital resources sharing. According to the uniform standard each university library should build respective characteristic collection, undertake all or part of the database of a project subject or the characteristic database construction by using local characteristic history culture or subject resource advantages; and all libraries joining in the construction can use the database built.



Meanwhile, universities should establish a digital library cooperation committee in a certain area or a province. With large and medium-sized university library as the center, it realizes the reasonable management to each university library fund and the effective utilization of human resources. It makes full use of modern high-tech means, to consider overall global idea, to do well the division of each school, to avoid duplication and to improve the service efficiency of funds.

Note the standardization and compatibility of resource construction. To achieve high share of digital resource, the construction of digital resources of universities should be in uniform standard. Digital library standards involve very wide range, including data format standard, literature markup standards, data exchange standard, service standards, information retrieval standards, inter-library loan standard etc. Universities should consider the international standard and the information resources compatibility in China when to select equipment or data standard, so as to facilitate the long-term development of the digital library of the university.

Perfect the policy and regulations of digital intellectual property. In order to adapt to the development requirements of international society, it must discover in time and adjust the copyright policy, reintegrate loopholes in various interests, and actively develop the information legislation, perfect the intellectual property system, make the information industry development on track. It is very important to strengthen intellectual property protection and ensure the healthy development of the digital library. It needs the builders of the digital library from the law, technology, and consciousness, cooperation and management, etc., and adopts effective coordination measures to realize the vision.

Ensure the safety and reliability of digital library. The security of digital library can use the security management mechanism of operating system to prevent external malicious attacks to digital library. To legitimate users, through the user accounts and passwords can confirm the user's identity. Meanwhile, the user access should be controlled. The system reliability should guarantee system long-term and stable operation. The data can restore when destroyed, and also the system recovers rapidly and operate normally.

Establish and perfect the incentive mechanism of librarian. Librarian of the library is the most important human resource; excellent librarian is the most important resources of library and wealth. The actual role of librarian will play a decisive role in the whole library management. Therefore, it is the fundamental problem to arouse the librarians' enthusiasm, initiative and creativity in maximum limit through effective method. To achieve this goal, the first measure to be taken is to establish and perfect the incentive mechanism of librarians, thus to stimulate their dedication and sense of responsibility, and eventually to complete effectively the mission and goals of the library.

Complete the talent cultivation and reserve work to adapt the development of digital library. Compared with the traditional library, digital library put forward higher request on the librarians. Due to the long-term of the construction of digital library, the university libraries should prepare to cultivate a librarian team with information requirements. First, in order to adapt to the development of the digital

library, the role of librarians should also change into a navigator and information expert who can help the public obtain information. Secondly, librarian as the worker of professional information should have various skills including information classification, organization, and retrieval and gaining, understand all kinds of information network system function more than readers, and should have rich practical experience, can help and guide users to use the new information environment of digital library. Librarians should not only have the basic knowledge, the management and maintenance skills of computer hardware and software, and the retrieval ability of information, also the stronger information processing ability and scientific research ability by learning and improving their own business level constantly.

Therefore, the library in university should grasp training librarian to adapt to the need of digital library as an important work. From the following two aspects: first, the library should introduce the high-level talents; second, the library should strengthen to educate the existing library staff and meet the need of working environment continuously.

4 Conclusions

With the popularization of information in university library, the digital library is the direction of development in the information age in future. It is the important guarantee for a university to survive in the fierce competition. In this paper, SWOT analysis on the strength, weakness, opportunity and threat of the digital library in universities, the development strategy of the digital library in universities is obtained. The results are useful for the digital library in universities in China to develop healthily.

References

1. Li, M.: submitted to China Science and Technology Information (2005)
2. Information on, <http://home.wlu.edu/~whaley/classes/digilib/whaley/definition.html>
3. Chu, J.: Journal of Yan'an Vocational & Technical Institute Vol 24, 103 (2010)
4. Wang, X.: submitted to China Education Outside (2009)
5. Zhang, J.: Journal of Changchun Normal University (Natural Science) 29, 152 (2010)
6. Yin, M.: submitted to Inner Mongolia Science Technology & Economy (2009)
7. Liu, W.: submitted to Journal of Jiamusi Education College (2011)
8. Chen, S.: submitted to Science & Technology Information (2010)
9. Wang, T.: submitted to China Science and Technology Information (2010)
10. Zhou, H.: Journal of Wuhan University of Technology (Social science edition) 23, 779 (2010)

Quality Evaluation on Troops Transport Capacity Based on Interval Grey Numbers Multi-attribute Decision-Making Method

Lu Bangjun¹ and Dai Xiliang²

¹ Dept. of Transportation Command

² Section of Basic Courses, Automobile Management Institute, Bengbu, China
luu_luu@sina.cn

Abstract. taking quality evaluation on troops transport capacity as example, decision problems as to the standard interval grey number index value for the interval grey number are studied, and entropy weight is used as the main means to calculate the entropy value of each properties in order to determine the weight and to present scientific evaluation methods for multi-attribute decision-making problems.

Keywords: troops transport capacity quality, multi-attribute decision-making, interval grey numbers, entropy method, Grey entropy.

1 Introduction

Multi-attribute decision making, under the condition of considering the case of multiple properties, designs the best alternative programs or rearranges programs order. In solving multiple attribute decision making problems, how to determine the attribute weights is very important. At present, a number of valuable results concerning the determination of weight coefficient have been achieved, and how to use entropy to determine weight of the evaluation index is already explained in detail in document [1]. In actual practice, however, the program's complexity and ambiguity of indexes always make the resulting index values tend to be expressed by interval numbers in problem evaluation. In other words, Evaluation of the index value changes in a certain interval, but it is rather than a fixed value. Currently, there are no sufficient researches on determining the weight of intervals [3-6]. First, interval grey number is converted into standard interval grey number [2] (i.e. the sum of white number [3] and a variable), then according to the entropy principle, the weight of every evaluation index is to be determined to provide decision makers with useful information.

2 Standard Interval Grey Number

Definition 1 (Standard Interval Grey Number). Given that an interval grey number can be expressed as shown in the following form of the equation, and in the formula, g_i is known as the white section of G_i , and

$$G_i = g_i + c_i \gamma_i \quad i = 1, 2, \dots \tag{1}$$

$c_i \gamma_i$ is termed as the grey section of G_i , thereunto, c_i is called grey coefficient, γ_i is the unit of grey (or grey unit), then grey form expressed formula (1) is called standard interval grey number, or simply the standard grey number.

Theorem 1. (standard expression grey number) any interval grey number,

$$G_i \in [a_i, b_i], a_i \leq b_i, i = 1, 2, \dots$$

can be expressed as the standard definition of form 1 $G_i = g_i + c_i \gamma_i \quad i = 1, 2, \dots$

Definition 2 (standard interval grey number operation rules) [2]. If algebraic operation $F(g) = g(G_1, G_2, \dots, G_n)$ of any standard interval grey numbers can still apply the classic rules, and taking a number $\gamma_i, \gamma_i \in [0, 1], i = 1, 2, \dots$ of the these grey numbers as a constant, the we can get the minimum and maximum:

$$\min\{F(g)\} = \min g(G_1, G_2, \dots, G_n) | \gamma_i = c_i, c_i \in [0, 1], i = 1, 2, \dots, n$$

$$\max\{F(g)\} = \max g(G_1, G_2, \dots, G_n) | \gamma_i = c_i, c_i \in [0, 1], i = 1, 2, \dots, n$$

The computing process endpoint values around the left and right of $\{F(g)\} = g(G_1, G_2, \dots, G_n)$ is respectively known as the standard interval grey number operation process.

3 Entropy Method and Its Basic Principle

Entropy method is a mathematical method synthesizing in consideration of various factors on the basis of information provided to calculate a composite index. As an objective evaluation method, it is mainly based on the indexes of the amount of information passed to the decision makers to determine the size of their weights [4]. Entropy was originally a thermodynamic concept first introduced into information theory by the Shannon C.E.; it is widely used now in the engineering, socio-economic and other fields. According to the basic principles of information theory, information is a measurement of the degree of system order, while entropy is the measurement of degree of the system disorder, the two absolute values are equal, but opposite in symbol.

The basic principle of entropy method is: supposing study objects as a set of m , denoted by $Q = \{Q_1, Q_2 \dots Q_m\}$. Evaluating indexes are a set of n , denoted by $I = \{I_1, I_2 \dots I_n\}$, and with raw data measured. Supposing the actual measurement of the original data matrix as:

$$R^i = (r_{ij}^i)_{m \times n} \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n. \quad r_{ij}^i \text{ is the score of sample } j \text{ on}$$

the indexes i if the indexes are measured in different units or different

measurements. To standardize R is a must in order to obtain a standardized score matrix of each index. After considering the data by standardizing is affected r'_{ij} , $\min|r'_{ij}|$ and $\max|r'_{ij}|$, extremum method is used to standardize the original data. Given the original data matrix, $R' = (r'_{ij})_{m \times n}$, and the standardized matrix is $R = (r_{ij})_{m \times n}$, then the specific standard formula is:

$$r_{ij} = \frac{r'_{ij} - \min|r'_{ij}|}{\max|r'_{ij}| - \min|r'_{ij}|} \tag{2}$$

After the raw data is standardized, the information entropy of each index can be calculated. The entropy H_i of index i can be defined as:

$$H_i = -k \sum_{j=1}^n f_{ij} \ln f_{ij} \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad , \quad \text{of} \quad f_{ij} = \frac{r_{ij}}{\sum_{j=1}^n r_{ij}} \quad ,$$

where k is a constant.

After the entropy value is confirmed, and then entropy weight w_i of index i can be determined according to the following formula:

$$w_i = \frac{1 - H_i}{m - \sum_{i=1}^m H_i}$$

As can be seen from the above basic principle, the smaller information entropy H_i of a certain index, the greater variation of the value of the index it indicates, and the greater amount of information it can provide, the bigger role it plays in synthesizing evaluation and the greater weight it is. Therefore, in the specific analysis of the process, entropy can be used to calculate each index value according to the degree of variation of each index. In conducting more quantitative evaluation indexes, the determination of index weight is particularly important, for, it relates to the accuracy evaluation of the results.

4 Interval Grey Methods in Determining Entropy Weight in Multiple Attribute Decision

Supposing that an established program number to be evaluated is m , denoted by $Q = \{Q_1, Q_2 \dots Q_m\}$, the interval number \otimes_{ij} denotes the evaluation value of program Q_i to I_j of index j , then the index value mn of the program m

constitutes a matrix $A = [\otimes_{ij}]_{m \times n}$, which is known as the evaluation matrix of program sets to index sets.

Evaluation index by its attribute is usually divided into two categories: the cost-based type and efficiency-based type. The index value of the former is the one which the smaller it is, the better it becomes, while index value of the latter is the one which the greater it is, the better it becomes. In general, different evaluation index tends to have different dimensions and dimensional units, in order to eliminate their Incommensurability during multi-index evaluation, non-dimensional treatment must be conducted in evaluation index so as to get the standardized matrix. So far as the number of index system for the white (white number) is concerned, by equation (2) mode can be standardized while for the interval grey number, different methods of standardization are used.

For efficiency-based indexes, supposing

$$\begin{cases} r_{ij}^L = \frac{a_{ij}^L}{\sum_{i=1}^n a_{ij}^U} \\ r_{ij}^U = \frac{a_{ij}^U}{\sum_{i=1}^n a_{ij}^L} \end{cases}, \quad i = 1, 2, \dots, n \tag{3}$$

For efficiency-based indexes, given

$$\begin{cases} r_{ij}^L = \frac{1/a_{ij}^U}{\sum_{i=1}^n (1/a_{ij}^L)} \\ r_{ij}^U = \frac{1/a_{ij}^L}{\sum_{i=1}^n (1/a_{ij}^U)} \end{cases}, \quad i = 1, 2, \dots, n \tag{4}$$

The basic steps to multiple attribute decision making evaluation of interval grey number of are as follows:

The first step: changing interval grey number decision making matrix into the form of a standardized grey matrix.

By (2), change the evaluation matrix $A = [\otimes_{ij}]_{m \times n}$ into $a = [g_{ij} + c_{ij}\gamma_{ij}]_{m \times n}$.

Step Two: Calculating the entropy of each index

The entropy of \tilde{H}_i of index i can be defined as:

$$\tilde{H}_i = -k \sum_{j=1}^n \tilde{f}_{ij} \ln \tilde{f}_{ij} \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad \text{in which}$$

$$\tilde{f}_{ij} = \frac{g_{ij} + c_{ij}\gamma_{ij}}{\sum_{j=1}^n g_{ij} + c_{ij}\gamma_{ij}}, \quad k = \frac{1}{\ln n}, \quad 0 \leq \gamma_{ij} \leq 1, \quad \text{the maximum value } \tilde{H}_i \text{ of}$$



\tilde{H}_i^U and the minimum value of \tilde{H}_i^L can be obtained by calculating, that is: $\tilde{H}_i \in [\tilde{H}_i^L, \tilde{H}_i^U]$.

Step three: Calculate the weight of composite index grey entropy weight is obtained by the above formula; the interval entropy weight of attribute I_j of j is as the following:

$$\tilde{w}_j = \frac{1 - \tilde{H}_j}{n - \sum_{j=1}^n \tilde{H}_j}, j = 1, 2, \dots, n \tag{5}$$

Algorithms[8] based on the interval, the above equation can be further written as

$$\left\{ \begin{aligned} w_j^L &= \frac{1 - H_j^U}{n - \sum_{j=1}^n H_j^L} \\ w_j^U &= \frac{1 - H_j^L}{n - \sum_{j=1}^n H_j^U} \end{aligned} \right., j = 1, 2, \dots, n \tag{6}$$

Step four: program sorting procedure changing $\tilde{w}_j \in [w_j^L, w_j^U]$ into a standard interval grey number $\tilde{w}_j = w_j^L + (w_j^U - w_j^L)\gamma_j$, then each unit is calculated by evaluating the value of synthesizing evaluation

$$z_i = \sum_{j=1}^n [w_j^L + (w_j^U - w_j^L)\gamma_j](g_{ij} + c_{ij}\gamma_{ij}) \tag{7}$$

$i = 1, 2, \dots, m; j = 1, 2, \dots, n$

Calculating the maximum values z_i^U and minimum values z_i^L of z_i , that is, $z_i \in [z_i^L, z_i^U]$, sorting the interval number z_i in the synthesizing evaluation of interval numbers according to the comparison of standard interval grey value given in the document [2], concrete steps are not described here.

5 Experimental Analyses

This paper from three aspects makes comparative evaluation on four military units transportation quality and capacity, the purpose of which is to provide a evaluation train of thought for the troops, including: I_1 -vehicle quality; I_2 -driving quality;

I_3 - synthesizing training quality. The specific attribute values of each transportation unit's capacity are in the below.

| Differentiation | I_1 | I_2 | I_3 |
|-----------------------|---------|---------|---------|
| Transportation unit 1 | [20,30] | [40,60] | [60,80] |
| Transportation unit 2 | [30,40] | [30,50] | [45,55] |
| Transportation unit 3 | [25,30] | [40,70] | [60,70] |
| Transportation unit 4 | [25,35] | [45,60] | [40,60] |

Changing interval grey number decision matrix into the form of a standard grey matrix.

| Differentiation | I_1 | I_2 | I_3 |
|-----------------------|----------------------|----------------------|----------------------|
| Transportation unit 1 | $20 + 10\gamma_{11}$ | $40 + 20\gamma_{12}$ | $60 + 20\gamma_{13}$ |
| Transportation unit 2 | $30 + 10\gamma_{21}$ | $30 + 10\gamma_{22}$ | $45 + 10\gamma_{23}$ |
| Transportation unit 3 | $25 + 5\gamma_{31}$ | $40 + 30\gamma_{32}$ | $60 + 10\gamma_{33}$ |
| Transportation unit 4 | $25 + 10\gamma_{41}$ | $55 + 5\gamma_{42}$ | $40 + 20\gamma_{43}$ |

Calculate the entropy weight of each factor.

| Differentiation | I_1 | I_2 | I_3 |
|-----------------------|---|---|--|
| Transportation unit 1 | $\frac{20 + 10\gamma_{11}}{100 + 10\gamma_{11} + 10\gamma_{21} + 5\gamma_{31} + 10\gamma_{41}}$ | $\frac{40 + 20\gamma_{12}}{165 + 20\gamma_{12} + 10\gamma_{22} + 30\gamma_{32} + 5\gamma_{42}}$ | $\frac{60 + 20\gamma_{13}}{205 + 20\gamma_{13} + 10\gamma_{23} + 10\gamma_{33} + 20\gamma_{43}}$ |
| Transportation unit 2 | $\frac{30 + 10\gamma_{21}}{100 + 10\gamma_{11} + 10\gamma_{21} + 5\gamma_{31} + 10\gamma_{41}}$ | $\frac{30 + 10\gamma_{22}}{165 + 20\gamma_{12} + 10\gamma_{22} + 30\gamma_{32} + 5\gamma_{42}}$ | $\frac{45 + 10\gamma_{23}}{205 + 20\gamma_{13} + 10\gamma_{23} + 10\gamma_{33} + 20\gamma_{43}}$ |
| Transportation unit 3 | $\frac{25 + 5\gamma_{31}}{100 + 10\gamma_{11} + 10\gamma_{21} + 5\gamma_{31} + 10\gamma_{41}}$ | $\frac{40 + 30\gamma_{32}}{165 + 20\gamma_{12} + 10\gamma_{22} + 30\gamma_{32} + 5\gamma_{42}}$ | $\frac{60 + 10\gamma_{33}}{205 + 20\gamma_{13} + 10\gamma_{23} + 10\gamma_{33} + 20\gamma_{43}}$ |
| Transportation unit 4 | $\frac{25 + 10\gamma_{41}}{95 + 10\gamma_{11} + 10\gamma_{21} + 5\gamma_{31} + 10\gamma_{41}}$ | $\frac{55 + 5\gamma_{42}}{165 + 20\gamma_{12} + 10\gamma_{22} + 30\gamma_{32} + 5\gamma_{42}}$ | $\frac{40 + 20\gamma_{43}}{205 + 20\gamma_{13} + 10\gamma_{23} + 10\gamma_{33} + 20\gamma_{43}}$ |

$$\begin{aligned}
 H_1 = & -\frac{20 + 10\gamma_{11}}{100 + 10\gamma_{11} + 10\gamma_{21} + 5\gamma_{31} + 10\gamma_{41}} \ln \frac{20 + 10\gamma_{11}}{100 + 10\gamma_{11} + 10\gamma_{21} + 5\gamma_{31} + 10\gamma_{41}} \\
 & -\frac{30 + 10\gamma_{21}}{100 + 10\gamma_{11} + 10\gamma_{21} + 5\gamma_{31} + 10\gamma_{41}} \ln \frac{20 + 10\gamma_{11}}{100 + 10\gamma_{11} + 10\gamma_{21} + 5\gamma_{31} + 10\gamma_{41}} \\
 & -\frac{25 + 5\gamma_{31}}{100 + 10\gamma_{11} + 10\gamma_{21} + 5\gamma_{31} + 10\gamma_{41}} \ln \frac{25 + 5\gamma_{31}}{100 + 10\gamma_{11} + 10\gamma_{21} + 5\gamma_{31} + 10\gamma_{41}} \\
 & -\frac{25 + 10\gamma_{41}}{100 + 10\gamma_{11} + 10\gamma_{21} + 5\gamma_{31} + 10\gamma_{41}} \ln \frac{25 + 10\gamma_{41}}{100 + 10\gamma_{11} + 10\gamma_{21} + 5\gamma_{31} + 10\gamma_{41}}
 \end{aligned}$$

Solving H_1 with MATLAB software, its maximum value H_1^U and minimum value H_1^L is obtained, of which $H_1^U = 0.9993$

Similarly: $H_2^L = 0.9771, H_2^U = 0.9955; H_3^L = 0.9706, H_3^U = 0.9996$

Computing method of synthesis weight. According to equation (6), \tilde{w}_1 is obtained, the interval of \tilde{w}_2 and \tilde{w}_3 is:



$$[0.9251, 4.4392], [0.5856, 4.0205], [0.5530, 5.1749]$$

To facilitate the calculation, standardizing entropy weight, the normalized formula is:

$$\tilde{w}'_j = \frac{\tilde{w}_j}{\max(w_j^U)}, j = 1, 2, \dots, n \quad \text{the result is: } \tilde{w}'_1 = [0.1788, 0.8578] ,$$

$$\tilde{w}'_2 = [0.1132, 0.7769] , \tilde{w}'_3 = [0.1069, 1]$$

Programs sorting. According to (7) the calculated evaluation values of synthesizing evaluation of the objects are as the following:

$$z_1 = [14.51, 152.35] , z_2 = [13.57, 128.57] , z_3 = [17.42, 150.12] ,$$

$z_4 = [13.84, 136.64]$ According to the introduction of interval number comparison

offered in document [2], the sorting of z_1, z_2, z_3, z_4 is $z_3 \succ z_1 \succ z_4 \succ z_2$, that is: Transportation unit 3 \succ transportation unit 1 \succ transportation unit 4 transportation unit 2.

6 Conclusions

In this paper, the standard form of interval grey number of interval grey number is changed into the white one with the number of variables, which the calculation and comparison among the grey numbers can be realized through the use of the classical mathematics, thus, it to overcomes the decision-making problems due to the existence of matrix, which is difficult to compare the size of interval grey number of index and hard to calculate. Therefore, such decision-making problems as quality sorting of evaluation objects is identified, and an effective method to solve interval grey number decision-making problems is provided through the size comparison between standard interval grey numbers. And the weight concerning decision-making problems based on interval grey numbers are calculated.

References

1. Xia, Y., Wu, X.: A technique of order preference by similarity to ideal solution for hybrid multiple attribute decision making problems. *Journal Of Systems Engineering* 09(6), 142–144 (2004)
2. Fang, Z.: Grey Interval Characterization and arithmetic improving in GM(1,1) Modely. *Journal of Engineering Science* 7, 58–61 (2005)
3. You, T.: A Method of the Entropy Power On Uncertain Multiple Attribute Decision Making. *Journal of DongBei University* 25, 598–601 (2004)
4. Wang, Z., Wang, C., Zhang, J.: The Multi-object Weights Evaluation Method Based on Ideal Interval Number and Entropy. *Chinese Journal of Engineering Mathematics* 23, 13–19 (2006)
5. Zhu, F.X., Chen, H.: A method of entropy for obtaining the attribute weights of interval numbers decision - making matrix. *Fire Control & Command Control* 30, 4–6 (2006)

6. Liang, C., Wu, J., Lu, W., Ding, Y.: A New Method on Hybrid Multiple Attribute Decision-Making Problem for Choosing the Supplier. Chinese Journal of Management Science 14, 71–76 (2006)
7. Liu, S., Dang, Y.: Quantitative Analysis Method and Application. BeiJing science and Publishing House (2004) (in Chinese)
8. Xu, Z.: Uncertain Multi - attribute Decision- making Method and Application. BeiJing Tsinghua Publishing House (2004) (in Chinese)

Explorations of and Reflections on the Educational Model of Software Schools

Jingfeng Yan and Guoqing Li

School of Computer, Xuchang University, Xuchang, Henan, 461000, P.R. China
li_1i2626@yeah.net

Abstract. The Based on an analysis of the current situation of both the demand in China for software professionals and the software school of xuchang University, this paper proposes some new schemes for the disciplinary development of software technology. With the "major plus company plus professional training" disciplinary development model as well as the "project-driven, case-based teaching, and integrated classroom" teaching model, a professional team and a propeller-style curricular system will be established, a base will be developed for practice teaching, and teaching quality supervision will be improved to train professionals who meet the requirements for the application of the software discipline and serve the purpose of regional economic and industrial development of xuchang. And the performance of the students who finished their studies at university over the past two years proves that the schemes proposed by this paper have a significant effect on the training of software professionals.

Keywords: List the case-based teaching, university-industry collaboration, software schools.

1 Demand Analysis of Disciplinary Development

The Specialized Plan for informatization in the Eleventh Five-Year Plan for National Economic and Social Development states that it is strategic measures of modernization to vigorously promote the informatization of national economy and society. In the "Eleventh Five-Year Plan" period of China, the total demand for all manner of information professionals increased by 20 million among which the demand for software professionals increased by 3 million[1]. However, the annual number of software graduates is about 500 thousand, thus the huge gap between supply and demand. In the "Training Program for Urgently Needed Professionals in the Manufacturing Industry and Modern Service Industry" jointly launched by the Ministry of Education and other five ministries, the major of computer software technology was listed as one of the four urgently needed majors.

2 Service to the Regional Economy of Xuchang

As a software school undertaken by the only undergraduate institution of xuchang, it aims to train applied professionals of production, technology, service and

management for the businesses and all walks of life throughout the city. In view of the demand for software professionals in xuchang and the neighboring cities[2][3], xuchang University has been establishing the major of software technology from 2009 to train software professionals and serve the purpose of regional economic and industrial development.

3 Current Problems

The freshly established Software School is affiliated to the School of Computer Science and Technology, xuchang University. And it has the following three problems:

1. The Software School is currently staffed by 40 teachers, and most of them have to deliver lectures at the School of Computer Science and Technology as well as the Software School, thus causing a severe shortage of teaching staff of the Software School.
2. Currently, most teachers of the Software School have to deliver lectures in two majors, thus the problem of transformation in teaching approaches from undergraduate curricula into software school curricula.
3. Some teachers lack the experience from practical project development, thus causing difficulties for the practice teaching and students' professional training.

4 Guiding Principles and Objectives of Disciplinary Development

The Software School of xuchang University adheres to the scientific development concept and the educational principle of "job-oriented and competency-based education", aims to serve xuchang and the society to improve in a comprehensive manner the market relevance of the software discipline as well as the quality and employment rate of graduates and to meet the demand for software professionals from economic development in the 21st century.

5 Training of Applied Software Professionals with Great Coding Competency

The major of software technology aims to train applied professionals engaged in software development such as software programmers, software testers, technical service and management personnel, sales and promotion personnel. And it also aims to enable students to acquire the fundamental knowledge and skills of software development on the basis of gaining necessary theoretical and professional knowledge and become the "software blue-collar" with good knowledge application and problem-solving ability, innovation and sustainable development capacity, professional ethics and integrity.

6 Establishment of the Practice Teaching

The plan is aimed at developing the practice teaching base of software major into the first-class regionally shared software training base which integrates multiple functions such as teaching, training, product development and technical innovation with increased investments and standardized management on the basis of the existing practice teaching base within three years, according to the principles of real environment, serial functions, enterprise management, advanced equipment and professional staff.

7 Training Framework

Software outsourcing talents' cultivation method is undergoing major transformation, through the curriculum and teaching method reform and innovation make students with the knowledge and skills to meet the needs of outsourcing enterprise. Many software engineering software talents training of scholars and workers will be teaching way from the traditional way of Learning to Lecture primarily LBI - Lecture Based (Instruction) to case studies Based on problems mainly promotion training (PBL - Learning) way bought Based/transformation.

(1) based on the problem of training (PBL)

Traditional training mode focuses on the knowledge learning, PBL emphasize ability. Table 1 gives two training difference of method[4].

Table 1. Traditional training method based on problems with the training methods

| | |
|---|---|
| Traditional methodologies | Problem-based learning |
| Teaching centered | Student centered |
| Linear and rational | Coherent and relevant |
| Teacher as transmitter | Instructor as facilitator.active participants |
| Stuctured environment | Flexible environment |
| Individual and competitive learning | Co-operative learning |
| Assessment is the responsibility of the teacher | Assesment is the shared responsibility of the Students, the group and the teacher |

PBL is aimed at a software development activities and puts forward corresponding problems setting up the background material, teachers inspire students to find out the methods to solve the problems, this is a kind of interactive teaching methods, exercise the student ability, judgment, to solve the question ability, from the actual case is also developing projects.

Outsourcing talents training can be combined with LBI and PBL, realize the knowledge and skill training of dual goal.

(2) the Capstone project training (Projects)

Project practice middle school students play a major role in a practical project, and student teams make use of knowledge proposes a set of complete solution and put into practice, teacher's role is more like coach, provide advisory services. Through the

project training students exercise your communication skills, and ability of organization, self-study ability and collaboration capabilities, decision-making ability and the ability to solve problems. Project from real life or a software enterprise, which can will service outsourcing enterprise post training go up to the campus completed, help graduates work directly mount guard.

(3) cooperation projects (Co - op bought)

Co - op is a software enterprises and institutions of talent training, effective cooperation way, particularly suitable for ordering cultivation. It has three advantages: one is the on-demand training, students before graduation can sign employment contract with the enterprise or enterprise employees sent to specific skill training colleges; Second, the school and enterprise establish long-term cooperation relations, but part ease teaching shortage, through the double system, can will project development experiences to students, to improve the students' ability to solve engineering problems; Three is the enterprise can provide programs for students to practice, solve the problem of insufficient school training project.

Through effective training methods, advanced can improve the engineering ability, outsourcing talent to help the students to study the knowledge systematically into the ability to solve practical problems.

8 Building of a First-Class Professional Team "Excelling in Both Software Development and Teaching"

By the end of 2011, the major of software technology shall be staffed by 20 teachers, with the number of double-qualified teachers accounting for 90%, and 67% of the full-time teachers shall hold a master's degree and an intermediate or above technical title[4]. There shall be two-way exchanges between the software teachers of our school and the excellent technical personnel of the software industry, and the proportion of full-time teachers to part-time ones shall be equal. One or two pace-setters in this major shall be trained, so that a first-class professional team with good teaching ethics, performance, quality and structure can be formed under the leadership of the pace-setter.

9 Establishment of Curricular System Which Is "Seamlessly Connected" with Jobs

The curricular system of software technology shall follow the developmental trend of software industry in a timely manner. The curricular system shall be developed according to the requirements from posts (post groups) of the software technology field on the basis of relative professional qualification standards[5]. Since "we learn for the purpose of using", the content of courses shall be closely connected with the requirements for skills from relative posts of the current software industry such as programmers, so that the content of courses can be "seamlessly connected" with job competence.

10 Conclusion

With the adoption of the teaching reform scheme proposed in this paper for the first crop of two-year junior college students majoring in software technology, the performance of this crop of students proves that positive results have been achieved. Students' software development skills which would lay a solid foundation for software development have been greatly improved, and most students have secured employment with ease after graduation, hence this paper also provides valuable experience for the teaching reform of other majors of the software school.

References

1. Cui, H.: An Analysis of the Developmental Trend of Software Industry of China. Industrial Technology & Economy (2009)
2. Yang, F.: Knowledge Economy and Software Industry. Science & Technology Industry of China (2006)
3. Cui, H.: Study on the Development and Countermeasures of Software Industry in China. Doctoral Thesis of Jilin University (2005)
4. Zhang, Y.: Reflections on the Establishment of Exemplary Software School. Higher Education Press (2009)
5. Gao, L.: Software Elite Were Thus Formed. Higher Education Press (2010)

The Research and Contrast of the Hybrid Intrusion Detection

Meng JianLiang¹ and Yu Yang²

¹ North China Electric Power University he bei sheng bao ding shi hua dian lu
689 hao hua bei dian li da xue 358 xin xiang

² North China Electric Power University he bei sheng bao ding shi hua dian lu
689 hao hua bei dian li da xue 358 xin xiang

w.l.l_so@qq.com

Abstract. This paper firstly does some analysis on the system based on data mining and explains the preliminary concept and the characteristics of the data mining, then it analyses the system based on ummune theory According to the principle and the characteristic of the immune theory,we put it into the intrusion detection system and propose the instrusion system model which is mixed with the immunology. At last,this paoer introduces the typical hybrid intrusion detection system and the new hybrid intrusion detection system based on the immune algorithm.

Keywords: Data mining, Mixed intrusion detection, Immune algorithm.

1 The Typical Hybrid Intrusion Detection System

This paper describes a typical hybrid intrusion detection system. It uses both the technologies of misuse detection and anomaly detection which includes three sub modules: The misuse detection module,the anomaly detection module and the signature generation module.Its architecture shown in Figure 1:

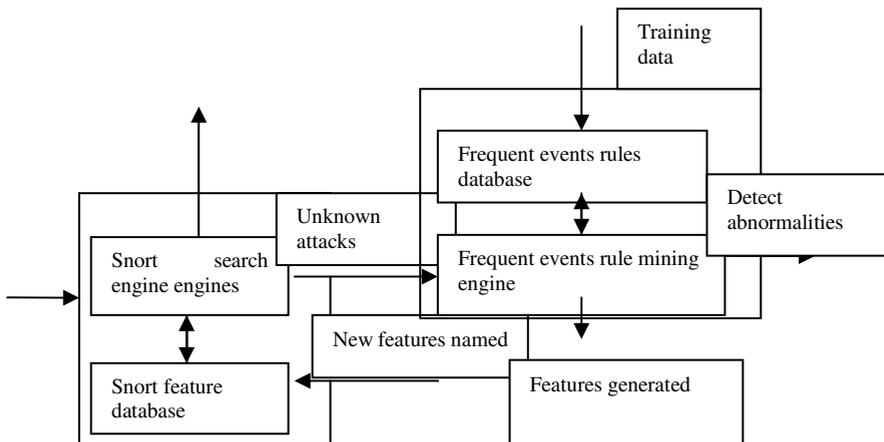


Fig. 1. The structure of hybrid intrusion detection system

The Misuse Detection Module. Using the lightweight of the snort as the core construction and its own rules to establish snort signature database. As snort is flexible ,customizable and extensible, it can easily add its own defined rules which lays a good foundation for the system to do the misuse detection and anomaly detection.

Anomaly Detection Module. The module construction is based on Frequent Episode algorithm, by using association analysis of the data structures and library functions to do sequential pattern analysis and get frequent events ,meanwhile by comparing with the frequent event rule, we can get the current activity is normal or attack.

Signature Generation Module. Signature generation module is responsible for the attack which is detected by anomaly detection module and map rules corresponding to snort, then insert it into snort rules library. Apriori algorithm is the basic algorithm required for mining association rules generated frequent item sets. Apriori algorithm is mainly composed of connections and pruning operation.

Connection: To find L_k , through L_{k-1} and the candidate k-item which is connected with it. The candidate set is denoted as C_k . Let l_1 and l_2 be the entry set in the L_{k-1} . Mark $L_{l_1}[j]$ as the term j in l_1 , For convenience, we assume that the dictionary entry is by order of a centralized order entry. When the L_{k-1} has the same in first $k-2$ entries ,we say that L_{k-1} is reachable, denoted by $l_1 \propto l_2$, then connect l_1 and l_2 , that is when $(l_1 [1] = l_2 [1]) \wedge \dots \wedge (l_1[k-2] = l_2 [k-2]) \wedge (l_1[k-1] < l_2 [k-1])$, the obtained results item set $l_1 [1]l_2 [2] \dots l_1[k-1] l_2 [k-1]$.

$$confidence(A \rightarrow B) = P(A | B) = \frac{sup\ port_count(A \cup B)}{sup\ port_count(A)}$$

Among them, expresses contain item sets of; support_count(A) expresses the number of records for the item sets A.

When to find frequent item sets, you can use them frequently to generate strong association rules, credibility can be calculated.

2 The Intrusion Detection System Based on Hybrid Immune

Technical Overview. The system is the immune system by simulating the performance of antigen invasion to complete intrusion detection and treatment. Important physiological functions of the immune system is "self" and "non-self" antigen recognition and response. Never matured to the mature immune cells, antigen recognition will kill line into immune memory, immune feedback. This is an important immune mechanism in three stages: self tolerance, immune response and immune feedback.

Table 1. The contrast between computer immune system and biological immune system

| Biological immune system | Intrusion Detection System |
|------------------------------|--|
| Antigen | Network Intrusion |
| Antibody | Binary string representation of the detector |
| Self tolerance | Negative selection algorithm |
| Antigen detection / response | Binary string of non-self recognition and response |

Clonal Selection Algorithm. Clonal selection theory is used to explain how the immune system is fighting with the antigen. When the external bacteria or viruses invade the body after, B cells begin to clone and destroy a large number of intruders, who can identify the antigen recognition based on the degree of cell cloning to achieve through the proliferation of purposes.

3 The Comparison and Conclusions of the System

The test data comparison. Experimental data sets takes KDDCUP99 experiments, Table 2 is a typical mix of data mining-based intrusion detection system which detects the rate.

Table 2. The detection rates and testing time of the rule mining algorithm

| Views | The number of records | The number of training rules | Time (sec) | Detected events | Detection Percentage (%) |
|-------|-----------------------|------------------------------|------------|-----------------|--------------------------|
| 1 | 2000 | 125 | 71 | 1914 | 95.7 |
| 2 | 10000 | 1523 | 871 | 9432 | 94.32 |
| 3 | 20000 | 2213 | 1512 | 18436 | 92.18 |

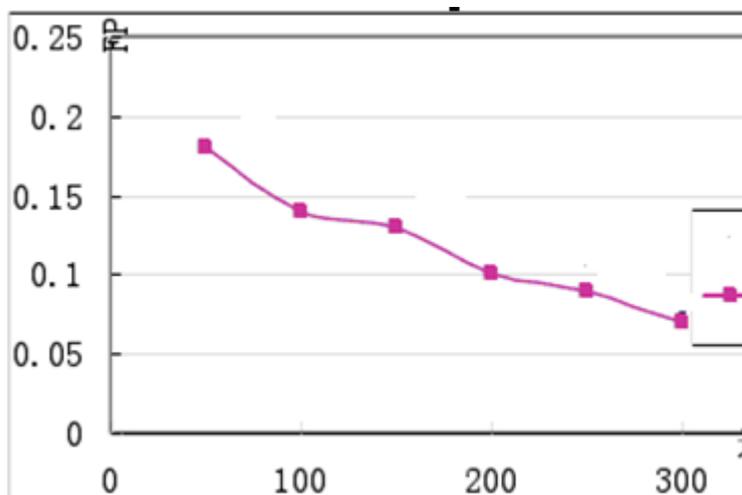


Fig. 2. Detection rates comparison

Figures 2 and 3 are based on immune principle intrusion detection system in the detection rate and false alarm rate changes.

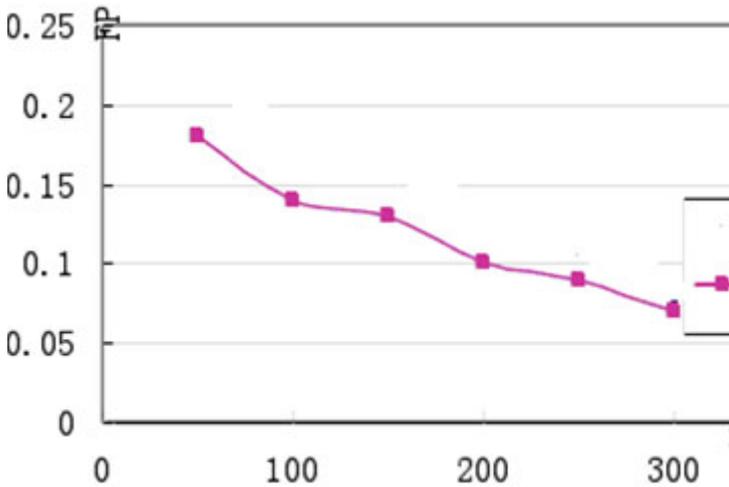


Fig. 3. False alarm rate comparison

Comparative Results. Compared with typical mixed intrusion detection technology, although the mixed immune intrusion detection system which is based on the principle can not reach as the high detection rate and seems not very mature. as the operating algebra increases ,it reflects the immune system of learning and cognitive characteristics. The system for the detection of unknown attacks is beneficial, but the intrusion detection based on artificial immune system can not avoid detection during the build process and a large number of strings Comparison of most inefficient. Data mining-based anomaly detection system also needs to generate rules which can scan the database repeatedly, also, it spends a lot of resources to make real-time data processing system a bit less than the detection which needs longer time. From Table 3, it is visible. the two systems are used to supplement the misuse detection system and the real-time detection system is to enhance the handling capacity of hybrid systems.

Table 3. Comparison of the data detection system

| | BP | SOM | SOM+LVQ |
|------------------|-------|-------|---------|
| Detection rate | 86.1% | 90.7% | 92.5% |
| False alarm rate | 20.9% | 18.1% | 12.0% |

4 Summary

Intrusion detection is the second access control, firewalls and other traditional techniques is after the generation of security protection technology which is a traditional security protection technology and reasonable supplement. However, the intrusion detection technology is currently far from mature. Especially with the direction of the network towards the development and invasion of high-speed means of constantly updated, the intrusion detection is also put forward higher requirements. The

traditional packet capture mechanism and detection engine are difficult to resolve the high-speed packet network environment and the leakage problem of false positives. Real-time intrusion detection system can capture and analyze data, meanwhile the network monitoring system can take notes in specific security data, through the analysis of the characteristics of the rapid detection of dangerous attacks and the alarm, it provides some protection measures.

References

1. Jiang, Z., Wu, T.: The Information Security Principles and Techniques. China Railway Publishing House (May 2009)
2. Wu, C., Zhang, C.: The information Security based on. Wuhan University Press (January 2007)
3. Xin, Y., Xiao, M.: The research and implementation on snort based on hybrid intrusion detection system. In: Proceedings of the Eighth International Conference on Machine Learning and Cybernetics (December 2009)
4. Chen, W., Zhou, J.: The snort lightweight intrusion detection system Raiders. Beijing University Press (July 2009)
5. Yong, Q., Shao, Y.: Intrusion detection. Xi dian University Press (August 2009)
6. Cui, M.: Based on the principle of mixed immune intrusion detection system. Nanjing University of Information Science, Nanjing (2007)
7. Sheng, X.: Based on Immune Intrusion Detection System. Nanjing University of Aeronautics and Astronautics, Nanjing (2008)
8. Ning, P., Reeves, D., Cu, Y.: Correlating alerts using Prerequisites of intrusions. Technical Report TR-2001-13, North Carolina State University, Department of Computer Science, 25 (1):55-63 (December 2005)
9. Zhao, H.: Network Intrusion Detection. Sichuan University 24(12), 13-15 (2003)
10. Yang, X., Liu, Z.: Comparison of Intrusion Detection System. Computer Engineering 30(6), 37-38 (2004)
11. Bai, Y., Kobayashi, H.: Intrusion Detection Systems: Technology and Development. In: Proceedings of the 17th International Conference on Advanced Information Networking and Applications, Lyon, pp. 607-610 (2003)

A New Adaptive Stopping Criterion for BICM-ID System

Xiao Ying, Li Jianping, and Cai Chaoshi

Communication University of China, Beijing, China
86400@home.news.cn

Abstract. This paper devises a new adaptive stopping criterion for BICM-ID system, which applies none-iterative-decoding scheme at low SNRs, SCR stopping criterion at medium SNRs, and CE stopping criterion at high SNRs respectively. Simulation results have verified that the proposed adaptive scheme can not only get excellent BER performances, but also can greatly reduce the average iteration numbers at the same time, which makes it a more optional scheme for BICM-ID system.

Keywords: BICM-ID, Stopping criterion, CE, SCR.

1 Introduction

Bit-interleaved coded modulation with iterative decoding (BICM-ID) can get remarkable performances both in AWGN and in Rayleigh fading channels [1,2]. However, the iterative algorithm increases the complexity at the receiver and leads to much decoding delay. Therefore, the stopping criterions [3-9], which can timely terminate the unnecessary decoding iterations, are put forward, such as the cross-entropy (CE) stopping criterion [3], the sign-change-ratio (SCR) stopping criterion [4], the hard-decision-aided (HDA) stopping criterion [4], the measurement of reliability (MOR) stopping criterion [5], the convolution-sum (CS) stopping criterion [6], the bit-based partial iteration criterion [7], the adaptive variable iterative decoding [8] and various simplified criterions [9] etc. Although some are proposed for turbo receiver, if properly modified, they can also be used for BICM-ID receiver, as the principle is the same. In this paper, we propose a new adaptive stopping criterion for BICM-ID system, which uses the none-iterative-decoding scheme at low SNRs, the SCR stopping criterion at medium SNRs, and the CE stopping criterion at high SNRs respectively. Simulation results have made a validation that the proposed new adaptive scheme can take a better trade off between the average iteration numbers and the BER performances.

This paper is organized as follows: General reviews about the typical CE stopping criterion and SCR stopping criterion are included separately in section 2 and section 3. Section 4 describes the prerequisite of the adaptive scheme. Section 5 detailedly introduces our design of the new adaptive stopping criterion. Finally, section 6 concludes the paper.

2 CE Stopping Criterion

CE (cross-entropy) is a measurement of how close two distributions are. For the distributions p and q of a finite alphabet χ , the CE is defined as

$$H[p, q] = \sum_{x \in \chi} p(x) \log \frac{p(x)}{q(x)} \tag{1}$$

In BICM-ID system, we assume that $L_{DEC}^i(\hat{c}_k)$ denotes the LLR (log-likelihood-ratio) of the decoded bit c_k ($k=1, 2, 3, \dots, N$) after the i_{th} iteration and $L_{DEC}^{i-1}(\hat{c}_k)$ represents the LLR of the $(i-1)_{th}$ iteration respectively. Then, the CE value between the decoding outputs of the two consecutive iterations is

$$\begin{aligned} & H[P_{DEC}^i(\hat{c}_k^i), P_{DEC}^{i-1}(\hat{c}_k^{i-1})] \\ & = P_{DEC}^i(\hat{c}_k^i = 0) \log \frac{P_{DEC}^i(\hat{c}_k^i = 0)}{P_{DEC}^{i-1}(\hat{c}_k^{i-1} = 0)} \\ & + P_{DEC}^i(\hat{c}_k^i = 1) \log \frac{P_{DEC}^i(\hat{c}_k^i = 1)}{P_{DEC}^{i-1}(\hat{c}_k^{i-1} = 1)} \end{aligned} \tag{2}$$

where

$$P_{DEC}(\hat{c}_k = 0) = \frac{e^{L_{DEC}(\hat{c}_k)}}{1 + e^{L_{DEC}(\hat{c}_k)}} \tag{3}$$

$$P_{DEC}(\hat{c}_k = 1) = \frac{1}{1 + e^{L_{DEC}(\hat{c}_k)}} \tag{4}$$

Substituting (3) and (4) into Eq. 2, we get CE:

$$\begin{aligned} & H[P_{DEC}^i(\hat{c}^i), P_{DEC}^{i-1}(\hat{c}^{i-1})] \\ & = \sum_{k=1}^N \left[\frac{L_{DEC}^{i-1}(\hat{c}_k^{i-1}) - L_{DEC}^i(\hat{c}_k^i)}{1 + e^{L_{DEC}^i(\hat{c}_k^i)}} + \log \frac{1 + e^{-L_{DEC}^{i-1}(\hat{c}_k^{i-1})}}{1 + e^{-L_{DEC}^i(\hat{c}_k^i)}} \right] \end{aligned} \tag{5}$$

As the iteration continues, the outputs between the two consecutive iterations i and $i-1$ come to be approximately the same. Thus, the CE value becomes smaller and smaller. When the iteration proceeds to a certain extent, although the iteration number continues increasing, the CE value won't decrease any more. This means that the iteration has reached the decoding limit. Thus, the iterative process can be stopped. We usually set a threshold to stop the unnecessary iterations:

$$T(i) = \sum_{k=1}^N \left[\frac{L_{DEC}^{i-1}(\hat{c}_k^{i-1}) - L_{DEC}^i(\hat{c}_k^i)}{1 + e^{L_{DEC}^i(\hat{c}_k^i)}} + \log \frac{1 + e^{-L_{DEC}^{i-1}(\hat{c}_k^{i-1})}}{1 + e^{-L_{DEC}^i(\hat{c}_k^i)}} \right] < threshold \tag{6}$$

3 SCR Stopping Criterion

The SCR (sign-change-ratio) criterion is a simplified version of CE criterion. From literature [4], we know

$$T(i) \approx \delta_i C(i) \quad (7)$$

where δ_i is an average value, $C(i)$ represents the sign differences in $L_{e2}(\hat{h})$, which means the extrinsic value of the estimated information bit \hat{h} of the second decoder.

We can see obviously from Eq. 7 that $T(i)$ is directly related to the number of sign changes $C(i)$ from iteration $i-1$ to iteration i . The decoding process won't stop until the iteration reaches the inequality $C(i) \leq (0.005 \sim 0.03)N$. Here, N denotes the length of the information bits. The advantage of SCR compared with CE is that the decoder only needs to count the number of sign changes of $L_{e2}^i(\hat{h})$ between the two consecutive iterations, and there is no need to calculate the actual CE value. Obviously, the SCR criterion requires less complex computation.

The SCR criterion introduced above is for turbo receiver. For BICM-ID receiver, as the principle is the same, $L_{e2}(\hat{h})$ only need to represent the extrinsic value of the estimated information bit \hat{h} of the only decoder.

4 The Prerequisite of the New Adaptive Stopping Criterion

We first define the simulation circumstances: BICM-ID system, rate 1/2 convolutional code, 8PSK modulation, SP mapping, Rayleigh fading channel, frame size 2046, maximum iteration 10. In CE scheme, when $T(i) < 10^{-4}T(1)$, stop iteration. In SCR scheme, when $C(i) \leq 0.02N$, stop iteration. We should note that a fixed scheme, which applies 10 as the fixed iteration number is included for comparison.

We can obviously see from Fig. 1 that when SNR is below 3dB, the performance gap between iteration 1 to iteration 10 of the conventional iteration in BICM-ID system is negligible. Therefore, there is no need to conduct the iterative decoding when SNR is below 3dB.

Fig. 2 shows that when SNR is between 3dB and 4.8dB, the BER performance curve of SCR criterion almost coincides with the curve of CE criterion. However, the SCR scheme uses nearly 1 less iteration than CE as Fig. 4 shows. Thus, when SNR is between 3dB and 4.8dB, the SCR scheme is more optional.

Fig. 3 and Fig. 4 make a joint display that when SNR is higher than 4.8dB, although the CE criterion need nearly 1 more iteration than the SCR scheme, it can get distinctly better BER performance as compensation. At high SNRs except the point of 6dB, the CE scheme can get nearly 0.1dB coding gains than SCR. So, the CE criterion is more optional at high SNRs.

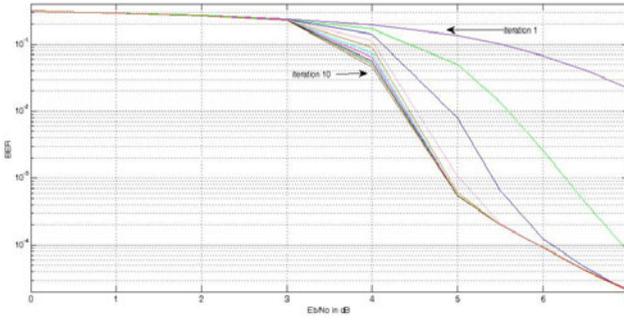


Fig. 1. The conventional iteration from 1 to 10

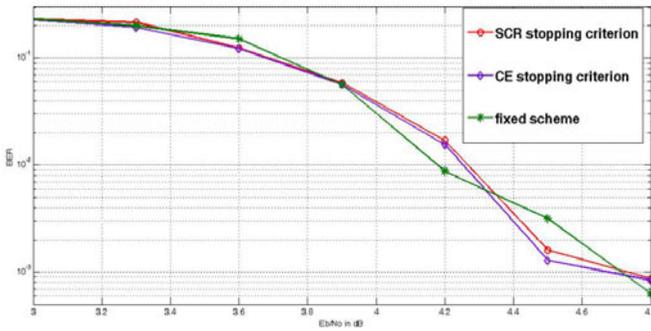


Fig. 2. BER performances at relatively low SNRs

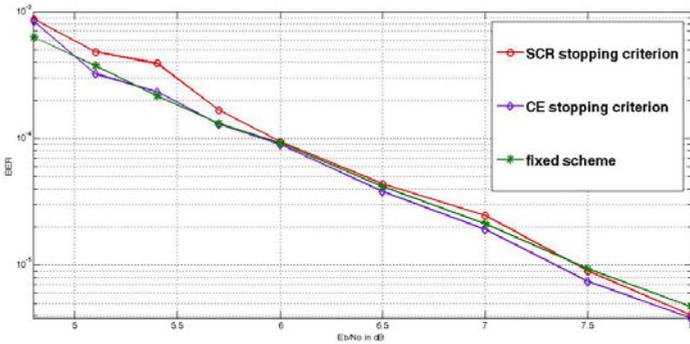


Fig. 3. BER performances at relatively high SNRs

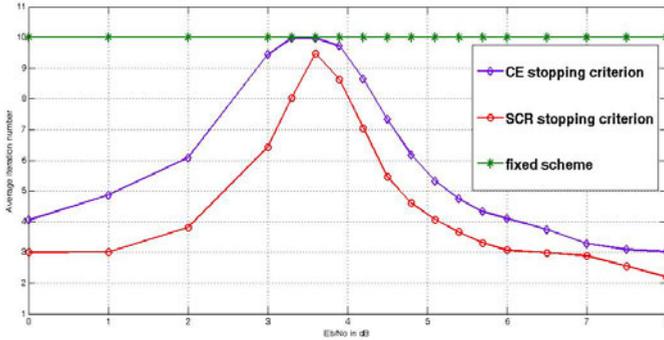


Fig. 4. Comparison of average iteration numbers

5 The Proposed New Adaptive Stopping Criterion

According to the analysis of section 4, we design our adaptive stopping criterion as: When SNR is below 3dB, we don't use iterative decoding, when SNR is between 3dB and 4.8dB, we apply SCR criterion, when SNR is higher than 4.8dB, we use CE criterion instead of SCR. To realize our design, we join a SNR estimator at the BICM-ID receiver as shown in Fig. 5.

Simulation results of Fig. 6, Fig. 7 and Fig. 8 show that the proposed adaptive scheme is superior than others. At low SNRs, without obvious BER performance degradation, the adaptive scheme can greatly reduce the iteration numbers. The iteration number is reduced to only 1 especially when SNR is below 3dB. At high SNRs, the adaptive scheme performs the best both from the BER performance and from the average iteration number points of view. Considering the BER performance, it can get nearly 0.2dB coding gains than the fixed and SCR and can surpass CE at the same time. When it comes to the iteration numbers, take the point of SNR=6dB as example, the adaptive scheme requires only 2.5 iteration, while the SCR, CE and fixed need 3.1, 4.1 and 10 iteration separately.

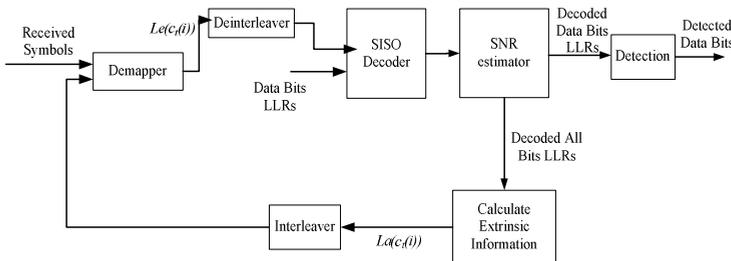


Fig. 5. The receiver of BICM-ID in our adaptive scheme



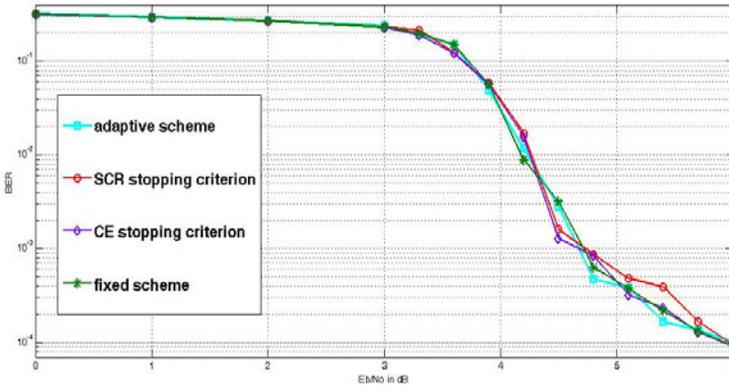


Fig. 6. BER performances at relatively low SNRs

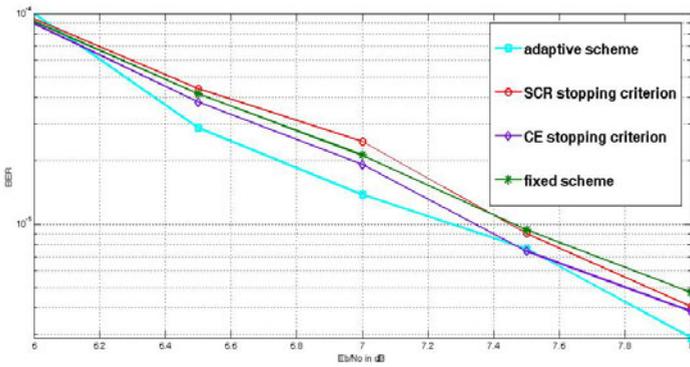


Fig. 7. BER performances at relatively high SNRs

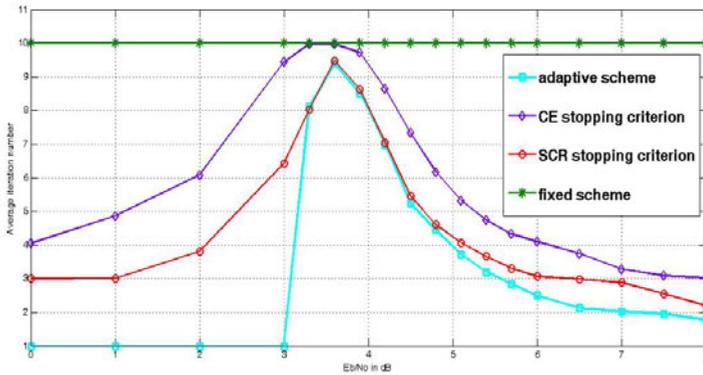


Fig. 8. Comparison of average iteration numbers

6 Summary

This paper proposes a new adaptive stopping criterion for BICM-ID system. It joins a SNR estimator at the receiver to decide whether to use CE or to use SCR or even not use iterative decoding at all. Simulation results have shown that the new adaptive stopping criterion can get excellent performances both from the BER properties and from the iteration numbers points of view. The superior properties make the proposed new adaptive stopping criterion a more desired scheme for BICM-ID system.

References

1. Li, X., Ritcey, J.: Bit-interleaved coded modulation with iterative decoding. *IEEE, Communications Letters* 1(6), 169–171 (1997)
2. Li, X., Ritcey, J.: Bit-interleaved coded modulation with iterative decoding using soft feedback. *IEE Electronic Letters* 34(10), 942–943 (1998)
3. Hagenauer, J., Offer, E., Papke, L.: Iterative decoding of binary block and convolutional codes. *IEEE Trans. Inform. Theory* 42, 429–445 (1996)
4. Shao, R.Y., Lin, S., Fossorier, M.P.C.: Two simple stopping criteria for turbo decoding. *IEEE Trans. Commun.* 47, 1117–1120 (1999)
5. Li, F.-M., Wu, A.-Y.: A new stopping criterion for efficient early termination in Turbo decoder designs. In: *ISPACS 2005*, pp. 585–588 (December 2005)
6. Li, L., Wang, Q., Lu, C.Y.: A novel stopping criterion for Turbo decoding. In: *First International Conference on ICICIC 2006*, vol. 1, pp. 201–205 (2006)
7. Wu, J., Wang, Z., Vojcic, B.: Partial iterative decoding for binary Turbo codes. *IEEE Trans Communications* 57(11), 3298–3306 (2009)
8. Zhang, S., Li, J., Yang, G., Hu, F.: “An adaptive variable iterative decoding scheme for BICM-ID,”. In: *International Conference on CMC 2010*, vol. 2, pp. 282–286 (2010)
9. Zhang, S., Li, J., Chen, J.: “Three simple iterative decoding schemes for BICM-ID,”. In: *International Conference on ICBECS, 2010*, pp. 1–4 (2010)

Data Integration of Heterogeneous Data Source in Multi-parameter Test Processing

Wang Guitang, Liu Wenjuan, and Jiang Yuelong

School of Information Engineering,
Guangdong University of Technology Guangzhou, P.R. China
wangbb33@126.com

Abstract. A variety of specific instruments are used to test the corresponding parameters in the process of multi-parameter test. Function of traditional instrument is simple and it can only test a single type of parameter, therefore we require a variety of instruments. Basic data of different digital instruments bought by companies at different times has large differences in the standards of their structures and amount. Decentralized data storage management, data over redundancy and lack of networks result in worse isolation of information. But the multi parameter test result data of tests analysis and report form need to be integrated, this article uses LabVIEW tool bag LabSQL to realize the visit to Heterogeneous Database and data format transformation of different standards. Then order, event, for circulation and condition structure are used to generate the data preliminary conformity, finally proposed that LabVIEW and the Excel hybrid programming combined to generate the report form printing function. It satisfies requests of integration, uniformity and visit security on the data conformity.

Keywords: multi-parameter test, heterogeneous data source, data integration, LAN communications.

1 Introduction

A variety of specific instruments are used to test the corresponding parameters in the process of multi-parameter test. Digital instruments can generate electronic form data, but the information basic data of different types of digital equipments have larger differences in the standards of their structures and amount. For example, the test results data of the detection system are stored by different means, including the quite different data management system. The simple file database or other text with large amount of data which can be saved as text forms or Access database and complicated network database, they form the heterogeneous data sources.

Although these dispersed data management system of different test equipment, can meet the requirements of data storage and management with a single test result, in many cases, lots of data distributed in different Data management systems of different test instruments need to be visited to analyze some kind of experiment comprehensively. And multi-parameter testing results data need to be integrated in

the test analysis and report form, mainly rely on artificial analysis, fusion and arrangement.

But such question appeared frequently: a unified data interface cannot be provided for a great deal of test results data, universal standards and specifications cannot be used. Therefore, the comprehensive information warehouse based on the whole parameter testing system and the integration of various digital testing instrument heterogeneous data sources is needful in a powerful distributed application system. In addition, the integration of heterogeneous data sources is a real problem in the construction process of IT promotion in enterprises. How to realize data integration of Heterogeneous Database System on the premise of avoiding destroying enterprise original system resources and increasing the additional investment in equipment, has the vital significance and high application value.

This paper studies the access to heterogeneous database and conversion among different standards and specifications of the data format, develop a design method using LabVIEW toolkit of LabSQL to realize the data integration of heterogeneous database in the comprehensive test system of metal materials of an enterprise and the automatic generation of test result report.

2 LabVIEW and the Access to Database

LabVIEW is a tool for virtual instruments' software development which is based on language G and launched by the company NI. We can use LabVIEW to create a virtual instrument system easily, and complete missions such as signal acquisition, data processing and so on, then constitute a whole testing system. LabVIEW has powerful function that can generate reports and communicate with various applications, which makes it the first choice of report generation. Right now LabVIEW has been used in the late data processing of testing system.[2]

There are several methods to access to the database under the environment of LabVIEW[5]:

(1) Take use of the additional SDK LabVIEW SQL Toolkit launched by th company UI to access to the database. However, this SDK is costly that is unaffordable for many LabVIEW users.

(2) Using the DULL program based on other languages such as Visual C++ to access to the database, then access to the program through the DULL interface by LabVIEW, which can connect the database indirectly but has big workload.

(3) Take use of the function of ActiveX by LabVIEW, call the controls of Microsoft ADO and use the language SQL to access to the database. However, this method calls for much knowledge about controls of Microsoft ADO and the language SQL, and complex bottom programming, which is unrealistic for most users.

(4) Using the free database accessing tool LabSQL by LabVIEW. LabSQL is a free, multi-database, multi-platform database-access SDK. LabSQL takes use of the function ActiveX, and calls controls Microsoft ADO, combined language SQL and ODBC interface to access to database. Compared with other methods above, it is the most convenient and effective. When using LabSQL to access to database, it need to be installed and configured. The installation is simple. Unzip the file LabSQL.rar to

the folder LabSQL, then copy the folder LabSQL to the folder user.lib which is under the Installation Directory of LabSQL. When run LabVIEW, the function template of LabVIEW will be loaded to LabSQL automatically.

3 Data Intergating of Heterogeneous Data Source

Access and integration process of heterogeneous database mainly consists of three layers: application layer, data integration layer and data source layer. The operation process can be described as follows: The application layer through a global query access may be formed after converting one or more local query, the query was submitted to the local data sources of various data access interface (API) to complete the final query, the query results are returned to Results synthesizer. Results synthesizer integrated under the appropriate rules and global query statement to query results returned to the application layer after synthesis, so a data inquiries is completed. Figure 1:

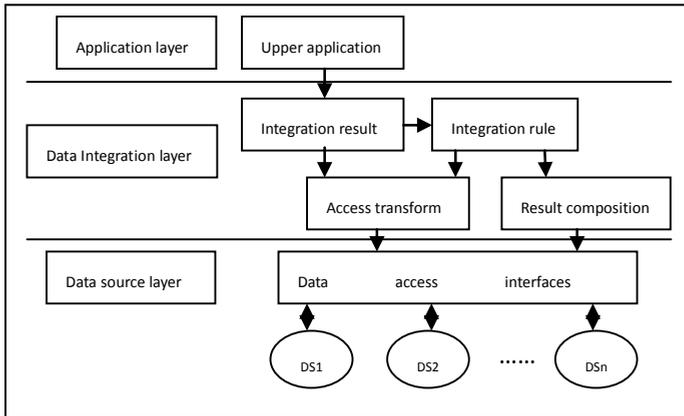


Fig. 1. Data accessing process

"Integration of production, teaching, and research " Hua du Si xiong base material testing laboratory test project has successfully applied this idea to achieve data integration of heterogeneous data bases of two independent detection system-the SXDHXE-7B PC automatic carbon and sulfur analyzer, SXDHXS-3AD PC multi-elemental analysis. The system can access to two different database systems, and make test results on two devices output at the same time. Using free LabVIEW Toolkit LabSQL access to the database, we do not need to know the underlying source code of the instrument development, only the data output ports of instrument testing system could be used, thus greatly reducing the difficulty and workload of software development. Block diagram as follows:

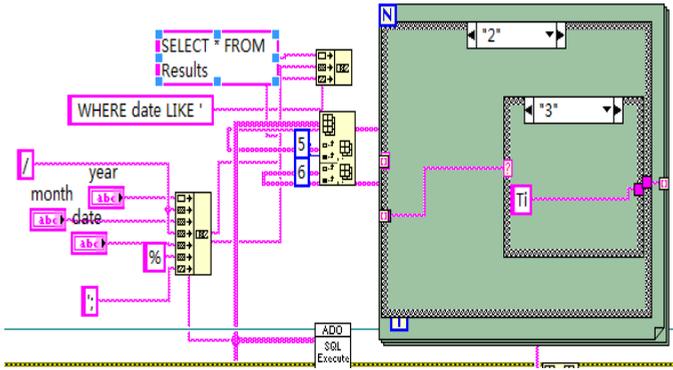


Fig. 2. Database access of SXDHXS-3AD PC multi-elemental analysis

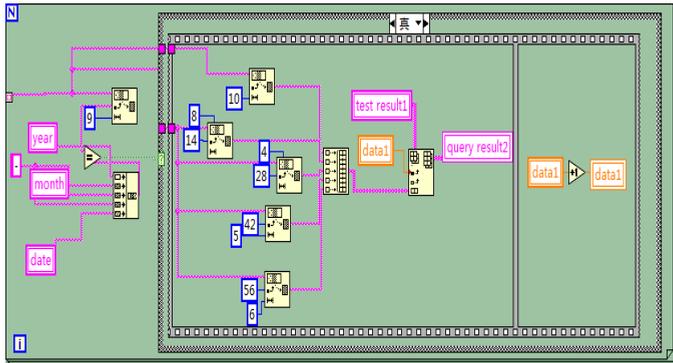


Fig. 3. Database access of SXDHXE-7B PC automatic carbon and sulfur analyzer

| rid | desp name | absorbance | density | charcde | curve nu | date | K | B | R | test eleme | date | time | C | S | SC |
|-----|-----------|------------|----------|---------|----------|--------------------|----------|---------|---------|------------|-----------|----------|-----|------|--------|
| 1 | 11 | 1.20474 | 5.822985 | 1 | 1 | 2009/5/16 11:20:34 | 43.65197 | 5647840 | 9999903 | Cr | 2008-5-16 | 9:35:51 | 194 | 0500 | 0.412 |
| 2 | 11 | 2.666553 | 12.20401 | 1 | 1 | 2009/5/16 11:21:08 | 43.65197 | 5647840 | 9999903 | Cr | 2008-5-16 | 9:56:30 | 470 | 0011 | 0.5735 |
| 3 | 11 | 1.20474 | 5.822985 | 1 | 1 | 2009/5/16 11:21:39 | 43.65197 | 5647840 | 9999903 | Cr | 2008-5-16 | 9:59:59 | 484 | 0028 | 0.8048 |
| 4 | 11 | 1.20474 | 5.822985 | 1 | 1 | 2009/5/16 11:23:58 | 43.65197 | 5647840 | 9999903 | Cr | 2008-5-16 | 10:03:53 | 496 | 0031 | 0.5172 |
| 5 | 11 | 2.66166 | 12.18345 | 1 | 1 | 2009/5/16 11:24:31 | 43.65197 | 5647840 | 9999903 | Cr | 2008-5-16 | 10:08:12 | 588 | 0027 | 0.6572 |
| 6 | 11 | 2327603 | 1.801140 | 1 | 3 | 2009/5/16 13:07:08 | 7.085266 | 151979 | 1 | Mo | 2008-5-16 | 10:14:12 | 620 | 0027 | 0.6650 |
| 7 | 11 | 2322134 | 1.804958 | 1 | 3 | 2009/5/16 13:07:42 | 7.085266 | 151979 | 1 | Mo | 2008-5-16 | 10:17:03 | 613 | 0049 | 0.5689 |
| 8 | 11 | 1.678041 | 1.337373 | 1 | 3 | 2009/5/16 13:08:00 | 7.085266 | 151979 | 1 | Mo | 2008-5-16 | 10:21:57 | 658 | 0027 | 0.4521 |

Fig. 4. Access results of two different databases

Front panel of the complete test system software was shown in Figure 5: its operation is simple and the software development cycle is short. It has not wasted and destroyed original system resources of the company even not increased the extra equipment investment. The system is high stability and strong maneuverability so that it has obtained high praise.



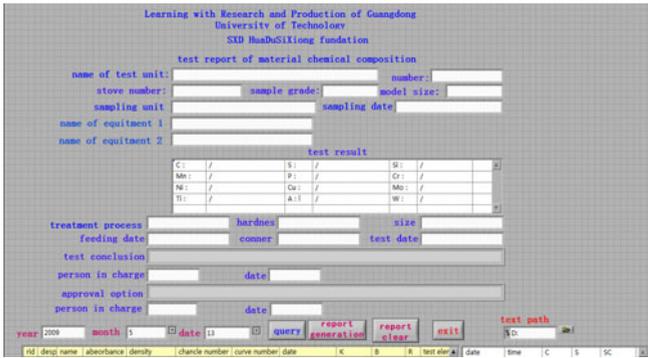


Fig. 5. Front panel of the complete test system software

The basic instrumentation meets the user's needs basically, but for later data mining, integration of treatment, it increases the complexity of the inspectors work, and can't really meet customer needs. In order to be convenient for storage of test results data and the secondary development, Testing channel and curve of multi-element analyzer based on PC SXDHXS-3AD are fixed, which greatly reduces the test work and make test data output to the report successfully. Full report output was shown in figure 6:

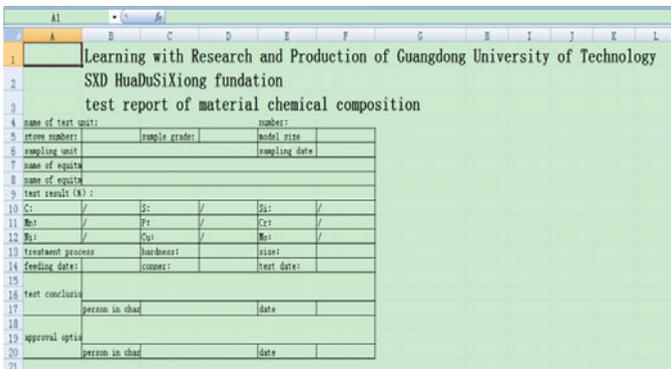


Fig. 6. Full report output

Finally local area network communications is successfully implemented to meet the user needs. The system is posted to the web page which bring greater flexibility and improve efficiency to test staff.

Here are a few applications LabVIEW release notes[2]:

1. The published VI had better be reentrant, so more users can access to it at the same time, or only one user can control and operating procedures at the same time. it prompted only one client has control when you can't be re-seeded.



2. Client needs LabVIEW Runtime Engine, and the browser can support the Labview control.
3. Do not forget to open port 80, otherwise it is impossible to access from other machines.

4 Conclusion

This comprehensive data integration report generation system has been used in material chemical analysis system and greatly reduces the data records burden on the staff. While it also increases the flexibility and reliability of data processing, and make the chart be printed the same as you see it really. The innovation of this paper is: integration of heterogeneous database, and SXDHXS-3AD PC multi-element analyzer test channel number and the curve number is fixed to compensate for the defective design of the instrument and make test data output to the report successfully. A fully functional report generator system is developed. It's simple, and portability of data is good. Meanwhile, the local area network communications is done for making the company achieve intelligent and networking of materials chemical test first in the same industry

References

1. Li, H., Zhao, Y., et al.: LabVIEW Advanced Programming. Tsinghua University Press, Beijing (2003)
2. Hui, C., Zhang, Y. (eds.): LabVIEW 8.20 from entry to the master. Tsinghua University Press, Beijing (2007)
3. Hou, G., Wang, K., Ye, Q., et al.: LabVIEW7.1 programming and virtual instrument design. Tsinghua University Press, Beijing (2005)
4. Hu, S., High, Y., Chotin: The LabVIEW-based Excel Report Generation Technology. Measurement and Control Technology 10, 64–69 (2007)
5. Li, J., Liu, D.X., Zhu, Q.: LabVIEW based on ADO technology to access the database approach. Automation Instrumentation (9), 5–18 (2003)

The Unbounded Parallel-Batching Scheduling Problem with Family Jobs to Minimize Makespan

Jintao Meng and Zhiyong Li

Zhengzhou Institute of Aeronautical Industry Management, Zhengzhou 450015, China
1278887919@qq.com

Abstract. In this paper we consider the single-machine parallel-batching scheduling problem with family jobs under on-line setting in the sense that we construct our schedule irrevocably as time proceeds and do not know of the existence of any job until its arrival. Our objective is to minimize the maximum completion time of the jobs (makespan). A batch processing machine can handle up to B jobs simultaneously. The jobs that are processed together form a batch, and all jobs in a batch start and complete at the same time. The processing time of a batch is given by the longest processing time of any job in the batch. The jobs from different families are incompatible and thus cannot be put in the same batch. We deal with the special variant of the general problem: the unbounded model in which the machine can handle infinite number of jobs simultaneously, the jobs only have two distinct arrival times and come from 2 families. We provide an on-line algorithm with a worst case ratio $1 + \alpha$, where $\alpha = \frac{\sqrt{5}-1}{2}$.

Keywords: Single machine scheduling, Worst-case analysis, Family, On-line algorithm, Parallel-batching scheduling.

1 Introduction

The parallel-batching scheduling problem is motivated by the burn-in operations in semiconductor manufacturing, in which a batch of integrated circuits are placed in an oven and exposed to a high temperature to test their thermal standing ability. The circuits are heated inside the oven until all the circuits are burned. The burn-in times (processing times) of the circuits may be different. When a circuit is burned, it has to wait inside the oven until all the circuits are burned. Therefore, the processing time of a batch of circuits is the processing time of the longest one in the batch.

Motivated by burn-in operation in semiconductor manufacturing, Liu and Yu [10] proved the NP-completeness of this problem even in the case with only two arrival times and provided a pseudo-polynomial time algorithm when the number of arrival times is fixed. Zhang et al. [1] provided on-line algorithms for minimizing makespan on parallel-batching machine. They dealt with several variants involving single or parallel machines, with unbounded or bounded batch size. The basic idea introduced in their algorithms is Delay, in which the current schedule and the information on the

available jobs are used to determine whether processing of a batch will be initiated or postponed. Some other research concerned with on-line batching scheduling can be found in [3] and [4]. Chen et al. [5] considered the on-line scheduling problem that involves assigning jobs to batches and determining the batch sequence in such a way that the total weighted completion times is minimized. They developed a linear time on-line algorithm with a worst-case ratio of $\frac{10}{3}$ for the unbounded model and an efficient algorithm with a worst-case ratio of $(4 + \epsilon)$ for any $\epsilon > 0$ for the bounded model.

In modern manufacturing, another approach from which may result in great benefits is the use of designing and processing similarities of products (jobs) to form product families and processing such families on exclusive machine groups. Scheduling parallel-batching machine with family jobs has been addressed by a number of researchers. Q.Q. Nong and J.J. Yuan et al. [8] proposed an on-line algorithm with a worst-case ratio of 2 for the problem $1|p\text{-batch, on-line, family, } B < n|C_{\max}$. Uzsoy [2] developed efficient algorithms for minimizing the total weighted completion times on a parallel-batching machine with family jobs. Potts and Kovalyov [9] reviewed on the literature on scheduling family job with batching and giving details of the basic algorithms. Azizoglu and Webster [7] considered the parallel-batching scheduling problem with family jobs where the size of each batch is limited and jobs belonging to a common family have identical processing times and arbitrary job sizes. They described a branch and bound procedure applicable to this problem for minimizing the total weighted completion times.

In our study, the given machine is a parallel-batching machine which can process up to infinite jobs simultaneously as a batch. Jobs from different families are incompatible and thus cannot be put in the same batch. All jobs in a common batch start and complete at the same time. In the standard classification scheme, this problem is denoted by

$$1|on\text{-line, } r_0 = 0, r' > 0, family, b = \infty|C_{\max}.$$

And we provide an on-line algorithm with a worst-case ratio $1 + \alpha$.

2 Preliminaries

The quality of an on-line approximation algorithm is usually measured by its worst-case ratio: the smaller the ratio is, the better the algorithm will be. The worst-case ratio of an algorithm is defined to be the smallest constant ρ such that, for any input instance, the algorithm always finds a schedule with an objective value no more than ρ times that of an optimal schedule.

In this paper, let $C_{OPT}(L)$ be the makespan of an optimal off-line schedule for the job list L , and let $C_{H^2}(L)$ be the makespan of job list L obtained by algorithm H^2 . For simplicity, in the following instead of $C_{OPT}(L)$ and $C_{H^2}(L)$, we use C_{OPT} and C_{H^2} to denote the corresponding makespans. Without loss of generality, we can

assume that the earlier arrival time $r_0 = 0$, the other arrival time, denoted by $r' > 0$, is unknown in advance. Let L be the set of all jobs, Let L_1 be the set of jobs arriving at time $r_0 = 0$, and Let L_2 be the set of jobs which will be arriving at r' , let F_i be the set of all jobs in the i -th family.

The jobs from the same family will be processed as a batch. to every family F_i , the job with the large processing time is denoted by j_i , which the processing time is denoted by p_i (the remaining jobs may be deleted and no affect the schedule result). Let T_1 be the total batch processing times of the job list L_1 , and let T and T_2 be the total batch processing times of the job lists L and L_2 , respectively (regardless of arrival time).

Throughout the rest of the paper we always assume that $\alpha = \frac{\sqrt{5}-1}{2} \approx 0.618\dots$.

3 A Lower Bounds

The following lemma can be drawn from the proof of the Theorem 1 in [1].

Lemma 3.1: There exist no on-line algorithm with a worst-case ratio less than $1 + \alpha$ for the batch processing problem $1|on-line, r_0 = 0, r' > 0, family, b = \infty|C_{max}$, where $\alpha = \frac{\sqrt{5}-1}{2}$.

4 Main Results and Proofs

In this section, we consider a special unbounded case where there are only two distinct arrival times and 2 families, the unbounded case where the capacity B is sufficiently large, i.e. all jobs can be processed simultaneously in a single batch. The idea is that we always delay processing of the currently longest batch so that some unexpected long jobs can join this batch. Now we describe the algorithm H^2 as follows.

Algorithm H^2 :

Step 1: At $c = 0$, if there is a idle set between the F_1 and F_2 in the set L_1 , without loss of generality, we assume $F_1 \neq \emptyset$ and $F_2 = \emptyset$, let $p_1 = \max\{p_i, j_i \subseteq F_1\}$.

Step 1.1: If the jobs of the set L_2 arrive at time $r' > \alpha p_1$, the jobs coming from F_1 as a batch will be processed at αp_1 .



Step 1.2: If the jobs of the set L_2 arrive at $r' \leq \alpha p_1$, go to Step 3.

Step 2: At $c = 0$, if there is no idle set between the F_1 and F_2 in the set L_1 , let j_1 and j_2 denote, respectively, the longest processing time jobs of F_1 and F_2 and the processing time be p_1 and p_2 , without loss of generality, we assume $p_1 \geq p_2$.

Step 2.1: The jobs of the F_2 as a batch will be processed at $c = 0$, let $c = p_2$.

Step 2.2: If $c \leq \alpha p_1$, go to step 2.5.

Step 2.3: If the jobs of the L_2 arrive before time c , i.e. $r' \leq c$, go to step 3.

Step 2.4: If $c > \alpha p_1$, the jobs of the L_2 arrive after time c , i.e. $r' > c$, the jobs of the F_1 as a batch will be processed at time c .

Step 2.5: If the jobs of the L_2 arrive before time αp_1 , let $c = \max\{c, r'\}$, go to step 3.

Step 2.6: The jobs of the F_1 as a batch will be processed at time αp_1 , let $c = \max\{(1 + \alpha)p_1, r'\}$.

Step 3: At time c , for the remaining jobs, we divide the two batches according to the family, then schedule the batches in arbitrary order, algorithm end.

Lemma 4.1: If there is a idle set between the F_1 and F_2 in the set L_1 , without loss of generality, we assume $F_1 \neq \emptyset$ and $F_2 = \emptyset$, and the jobs of the L_2 arrive at $r' > \alpha p_1$, then the worst-case ratio of algorithm H^2 is not greater than $1 + \alpha$.

Proof: By the algorithm H^2 , we can divide at most three batches, and $C_{H^2} = (1 + \alpha)p_1 + T_2$.

Case 1: In an optimal schedule, the job j_1 starts before r' . In this case $C_{OPT} \geq p_1 + T_2$. So

$$\frac{C_{H^2}}{C_{OPT}} \leq \frac{(1 + \alpha)p_1 + T_2}{p_1 + T_2} = 1 + \frac{\alpha p_1}{p_1 + T_2} < (1 + \alpha).$$

Case 2: In an optimal schedule, the job j_1 starts at or after r' .

If $T_2 \geq p_1$, then $C_{OPT} \geq r' + T_2 > \alpha p_1 + T_2$. So

$$\frac{C_{H^2}}{C_{OPT}} \leq \frac{(1 + \alpha)p_1 + T_2}{\alpha p_1 + T_2} = 1 + \frac{p_1}{\alpha p_1 + T_2} \leq 1 + \frac{p_1}{\alpha p_1 + p_1} = (1 + \alpha).$$



If $T_2 < p_1$, then $C_{OPT} \geq r' + p_1 > \alpha p_1 + p_1 = (1 + \alpha)p_1$. So

$$\frac{C_{H^2}}{C_{OPT}} \leq \frac{(1 + \alpha)p_1 + T_2}{(1 + \alpha)p_1} = 1 + \frac{T_2}{(1 + \alpha)p_1} < 1 + \frac{p_1}{(1 + \alpha)p_1} = (1 + \alpha).$$

Then $\frac{C_{H^2}}{C_{OPT}} \leq 1 + \alpha$.

Lemma 4.2: If there is a idle set between the F_1 and F_2 in the set L_1 , without loss of generality, we assume $F_1 \neq \emptyset$ and $F_2 = \emptyset$, and the jobs of the L_2 arrive at $r' \leq \alpha p_1$, then the worst-case ratio of algorithm H^2 is not greater than $1 + \alpha$.

Proof: In this case, we have

$$C_{H^2} \leq r' + T \leq \alpha p_1 + T \leq (1 + \alpha)T \leq (1 + \alpha)C_{OPT}.$$

Then $\frac{C_{H^2}}{C_{OPT}} \leq 1 + \alpha$.

At time 0, if there is no idle set between the F_1 and F_2 in the set L_1 , so we assume $r' < \max\{\alpha p_1, p_2\} + p_1$, otherwise algorithm H^2 will generate an optimal schedule with makespan $C_{max} = r' + T_2$.

Lemma 4.3: If $p_2 \leq r' \leq \alpha p_1$, then the worst-case ratio of algorithm H^2 is not greater than $1 + \alpha$.

Proof: In this case, we have

$$C_{H^2} \leq r' + T \leq \alpha p_1 + T \leq (1 + \alpha)T \leq (1 + \alpha)C_{OPT}.$$

Then $\frac{C_{H^2}}{C_{OPT}} \leq 1 + \alpha$.

Lemma 4.4: If $\alpha p_1 \leq p_2 \leq r'$, then the worst-case ratio of algorithm H^2 is not greater than $1 + \alpha$.

Proof: In this case, there is no idle time in the schedule produced by the algorithm H^2 . Then $C_{H^2} \leq p_1 + p_2 + T_2$. In fact, if the starting time p_2 of the batch B_1 is less than r' , then L_1 and L_2 are processed independently, so $C_{H^2} = p_1 + p_2 + T_2$. If the batch B_1 starts to be processed at time r' , i.e., $r' = p_2$, by the algorithm H^2 , the jobs in batch B_1 together with jobs in L_2 are rearranged according to family, obviously, the makespan for arranging $B_1 \cup L_2$ is not larger than the makespan for arranging B_1 and L_2 separately. Thus $C_{H^2} \leq p_1 + p_2 + T_2$.



Obviously, $C_{OPT} \geq \max\{r' + T_2, p_1 + p_2\}$. Because $\alpha p_1 \leq p_2 \leq r'$, we have $p_1 + p_2 \geq (1 + \alpha)p_1$.

If $p_1 \leq T_2$, because $r' + T_2 \geq r' + p_1 \geq p_1 + p_2$, we have $C_{OPT} \geq r' + T_2$. Then

$$\frac{C_{H^2}}{C_{OPT}} \leq \frac{p_1 + p_2 + T_2}{r' + T_2} \leq \frac{p_1 + p_2 + T_2}{p_2 + T_2} = 1 + \frac{p_1}{p_2 + T_2} \leq 1 + \frac{p_1}{\alpha p_1 + p_1} = 1 + \alpha.$$

If $p_1 > T_2$, we have $C_{OPT} > p_1 + p_2$. Then we have

$$\frac{C_{H^2}}{C_{OPT}} \leq \frac{p_1 + p_2 + T_2}{p_1 + p_2} = 1 + \frac{T_2}{p_1 + p_2} < 1 + \frac{p_1}{(1 + \alpha)p_1} = 1 + \alpha.$$

Then $\frac{C_{H^2}}{C_{OPT}} \leq 1 + \alpha$.

Lemma 4.5: If $p_2 < \alpha p_1 \leq r'$, then the worst-case ratio of algorithm H^2 is not greater than $1 + \alpha$.

Proof: By the algorithm H^2 , in this case, we can divide at most four batches, and $C_{H^2} = (1 + \alpha)p_1 + T_2$.

Case 1: In an optimal schedule, the job j_1 starts before r' . In this case $C_{OPT} \geq p_1 + T_2$. So

$$\frac{C_{H^2}}{C_{OPT}} \leq \frac{(1 + \alpha)p_1 + T_2}{p_1 + T_2} = 1 + \frac{\alpha p_1}{p_1 + T_2} \leq (1 + \alpha).$$

Case 2: In an optimal schedule, the job j_1 starts at or after r' .

If $T_2 \geq p_1$, then $C_{OPT} \geq r' + T_2 > \alpha p_1 + T_2$. So

$$\frac{C_{H^2}}{C_{OPT}} \leq \frac{(1 + \alpha)p_1 + T_2}{\alpha p_1 + T_2} = 1 + \frac{p_1}{\alpha p_1 + T_2} \leq 1 + \frac{p_1}{\alpha p_1 + p_1} = (1 + \alpha).$$

If $T_2 < p_1$, then $C_{OPT} \geq r' + p_1 > \alpha p_1 + p_1 = (1 + \alpha)p_1$. So

$$\frac{C_{H^2}}{C_{OPT}} \leq \frac{(1 + \alpha)p_1 + T_2}{(1 + \alpha)p_1} = 1 + \frac{T_2}{(1 + \alpha)p_1} < 1 + \frac{p_1}{(1 + \alpha)p_1} = (1 + \alpha).$$

Then $\frac{C_{H^2}}{C_{OPT}} \leq 1 + \alpha$.

Lemma 4.6: If $r' \leq p_2$, then the worst-case ratio of algorithm H^2 is not greater than $1 + \alpha$.

Proof: In this case, because of $p_2 \leq p_1$, so $(1 + \alpha)p_2 \leq p_1 + p_2$, then $p_2 \leq \frac{p_1 + p_2}{1 + \alpha} = \alpha(p_1 + p_2)$, we have

$$C_{H^2} \leq p_2 + T \leq \alpha(p_1 + p_2) + T \leq (1 + \alpha)T \leq (1 + \alpha)C_{OPT}.$$

Then $\frac{C_{H^2}}{C_{OPT}} \leq 1 + \alpha$.

The following theorem can be drawn from the above these lemmas.

Theorem 1: For any job sets, the worst-case ratio of algorithm H^2 is not greater than $1 + \alpha$, where $\alpha = \frac{\sqrt{5}-1}{2}$.

5 Conclusion

In this paper we have considered the problem of scheduling family jobs on a parallel-batching machine with dynamic job arrivals. and the capacity of the machine is infinite. For the schedule problem, we provide an algorithm H^2 with the worst-case ratio $1 + \alpha$, and it is implied that the algorithm H^2 is a best possible algorithm for the unbounded case according to the Theorem 1 and Lemma 3.1.

Acknowledgment. The paper is supported by The National Natural Science Youth Foundation (No. 10901144) and the Henan Education Department Natural Science Foundation (No. 2011A110021, No. 2011B11031, No. 2010B520030).

References

1. Zhang, G., Cai, X., Wong, C.K.: On-line algorithms for minimizing makespan on batch processing machines. *Naval Research Logistics* 48, 241 (2001)
2. Uzsoy, R.: Scheduling batch processing machines with incompatible job families. *International Journal of Production Research* 33, 2685 (1995)
3. Poon, C.K., Yu, W.: On-line scheduling algorithms for a batch machine with finite capacity. *Journal of Combinatorial Optimization* 9, 167 (2005)
4. Deng, X., Poon, C.K., Zhang, Y.: Approximation Algorithms in Batch Processing. In: Aggarwal, A.K., Pandu Rangan, C. (eds.) *ISAAC 1999*. LNCS, vol. 1741, pp. 153–162. Springer, Heidelberg (1999)
5. Chen, B., Deng, X.T., Zang, W.A.: On-line scheduling a batch processing system to minimize total weighted job completion time. *Journal of Combinatorial Optimization* 8, 85 (2004)
6. Azizoglu, M., Webster, S.: Scheduling a batch processing machine with incompatible job families. *Computers & Industrial Engineering* 39, 325 (2001)
7. Nong, Q.Q., Yuan, J.J., Fu, R.Y., Lin, L., Tian, J.: The single-machine parallel-batching on-line scheduling problem with family jobs to minimize makespan. *International Journal of Production Economics* 111, 435 (2008)
8. Potts, C.N., Kovalyov, M.Y.: Scheduling with batching: A review. *European Journal of Operational Research* 120, 228 (2000)
9. Liu, Z., Yu, W.: Scheduling one batch processor subject to job release dates. *Discrete Applied Mathematics* 105, 129 (2000)

Research on Layout Evaluation Indexes System of Dangerous Goods Logistics Port Based on AHP

Zhang Peilin, Mao Jian, and Yang Long

School of Transportation
Wuhan University of Technology
Wuhan, China
zyj752@yeah.net

Abstract. The reasonable layout of dangerous goods logistics port will largely affect the economic benefits of port and the security of port operations. Achieve a comprehensive, integrated, objective evaluation on layout of the program; it is needed to establish a set of comprehensive evaluation index system. And then, the indexes weights' determination is one of the keys of establish of the evaluation index system. In this paper, we do research on dangerous goods logistics port layout indexes and their weight determination based on AHP. The goal is more objective and strives to accurately reflect the reasonable conditions of dangerous goods logistics port layout.

Keywords: Dangerous goods, port layout, logistics, evaluation index, AHP.

1 Introduction

Dangerous goods Logistics is different from the general logistics, dangerous goods have flammable, explosive, corrosive and other high-risk characteristics, so the factors affected layout of dangerous goods logistics port are more complex. Whether the layout of dangerous goods logistics port arrangement is reasonable, it will largely affect the economic benefits of port and the security of port operations. Therefore, it is necessary to establish a comprehensive evaluation index system, the reasonable of evaluation index system of which will largely affect the validity of evaluation results. So, a good grasp of the index weight in evaluation system of dangerous goods logistics port is a top priority work. The studies on layout program evaluation of dangerous goods logistics port at home and abroad are very less. This paper based on previous studies proposed layout program evaluation index system of the dangerous goods logistics based on AHP, and make study on the determination of reasonable target weights. This paper has an important status guiding significance on objective and accurate assessment of layout program of dangerous goods port.

1.1 Analytical Hierarchy Process Theory

Analytical Hierarchy Process (AHP) is on the practical issues involved in classification of factors, construct a hierarchical model that Association between various factors, and then in the pair wise comparisons them in the pair wise between levels to determine

their relative importance, and then single Weights determine the level and consistency check can be obtained on all levels of factors on the weight of a hierarchical analysis method [1].

1.2 Advantages of AHP

1) *Applicability.* Decision-makers can use the AHP for decision making, which greatly increases the practicability and scientific of the decisions.

2) *Simplicity.* AHP simplifies a complex problem into its components, and then group them with a hierarchical structure.

3) *Practicality.* With great practicality, AHP can not only undertake quantitative analysis also the qualitative analysis.

4) *Systematic.* AHP treats a complex problem as a system, which is an effective decision thinking mode.

2 Model Building

This paper uses a combination of qualitative and quantitative analytical hierarchy process (AHP) to evaluate the layout program of dangerous goods logistics port. First of all, inviting 10 experienced experts in the shipping and engineering to discuss and determine quantitative indexes of the layout program of dangerous goods logistics port; then built the hierarchical model (see Figure 1) on the basis of first step; finally get the arrive at various indicators of degree of order.

2.1 Building AHP Model

Schematic diagram of evaluation index system as follows:

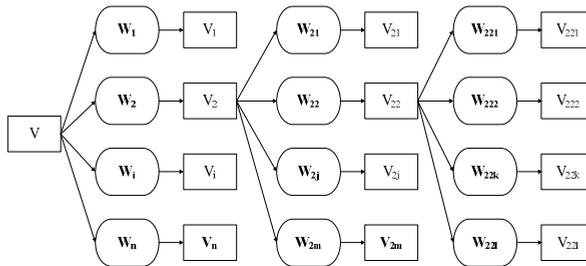


Fig. 1. Dangerous Goods Logistics Port Layout Evaluate System Schematic

W_i is Index share weights, V_x is Index evaluation scores, n is the number of A-level index, m is the number of B-level index, l is the number of C-level index

2.2 Determination Evaluation Indexes

When we evaluate dangerous goods logistics port layout program, at first, provide expert a detailed layout program of the port of dangerous goods logistics, using

5-point scale to rate. The bigger the better indicator of value, such as port economic benefits, from small to large score is set to 1-5; on the lower, the better indicators, such as ship berthing time, parking costs, from small to large score is set to 5-1 points. In addition, intermediate indicators, will be the best mid-range value is set to 5 minutes, lower or higher than the middle of the range of values followed by decreasing the fraction [4].

When experts get layout program, select the each layout scheme based on scoring, and according to their own experience to give the share weight of each index, form a set of dangerous goods logistics port evaluation index system.

2.3 Determination Indexes Weight

The index weights of dangerous goods logistics port reflect the importance of index in evaluation, the determination of weights will affect the science and effectiveness of the results of the program.

The investigation targeted on the comprehensive evaluation index of the layout program of dangerous goods logistics port, the mainly area of the survey is a number of experts in port research to evaluate the importance of indicators.

Experts participating in the survey investigated the importance of the evaluation indicators and scoring weights are given, is the base of further application of fuzzy comprehensive assessment.

Each grading and fuzzy sets can be seen as following:

Table 1. The fuzzy sets and grading of dangerous goods port evaluation program

| Fuzzy sets | Grading |
|--|---------|
| A: The essential elements of evaluation program ; very important | 10 |
| B: The required elements of evaluation program; more important | 7 |
| C: The general elements of evaluation program; general important | 4 |
| D: Little effect on evaluation program; unimportant | 1 |

The weight of A-level index is determined as following by experts:

Table 2. The weight of A-level index of dangerous goods port evaluation program

| Index | Weight |
|---------------------|--------|
| Production capacity | 0.4 |
| Economic benefits | 0.4 |
| Production security | 0.2 |

The weight of B-level index is determined as following by experts:

Table 3. The weight of B-level index(Production capacity) of dangerous goods port evaluation program

| Index | Weight |
|-------------------|--------|
| Berth capacity | 0.4 |
| Handling capacity | 0.4 |
| Yard capacity | 0.2 |

Table 4. The weight of B-level index(Economic benefits) of dangerous goods port evaluation program

| Index | Weight |
|------------------------------|--------|
| National economic benefits | 0.6 |
| Enterprise financial benefit | 0.4 |

Table 5. The weight of B-level index(Production security) of dangerous goods port evaluation program

| Index | Weight |
|-----------------------------|--------|
| Handing operations security | 0.6 |
| Safety of storage yard | 0.4 |

Because an excessive number of C-level index, we don't list its weight here.

2.4 Determinations Matrix Construction

Because the expert evaluation method has some subjectivity and inaccuracy, therefore it is necessary to further determine the index weights though the fuzzy analytic hierarchy process after determine the primary and secondary evaluation index weight through the Delphi method.

First, we have to determine the relative importance of different indicators, to compare any two different index, specific compared results can be converted to specific values as following:

Table 6. Completion of Evaluation Indexes of Ahp

| Importance of index X_i and X_j | The value of Determination b_{ij} |
|-------------------------------------|-------------------------------------|
| The same important | 1 |
| A little Important | 3 |
| More important | 5 |
| Highly important | 7 |
| Absolutely important | 9 |
| Importance between the above degree | 2,4,6,8 |

Based on the results of the above evaluation indexes, calculating the importance Matrix of the indexes:

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{1m} \\ b_{21} & b_{22} & b_{2m} \\ b_{m1} & b_{m2} & b_{mm} \end{bmatrix}$$

M is the number of evaluation indexes.

Calculating the Largest eigenvalue λ_{max} and Eigenvectors $\xi = (x_1, x_2, \dots, x_m)$ of

the matrix, make x_i as the weight of α_i , make $\frac{x_1}{\sum_{i=1}^m x_i}, \frac{x_2}{\sum_{i=1}^m x_i}, \dots, \frac{x_m}{\sum_{i=1}^m x_i}$ as the fuzzy

vector $A = (a_1, a_2, \dots, a_m)$ of the factors .

It is necessary, make Consistency test when calculating the largest eigenvalue λ_{max} to ensure the accuracy and reliability of the results. Consistency index $CI = \frac{\lambda_{max} - n}{n - 1}$, Test coefficient $CR = \frac{CI}{RI}$, and RI is the index to evaluate the Consistency.

3 Example

The following make the production capacity as an example, introducing the calculations.

The evaluation results of the production capacity by experts can be seen as following:

Table 7. The evaluation results of the production capacity

| Production capacity | Highest score | Lowest score | Average |
|---------------------|---------------|--------------|---------|
| Berth capacity | 10 | 7 | 8.11 |
| Handling capacity | 10 | 5 | 7.99 |
| Yard capacity | 10 | 4 | 7.63 |

Get the determination matrix:

$$A_0 = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 1 & 3 \\ 1/3 & 1/3 & 1 \end{bmatrix}$$



The consistency test shows that the determination matrix of this layer has fully consistent comparison. The impact weight to the higher index were 0.43, 0.43 and 0.14.

Similarly, the determination of other B-level and C-level indexes and the establishment of the evaluation model of dangerous goods port layout are show in Table 8.

Table 8. Dangerous Goods Logistics Port Layout Evaluation Form

| A-level index | Weight | B-level | Weight | C-level | Weight | | |
|---------------------|--------|------------------------------|--------|---------------------------------|--------|--|--------|
| Production capacity | 0.4 | Berth capacity | 0.166 | Berths number | 0.074 | | |
| | | | | Berths average utilization rate | 0.0268 | | |
| | | | | Port annual throughput | 0.0652 | | |
| | | Handing capacity | 0.166 | | | Crane number | 0.0476 |
| | | | | | | Crane average efficiency | 0.0412 |
| | | | | | | Quayside average utilization rate | 0.048 |
| | | | | | | Truck number | 0.0206 |
| | | | | | | Terminal front channels | 0.0086 |
| | | Yard capacity | 0.068 | | | Yard capacity | 0.0289 |
| | | | | | | Average stockpiling of goods | 0.0099 |
| | | | | | | Yard average utilization rate | 0.0065 |
| | | | | | | Yard average utilization rate of machinery | 0.0226 |
| Economic benefits | 0.4 | National economic benefits | 0.232 | ENPV | 0.1136 | | |
| | | | | IRR | 0.1184 | | |
| | | Enterprise financial benefit | 0.168 | | | FIRR | 0.0135 |
| | | | | | | FNPV | 0.0218 |
| | | | | | | Pre-tax investment recovery period | 0.0185 |
| | | | | | | Payback period after tax | 0.0235 |
| | | | | | | ROI | 0.0252 |
| | | | | | | Investment tax rate | 0.0218 |
| | | | | | | Capital profit | 0.0202 |
| | | | | | | Breakeven point | 0.0235 |



Table 8. (Continued)

| | | | | | | |
|---------------------|------------------|------------------------------|------|-----------------------------|--|--------|
| Production security | 0.2 | Handling operations security | 0.12 | Handling average efficiency | 0.0364 | |
| | | | | Handling shortest distance | 0.0288 | |
| | | | | Handling most long-distance | 0.0288 | |
| | | | | Emergency handling ability | 0.026 | |
| | Yard work safety | 0.08 | | | The shortest distance to the yard | 0.0144 |
| | | | | | The longest distance to the yard | 0.0144 |
| | | | | | The distance between the reservoir point | 0.0258 |
| | | | | | Emergency yard capacity | 0.0254 |

4 Conclusions

This method was established through the use of AHP dangerous goods logistics port layout comprehensive evaluation index system, which have changed the past method’s shortcomings. From the evaluation index system, we can clearly understand the impact of the port layout of dangerous goods logistics factors and secondary factors, so as to carry out qualitative and quantitative analysis to the layout of the program. Meanwhile, expert advice method and the fuzzy AHP method are combined to determine the index weights, which can avoid defects caused when the above two ways are used alone. Besides, in this way, the index weights are more in line with the actual situation and be good for comprehensive evaluation; most importantly, it has strong practical value to the evaluation of dangerous goods Logistics port layout program.

References

1. Liu, T.: Optimization of Multi-transportation Organization Mode for Complicated Freight Flow. Wuhan University of Technology (2010)
2. Wang, S.: Research on Container Terminal Layout Simulation Based on eM-Plant. Port Handling (6), 22–24 (2006)
3. Deng, S.: Complex engineering system design Assessment Method. Systems Engineering and Electronics (12), 1488–1491 (2003)
4. Yan, W.: Research on Determination of Indexes and Their Weights in Comprehensive Evaluation for Layout of Container Terminal. Port Handling (3), 42–43 (2007)
5. Gao, Z.: Research on Academic Standards and Assessment Based on Fuzzy Evaluation and AHP. Chongqing School Paper of Transportation (3), 96–100 (2006)



An Improved BP Neural Network Algorithm Embedded with Logistic Mapping and Its Application

Wang Tao

School of Software, East China Jiao Tong University, NanChang China 330013

Abstract. The algorithm uses double searching to avoid the long time of Logistic chaotic searching, namely BP-AA algorithm confirms error function descending monotonously, then using Logistic chaos algorithm to search in wide area, insuring to jump out local minimum point; at last careful search is in progress in ambient area of approximate optimum solution to get whole optimum solution. This algorithm raises searching efficiency effectively, independent variables' definition area of careful search reduce a little and the probability of seeking extremum point also is raised, confirming the convergence of BP-AAEC algorithm.

Keywords: BP, Chaos, Logistic Mapping.

1 Introduction

Along with the appearance of the great united power system and development of the power market, the safe stability of power system becomes more and more complex, and the requirement of stable operation becomes more and more strict. Both of these make the design, operation and control need simulation calculation exactly. Because the power load displays the complex dynamics behavior, the power load has some characteristic, such as time variation, multiplicity, no continuity and so on. The annual neural network is widely used in power load forecasting. Hereinto, the forward neural network gives priority to the BP algorithm. But the existed problem of the BP algorithm has not been resolved. How to overcome the shortcoming that convergence speed is slow and proneness to trapping in local minimum? This problem is hot point at present. This paper researches the method to resolve the problem. By way of overcoming the shortcoming that convergence speed is slow and network is prone to trapping in local minimum, the important factor-activation function which influences convergence in BP algorithm, and training algorithm is put forward which is based on three adjustable parameters activation function. Then, combined with randomness and ergodic property of chaos, the hybrid training algorithm embedded dual searching of chaos mapping is brought forward.

2 Amelioration of Bp Algorithm

BP algorithm has been applied widely just because of easiness of its realization. The essence of BP algorithm is the descending grads method. Although the descending direction of every time is the one in which the function value reduces the most quickly,

the nonlinear function is very complex, and partial grads in the error curved face do not point to minimal point. So, convergence route is zigzag, and the minimal point could be obtained after many steps. Objective function would need longer time to achieve schedule requirement. In application of using BP exist two important problems:

(1) Convergence speed is slow; usually convergence needs more than one thousand times;

(2) Objective function is prone to getting into local minimum.

At present, usually used methods to enhance convergence speed is for example: affiliating item of momentum ,using higher derivative or conjugate grads method or recurrence least-squares procedure or nerve cell spacing searching and so on; for local minimum, some scholars put forward several global optimization methods, such as: simulation anneal methods, genetic algorithm. But enhancing convergence speed and global optimization could not be satisfied contemporary.

Analysis of Activation Function. After a great deal of research on ANN, recognition to nerve cell activation function is as follows:

(1) In the traditional neural network model, the connection of nerve cell is deemed to memory information (adjustable), and activation functions which can simulate function of cytosine are fixed. So it makes network training progress could be simplified to modify connection weight value of every layer nerve cell. Simplification of traditional algorithm constricts nonlinear mapping ability, so it makes network training precision poor, and the convergence speed is also impacted.

(2) Activation functions exist in saturate area. When it output of nerve cell drop into the saturate area, the weight value need be adjusted greatly to make the processing unit out of the saturate area.

Based on analysis above, some adjustable parameter should be introduced. When output of nerve cell drop into the saturate area of activation function, that derived number of activation function could adjust following the change of the error avoids the appearance of anesthesia effectively and expedites convergence speed. Therefore, the hide layer nerve cell's activation function of Sigmoid function which has three online training variation parameter is put forward. Through improving activation function, connection weight value and activation function are adjusted in the training progress contemporarily, and the stronger nonlinear ability could be achieved, convergence speed and approaching precision are also expedited.

BP-AA (Adjustable activation function algorithm) three adjustable parameter Sigmoid activation function is as follows:

$$F_{a,b,\lambda}(x) = \frac{1}{1 + e^{-\frac{x-b}{\lambda}}} + a \quad (1)$$

Where, parameter a is excursion one, b is threshold, λ is gradient gene. It is obvious that function $F_{a,b,\lambda}(x)$ has more prolific nonlinear expression ability than $F(x)$.

Parameter a and b decide the uprightness and level position of function, and λ decides the form of function. When λ change from 0.1 to 10, the form of function is fig.1(a , b are equal to 0 in the figure). It could be gotten from the figure that in the saturate area, derived function of activation function has bigger change following the change of gradient gene λ .

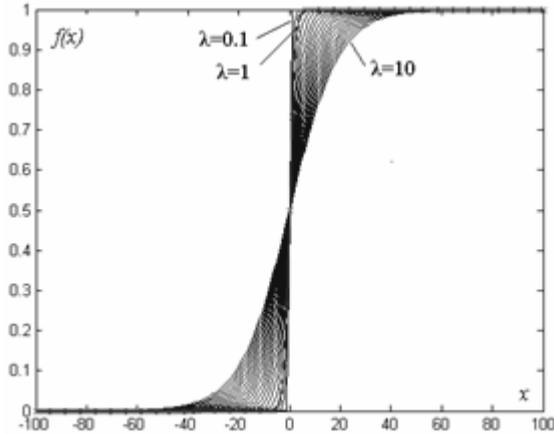


Fig. 1. Curve of activation function while λ changing from 0.1 to 10

The deduction of training algorithm presented in this paper, which adjusts network weight value, ω and activation function $F_{a,b,\lambda}(x)$ contemporarily will not give in details.

Simulation experiment and Capability analysis. For the sake of researching the improvement of improved algorithm, the equation (2) is used to test the improved algorithm and ANN using BP, namely:

$$f(x) = 10 \sin(2\pi x) + 20, \quad x = [0, 2] \quad (2)$$

the network adopt the structure of 1-8-1 single hide layer, and the maximum training times is 50000, and the precision is set to 0.01, 0.001, 0.0001 respectively. The ability of convergence of two algorithms is compared in 200 times in every one of the precisions.

For the sake of research the approaching ability of improved algorithm, the structure of 1-8-1 is used in the network using normal BP, and the structure of 1-6-1 is used in the network using BP-AA algorithm. The error of mean square of network output of the two algorithms is compared in equal times of iteration.

3 Ann Training Algorithm Embedded with Logistic Mapping

Theorem of BP-AAEC Algorithm. Although BP-Adjustable activation function algorithm raised convergence speed of neural network, the essence of algorithm is to seeking the most optimal value following the direction of digressive grads. When initial value is confirmed, the decent course is also confirmed. So the problem of trending local minimum also exists, and it has great contact with confirming initial value. We can say that BP-AA algorithm has the good ability of searching in partial area and ability of searching in global area. So, we can combine with two algorithms

to making the best of LOA (Logistic optimal algorithm)'s fully searching ability and BP-AA algorithm's partial searching ability. BP-AA algorithm is used to resolve the weight value and parameters in neutral network. When getting in partial least, using LOA algorithm can choose new initial value of every parameter to jump out partial least.

Therefore, we raise the Adjustable activation function and grad optimism training algorithm embedded Logistic chaotic mapping in BP network's training. This algorithm is not only efficient but also difficult to get in local minimum. The algorithm can converge to fully optimum with probability of 1. it is called BP-AAEC (Adjustable Activation function and Embedding Chaos algorithm) Algorithm.

Description of Algorithm is as follows:

Step1. Suppose that the excursion parameter of neutral network is a and threshold value is b, and gradient gene is λ , weight value w can be exchanged by X_k . Setting the initial value of parameters- X_0 ; setting error function threshold value e, the threshold value of repeating times m and m1, ordering $k=0, k_1=0, k_2=0$; letting X^* is the most optimum network parameter at present;

Step2. X_k can be used by BP-AA algorithm searching to get parameter X_k^* ;

Step3. Compare $E(X^*)$ with $E(X_k^*)$, if $E(X^*) - E(X_k^*) > e$, ordering $X^* = X_k^*$, $k_1=0, k=k+1$, then transferring to step2; if $E(X^*) \geq E(X_k^*)$ and $|E(X^*) - E(X_k^*)|$, ordering $X^* = X_k^*$

if $k_1 \leq m, k_1=k_1+1, k=k+1$, then transfer to step2 ;

if $k_1 > m$, ordering $k_1=0$, then transfer to step4 ;

if $E(X^*) < E(X_k^*)$, X_k^* holding the line, $k_2=k_2+1$;

if $k_2 \leq m_1$, then transfer to step4 ;

if $k_2 > m_1$, end ;

Step4 X_k^* is mapped to $[0, 1]$, and gained Y_k , namely $Y_k = \frac{X_k^* - a}{b - a}$, a, b is plus

coefficient, and their value can be confirmed by every material parameter. The function is that parameters needing optimum are mapped the bound of chaotic parameters getting value;

Step5 Y_k is inputted to Logistic chaotic equation, and gets chaos parameter Y_k^* ;

Step6 the new parameter is produced by chaos variable Y_k^* , namely

$$X_{k+1} = X_k^* + |X_k^*| \cdot (2Y_k^* - 1) \cdot r, \text{ transferring to step2 ;}$$

So you say, if the decrement of value of error function using BP-AA m times are all smaller than e, then transferring to step4 and using chaos mapping to help BP-AA jumping out partial least. In step3, if after m times of chaos searching, error function does not reduce, we can judge that the globally optimal solution has been found. When



the mathematic is ending, the globally optimal solution has been found according to ameliorating BP algorithm. In step6 new parameters can search with radius of $|X_k| r$, searching region can be adjusted using r.

4 Capability Testing of BP-AAEC Training Algorithm

For testing the capability of jumping out local minimum after the training algorithm embedded with chaotic searching, the function that has many partial minimal points is designed, just as equation (3),and compare with normal BP algorithm.

$$f(x) = 3 + \sin(0.5x) \log((x + 0.001) / 40), \quad x = [0,100] \tag{3}$$

This function has the maximum and minimum, when $x=9$ $f(x)$ has the maximum 4.458, when $x=1$, $f(x)$ has the minimum 0.41655, and this function has 7 maximal values and 8 minimal values.

A hide layer ANN is used, a hide layer including 25 nerve cell, separately using BP normal Algorithm and improved BP Algorithm embedded chaotic searching to training with 100 input and output which are presented. training error precision is set to 0.005, the whole iteration times is 50000. The network output simulation curves are given by Fig.2 and Fig.3. In the figures, "-" means the output of neural network, and the "+" means the sample data.

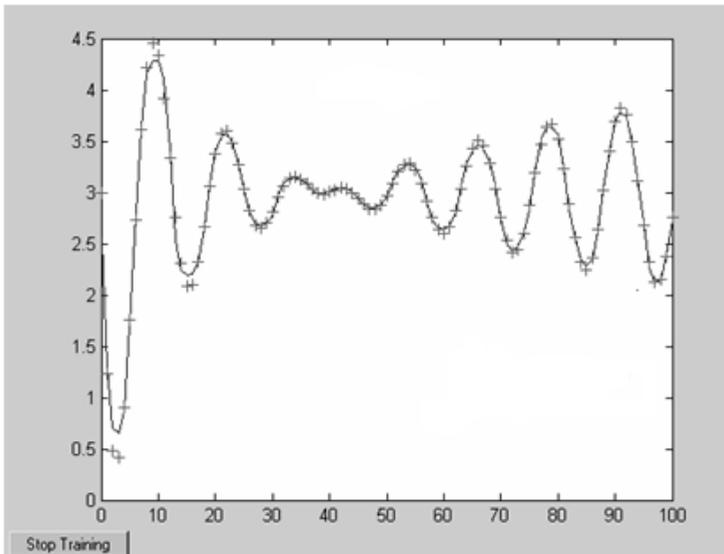


Fig. 2. Simulation curve of ANN using BP algorithm

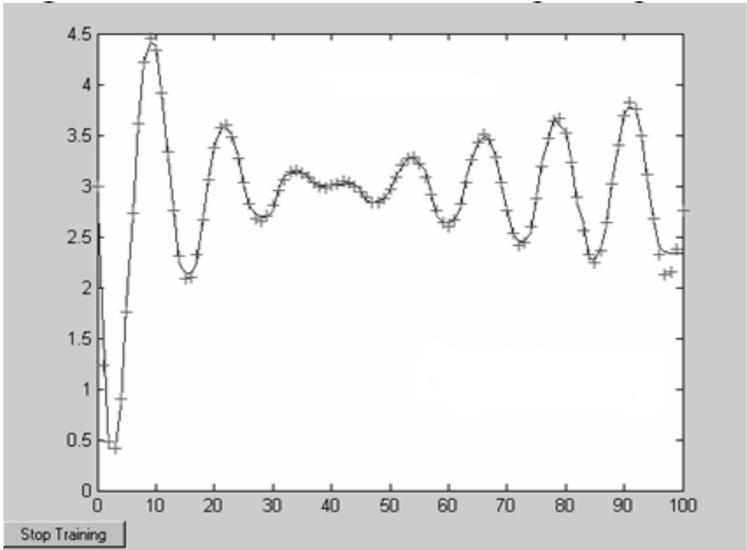


Fig. 3. Simulation curve of ANN using BP-AAEC algorithm

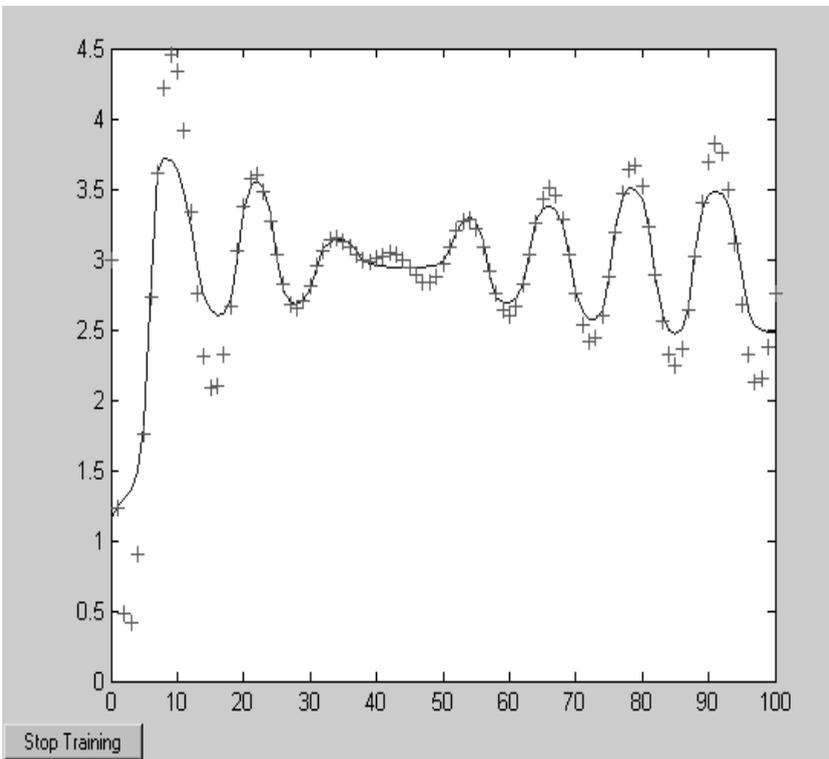


Fig. 4. Simulation curve of ANN using BP algorithm at loop 228

Fig.2 is the simulation curve of ANN using BP algorithm, the iteration time is 2080, error output is 0.00499616, and optimum point is not found. Fig.3 is the simulation output of ANN using BP-AAEC algorithm. The iteration times are 8 and error output is 0.00499796. At the same time, the least point has been found. Fig.4. is output of ANN using BP algorithm after iterating 228 times. After the comparison, we can make the conclusion that using improved algorithm embedded chaos searching to train network can get rid of disturbance of local minimum point and get the optimum point.

And at the same iteration times, ANN using BP algorithm comes under great disturbance of partial least. The error of output is 0.0876432. All of these illustrate that the ability of getting rid of local minimum of improved algorithm combining with chaos researching is stronger than ANN using BP Algorithm.

4 Conclusion

To improve precision of algorithm of power load forecasting is the impending problem, which should be resolved immediately. This paper researches the essential characteristics of ANN using BP which is widely used in power load forecasting and theorem of ANN using BP-AAEC algorithm. The theory analysis has indicated that the hybrid optimum training algorithm embedded chaos sequence would play a more and more important role in application of power load forecasting.

References

1. Park, D.C., El-Sharkawi, M.A., Marks II, R.J.: Electric Load Forecasting Using an Artificial Neural Network. *IEEE Trans. on Power Systems* 6(2), 442–449 (1991)
2. Hyde, O., Hodnett, R.E.: An adaptable automated procedure for short-term electricity load forecasting. *IEEE Trans. On Power Systems* 12(1), 84–93 (1997)
3. Hippert, H.S., Pedreira, C.E.: Estimating temperature profiles for short-term load forecasting: neural networks compared to linear models. In: *IEE Proceedings Generation, Transmission and Distribution*, July 11, vol. 151(4), pp. 543–547 (2009)
4. Baczynski, D., Parol, M.: Influence of artificial neural network structure on quality of short-term electric energy consumption forecast. In: *IEE Proceedings Generation, Transmission and Distribution*, March 2, vol. 151(2), pp. 241–245 (2004)
5. Michanos, S.P., Tsakoumis, A.C., Fessas, P., Vladov, S.S., Mladenov, V.M.: Short-term load forecasting using a chaotic time series. In: *International Symposium on Signals, Circuits and Systems, SCS 2009*, July 10–11, vol. 2, pp. 437–440 (2003); Maxwell, J.C.: *A Treatise on Electricity and Magnetism*, 3rd edn., vol. 2, pp. 68–73. Clarendon, Oxford (1892)

The Analytics of College Sports Participation and Implementation of the Path of Sun Sports

Zhang Run-tao

Art and Sport Institution of Si Chuan Agricultural University, Si Chuan, Ya An, 625014, China

Abstract. Carrying out the sun sports is an important expression of fully implement in the Party's education policy. Promoting the launching of the sun sports actively in universities is good to the health, so that students can input into studies better. On the current level of involvement of university students carry out a survey analysis, proposed effective implementation path of the sun sports countermeasures.

Keywords: Universities; Sports; Survey; Participation.

1 Introduction

Health is a college student contributing to our country; serve the society's basic premise. But, in recent years, China's college students' physical quality and health appeared problem that does not allow ignoring. In the campus extensively sunshine sports, strengthening students' sports consciousness, promote the students' participation in sports, make the student's physical quality and get fundamental change is the universities of the moment.

2 Analysis on College Students to Participate in the Sports

College Students' spare Time Arrangement. In universities, the learning tension in the secondary level is relatively easier. But, the classmates are doing what? The author survey found: college students in Yaan from table 1 shows, participate in sports, reading of the men and women living are respectively: 11.00 % 12.70; 15.15 %, 22.30%; To participate in athletic spare time girls slightly higher than the proportion of boys reason has : it is now many female college students out of the body, cover consideration, often use after school engaged in running, aerobics, fitness, badminton sport; Second, in recent years the girl one's deceased father grind proportion more and more high, many girls in order to keep the good study condition, so insist to exercise; Three is existing school sports appraisal criteria, including the national student physical health standard "; Some girls are bard, so some girls insist to take exercise. From 2002's "student physical health standard", to the 2007 national student physical health standard, I take the university student's male, 1500 meters, and female 800m: originally for three points with 55 seconds fail, now boy for four points 33 seconds, the girl is 4 minutes 23 seconds. In reality there are quite a number of boys test fail. In surveys found that a third of men and women who have a crush on network, this is a society,

also the common problems exist widely in universities. In addition, there is quite a part of the students thought empty, don't understand why you go to university, for whom in college? Especially in the not "211" and "985" university, the employment problem is more serious. The majority of students feel confused, don't know where your future? Including part family impoverished undergraduate, busy work-study, don't say not to participate in exercise, even studies also fallen into a state of neglect. These conditions the institutions engaged in higher education needs, leadership, workers attention, it is a national education system, an important symbol of national prospect.

Table 1. Things college student most want to do in the spare time want

| Content | reading | online | movement | sleep | shopping | other |
|-----------|---------|--------|----------|-------|----------|-------|
| male(%) | 11.00 | 44.72 | 12.70 | 17.70 | 7.60 | 6.28 |
| female(%) | 22.30 | 34.75 | 15.15 | 16.60 | 8.59 | 2.61 |

College Student’s Main Sports Activities. Form Now of college campus, students attend also exercise should form is varied. The author in the survey found: a physical exercise activities of the purpose of a few college students; Such as joining campus association or the number of accounts for only boys club 4.23 %; Girls calamitous close 251 % accounted for only. At this point, it more realistic today's college students, seldom someone to pursuit of an ideal, the surreal life things. Many students gym class afraid of hot is afraid of the cold, afraid of tired and so on. See, now from exercise alone activities to the students more. The boy has 29.53 %, girls have 48.07 % people like single movement; at the inquest, many girls found in the evening, the boy like exercise, ask them why? And they said one morning get up; it is evening exercise no one else can see. Others see will be embarrassed. It may have deeper issues to discuss, but the phenomenon is shown, or not loner collective concept. In addition, girl is on the prominence; in the same quarters with friends or prominence in boys. Table 2 shown here with the phenomenon of the younger generation students may personality, family environment, and social context factors.

Table 2. Main Method of Sport Activity

| Content | Separate activity | join in the fun | with classmates | Joining campus association or club |
|-----------|-------------------|-----------------|-----------------|------------------------------------|
| male (%) | 29.53 | 16.29 | 49.95 | 4.23 |
| female(%) | 48.07 | 38.13 | 11.29 | 2.51 |

College Student’s Sports Exercise Place. In polls can see, college students to participate in sports exercise good sites focused on track, gym and gymnasium. Track now speaking, or in the school student's main place to take exercise in all colleges, which is basically the same. The gym and gymnasiums difference, using different colleges is mainly development time, charge height, etc. In addition to a number of university students is in social venues to exercise or training. These students learn new technology is mainly by girls, fitness for purpose; Boys on fitness, for the purpose of social communication. The interview found that some college students to do well.



Table 3. Sports Exercise Place

| Content | Stadium | Gymnasium | Track | Glade | Other |
|-----------|---------|-----------|-------|-------|-------|
| male (%) | 22.04 | 21.28 | 54.61 | 0.22 | 1.85 |
| female(%) | 6.66 | 10.91 | 67.37 | 2.61 | 12.45 |

College Student's Sports Project. Physique, interests, hobbies, etc is different, its involvement with sports fitness program also different. The boy like to load, is adversarial of sport, with three ball (basketball, volleyball, football) primarily, 42.35% of small ball (badminton, tennis, table tennis) accounted for 19.54 %; And while girls' favorite intensity is small and conducive to develop its own form of sport, with calisthenics jogging and form three ball, mainly the accounted for 11.78 % 20.56 %, 10.71%, 10.23 %.

College Student's Sports Frequency and Time. The author in the survey found; College students to participate in the same exercise with their like frequency actually depend on. This shows a problem, only like or interest, just have the participation. Judging from participating in frequency, about one-third multi-point male college students in every week three times or more girls are in fifth appearance. Week insisted once or twice in the male classmate third levels; and while girls' account for half of the concept. From this perspective: male and female college students have nearly two-thirds of the students can weekly insist to take exercise one to three times. This means that education is a result of the constitution. Now most college students or recognized the importance of physical exercise. As for occasionally attend, or in addition to the outside never joined in gym class, they thought and understanding is more difficult to understand, exist various influence factors such as: social factors, education factors, family education factors, etc. "About develop in the national billionaire students notice of the sun sports should cooperate in physical education class teaching, points out that the average every weekday, ensure that students is an hour exercise time [1]. Boys than girls prefer sports fitness activities, take exercise every week 3 times or more, except 34.42 % of sports extracurricular never take part in physical exercises phenomenon is also very serious, the boy to account for % 15.96; But the enthusiasm of girl students in physical exercise is not very high, 12.74 % occasionally attending, 19.98 % take exercise every week 3 times above, have 50.10 % of female undergraduates weekly can insist on 1-2 times of physical exercise. And 17.18 %except sports extracurricular never take part in physical training. Students participate in sports a week of time problems, the writer survey found: college students to participate in sports exercise time men and women each are not identical, the boy appear normal distribution among, both sides is low, the trend, which also high tallied the realistic situation. Each exercise can in 1-2 hours; the number of few really No fixed exercise time of college student's number is not too much. But can adhere to take exercise every week of undergraduate men or after two-thirds. But the girl is different, is a from low to high or from high to low status. To insist on more than an hour a week on the number of %; only Can persist in half an hour to an hour there 13.42 %; Nearly one-third of the girls could insist on half an hour of exercise; more than half of the girl's

physical exercise is not fixed form. This show on each physical exercise of female college students the confidence, ability and physical problems, problems could is multifaceted. In the practical teaching process, many authors also found that female university students do exist the physical strength is insufficient the status quo. The newborn matriculation military training each year, is an example, the military training period, there are many stand training of female fainting. It also brought out to the female university students' physical exercise should be strengthened. Whether from the female physiological development characteristics based on education, or the consideration sports education weak. In order to reach the sun sports puts forward "ensure that students each weekday average is an hour exercise time" requirements.

3 College Student's Sports Affect Factor

Affect schools and college students to participate in the sun sports fitness factor is various; the main factor is the school, lack of guidance of lacking space equipment and not interested. Think of these three factors 30.29 %accounted for boys 16.72 %, 10.64 %; the girl thinks 19.02 % of 29.44 %, 24.32 %. The boys often attend sports mainly ball games projects that need large enough to carry the venues and enough equipment; While girls prefer aerobics, shape, etc projects that professional is stronger, need to have special teachers for guidance. 11.51% 16.60% boys and girls affected that intense physical fitness activities, which requires schools to reduce students' academic burden, make students have more time to participate in sports. In addition, there are 24.32 % girl is not interested in sports, future schools attach importance to cultivating the female university students about sports interest and strengthen their part in sports consciousness. As for other aspect factor effect is not the main reason, for example, learning tension, did not know, other considerations.

4 Suggestions for the Sun Sports Development

The students take part in physical exercises enthusiasm is not very high. How to encourage the students to walk out of the classroom, to a playground, going into the sun take exercise, is the school sports workers face important issues.

PE Add with Local Characteristics and Emerging Sports Teaching Content. How to develop good sunshine in college sports here? To a great extent, considering the course content Settings and the material hot choice. Our country region is vast, is also a multi-ethnic country. University student's face big, how can use source of the different geographies, nations of college students can lead to accord with the students' psychological needs and to meet the exercise? This requires in curriculum fluctuation is according to the college education as the overall requirements and PE curriculum, and conform to their rule of local university environment where [2]. In PE curriculum setting, follow the college students' physical and mental development regularity and hobbies, reflect regional nationality characteristics; Insist on scientific and acceptability, not only should consider combining students during school study needs, to consider the needs of students to society after; Not only should choose to have the

characteristics of local good traditional sports, also select in recent years popular prevalent, college students or popular sport. These projects are sports dance, aerobics, futsal, kickboxing and many indoor fitness project and so on. Therefore, the school should be in these respects put more some equipment, equipment to ensure the students' extracurricular exercise needs.

Conducting various forms of extracurricular sports activities and meet the needs of the development of students' individuality. Extracurricular sports activities of college physical education is an important part in extracurricular PE activity, how to put this position, can say college to occupy the success or failure of the college physical education. In the universities sports teaching time is limited, the current teaching concept is weekly, a two hours [3]. If not, well the extracurricular sports activities in class, also more seriously hours a week for two hours, exercise are far can't satisfy the requirements of exercising one hour a day. In addition, in the universities sports work, a high level of team training, Collective futsal counseling, amateur team training, other sports counseling and training needs, therefore, the extracurricular sports activities in the arrangement and organization is very important. In addition, there are many college students on sports like is very individualize, teaching high school missing that he likes, he likes the author and no open, so requires full use of extracurricular sports activities in time, "PE teaching is insufficient, those college students can very good according to own hobby and characteristics, and participate in sports fitness activities. Do not only satisfy the need of students' individual character development, promote the students' physical health.

Sponsor Sports Cultural Art Festival, Build the Campus Physical Exercise Atmosphere. Sports culture festival is a new type of campus sports organization form, its rich content, various in form, the wider, more in line with the nationality quality education requirements, be beneficial to build strong physical exercise atmosphere. Various colleges and universities can according to their own different geographies, nations customs characteristics, holding different types, forms of campus sports culture festival. In the sports culture festival, the students like project, a crucial decision take out the project, as many students can go on to these events. Also the maximum mobilization in, such staff also common participation and lets the teachers and students through games, and in so doing already active competitive game, and enhance the teachers and students atmosphere of friendship. Is kill two birds with one stone.

Develop Sports Leisure Forms, Enrich the Students' Holiday Sports. Sports leisure is a introduced into China from the us and Europe near a new sport and entertainment form, its essence is some of the sports activity content to join people's leisure activities [4]. It can be on different structure group, family, students, and colleagues, couples, even small groups of an AD hoc. College is a collection of students, teachers and student's compatriot, class, a variety of forms such as large groups. Sports leisure development has a good environment. We can make full use of these favorable conditions for sports leisure activities, to enrich student's holiday sports life. Tourism, cultural landscape more areas, use holiday to departments, class, with compatriot, dormitory and other forms of the climb, outing activities to attract students.

Often attend sports leisure activities can be in a good sport environment of exercise do easily extensive and interpersonal, achieve physical fitness, loosen the mood, the improvement of health, the adjusting mood role.

5 Conclusion

All-rounded development has been the party's education policy development of higher education is to reflect, cannot be suppressed sports teaching and school sports work. The sun sports development of the movement to fully implement the party's education policy is one of the important reflected, in colleges and universities in the positive lift push, the launch of the sun sports to promote college students' physical health, so that students can better into the study is utmostly important role. Education in college sports, also caught reality, focus on the future, not only to make college students have good health study, but also in education thought fluctuation kongfu lifelong sports, make them after getting out of school with abundant energy can also contribute to the party's each work, their life energy; Also for his I formed the good life and exercise habits, health and happiness, the life can lay a solid foundation.

References

1. Wang, Y.H.: Carry out the billionaire student sunshine sports. Understanding and Thinking of Shanghai Sports College 31(6), 81–83 (2007)
2. Liu, N.Y., Wang, Z.T., Wang, H.T., Kei: Science and carry out the sun sports teaching and management (2007)
3. Liu, D.K.: From student sunshine to talk about sports positioning information. Teaching and Management (11), 114–115 (2007)
4. Li, Z.M.: some problems of college sports reform thinking. Sports Science Research (3), 36–39 (2010)

Using the Least Repeaters to Connect Radios

Danying Chen¹, Dongping Yu¹, Zheng Wang², and Yuehua Wan^{3,*}

¹ Jianxing College, Zhejiang University of Technology, Hangzhou 310023, China

² College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China

³ The MOE Key Laboratory of Special Purpose Equipment and Advanced Processing Technology, Zhejiang University of Technology, Hangzhou 310023, China
chenpaths@gmail.com

Abstract. The proper locations of repeaters are the precondition to the efficiency of very high frequency (VHF) radio signal transmission. Thinking about using the minimum number of repeaters to connect all radios in one area, we divided the whole process into three procedures: User-Repeater Connection, Repeater-Repeater Connection and Frequency Distribution. Each radio should have at least one repeater settled in its signal emission range that can recognize its frequency. Likewise, each repeater requires another repeater within its signal transmission range. When we have to reuse frequencies, the co-channel interference cannot be neglected. We developed a complete set of arithmetic to work it out, using the continuous tone-coded squelch system (CTCSS) in a novel way in order to increase the number of radios that one repeater can serve simultaneously. We also quantify the number for the specific cases studied, and discuss further areas for research.

Keywords: Repeaters Distribution, VHF Radios, CTCSS Technology.

1 Introduction

Very high frequency (VHF) is the frequency ranges from 30MHz to 300MHz which has been widely used for television and FM radio broadcast. This spectrum involves line-of-sight transmission and reception [1,2]. However, with the help of repeaters which pick up weak signals, amplify and retransmit them, the distances that signals emitted by VHF radios can reach will increase largely [3]. Technologically, each radio has a unique transmission frequency, and each repeater can only recognize signals with a specific frequency emitted from radios [4]. Since repeaters can interfere with one another unless they are far enough apart or transmit on different frequencies, the continuous tone-coded squelch system (CTCSS) has been invented to attach a sub audible tone to the frequency [5]. It ensures that two signals with the same frequency but different sub audible tones will never be regarded as the same signal, thus the interfering problem will not occur.

The locations of repeaters may have a great influence on radio signal transmission, affecting its efficiency and expenditure. However, while Isberg, R.A studied how an underground communication system in London used radio repeater amplifiers [6], Jog,

S.A.D; Bhagwat,S.S introduced a system making use of VHF medium covered fairly long distance by hill-top repeaters in [7], Dixon, Chandler and Mohamed considered various network topologies of repeaters which provide reliable and efficient links [8], they all mainly focus on the technological aspects of repeater's insertion on special geographical environments, and the number of radios a repeater should serve is always small. Seldom have they discussed the problem in a general mathematical perspective, nor did they adopt the CTCSS technology.

However, in this paper, we're going to discuss such a problem: For a given flat area with some settled radios, how to locate repeaters to ensure that any two radios can communicate at any time. Since the setting of repeaters can be costly, we try to develop a whole set of arithmetic to work out the minimum number of repeaters it needs to serve n simultaneous users in one region. We'll also use the CTCSS technology, but we will use it in a different way. As far as we are concerned, if each repeater can recognize one frequency with m different sub audible tones simultaneously, the signals it can handle at one time will increase, thus the number of repeaters we need will decrease accordingly.

Each repeater will be distributed with a frequency, if some frequencies are reused, then considering the signal-to-interference ratio [9], two repeaters using the same frequency must be far enough apart.

It should be mentioned that we only discuss about on-frequency repeaters which receive a radio frequency signal and re-transmit it on the same frequency. So once a radio emits a signal, for any other radios who want to receive this signal, all they have to do is to adjust their reception frequency and sub audible tone to the same as that radio.

2 Process Division

In our work, we consider that each repeater can recognize a specific frequency with m different sub audible tones or without sub audible tones, so it can handle $(m+1)$ radios at one time at most. Here we assume the signal transmission between repeaters uses special technique so that repeaters don't have to match their frequencies just like the transmission between repeater and radio does.

When we locate repeaters, there are four constraints. First, for each radio, there must be at least one repeater within its signal emission range that can recognize its frequency so its signal will not die away. Second, the maximum number of radios a repeater can serve at one time is $(m+1)$. Third, for each repeater, there must be at least one repeater within its signal transmission range to retransmit its signals. Fourth, two repeaters handling the same frequency should be far enough apart.

To simplify this problem, we divide the whole process into 3 successive procedures: User-Repeater Connection, Repeater-Repeater Connection, and Frequency Distribution.

User-Repeater Connection: In this section, we need to make sure that each radio has at least one repeater within its signal emission range that can receive its signal. To satisfy this request, we consider it in the following steps.

The First Step: Determine the Candidates. Here we introduce a concept--*candidates*, which are the available locations where repeaters can be set on originally. Generally, we can use equilateral polygon to cover a plane area perfectly [10]. There are three general types of shape: regular triangle, square and regular hexagon (see Fig.1).

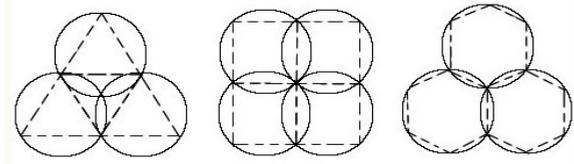


Fig. 1. Three general types of shape to cover a plane area

We adopt squares to cover the area. We make square grids to decide candidates. As is shown in Fig. 2, candidates are those points: centers of circles and intersections of circles which are denoted by red dots. For each radio, if the distance between it and some red dot is smaller than S (S is the signal emission range of radios), then the first step has been down.

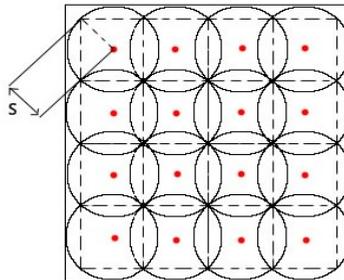


Fig. 2. The cover pattern and the candidates

The Second Step: Locate Original Repeaters. We will choose some candidates to locate original repeaters.

N : The number of chosen candidates

U : The number of radios

Z_i : If a repeater is set on the i^{th} candidate, the exact area its signal can reach is denoted by Z_i .

Q_i : The number of radios in Z_i that exceeds a repeater's serving capacity

M_{ij} : 0,1 variable. If Z_i can cover radio j , M_{ij} equals 1, otherwise M_{ij} equals 0.

Then a 0-1 matrix with N rows and U columns can be constructed:

$$\begin{pmatrix} M_{11} & \dots & M_{1U} \\ \vdots & \ddots & \vdots \\ M_{N1} & \dots & M_{NU} \end{pmatrix}$$

We should use least rows to ensure that in the 0-1 matrix made up of the selected rows, there exists at least one 1 in each column. Then apply DLX algorithm [11] to solve it, which is one kind of efficient solution to solve this kind of covering problem.

If the number of users in Z_i is smaller than $(m+1)$, jump to the fourth step, otherwise, go to the third step.

The Third Step: Add New Repeaters. In this step, we take Greedy Strategy [12] when adding repeaters, in which we retain the optimum choice in the current state in each step. The process is as follows:

Find maximum Q_i

If $Q_i \geq (m+1)$, add another repeater to Z_i to reduce the value of Q_i ;

$0 < Q_i < (m+1)$, consider the 4 candidates in the intersections between Z_i and the circles just

next to it, choose the most suitable candidate among them to put up a new repeater. $Q_i = 0$, go to the next procedure, otherwise, do the foregoing step again.

Repeater-Repeater Connection: Now there may still be some repeaters have no other repeaters in their signal transmission range. For instance, there are three repeaters A, B, C and three candidates X, Y, Z (See Fig.3). Repeater A and B are connected while repeater C is isolated. C can be connected to network which connects A and B through C-X-A or C-Y-Z-B.

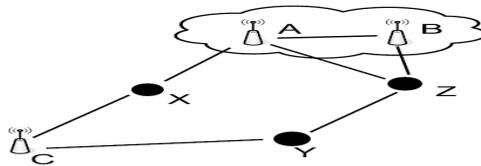


Fig. 3. Two possible link ways of repeater C

Repeaters (A, B, C) are regarded as nodes of an undirected graph. The link ways (C-X-A, C-Y-Z-B) are regarded as edges of the undirected graph, and the number of empty candidates on one edge is the weight of this edge.

We should make all nodes be connected with the minimum weight, which can be described as the Minimum Weight Spanning Tree Problem [13]. For there have been some classical algorithms to solve this problem, we are not going to describe it in detail.

Frequency Distribution: If there is no limitation on the number of frequencies that can be used in one area, we can give each repeater a unique frequency. But when frequencies have to be reused, the co-channel interference between repeaters which handle the same frequency must be taken into consideration. How they interfere with one another can be determined by the signal-to-interference ratio (SIR) [14].



$$\text{For repeater } i: SIR = \frac{R^{-n}}{\sum_{l=1}^L D_l^{-n}} . \quad (1)$$

Where: R is the radius of the signal transmission range of repeater i ;

L is the number of interference sources that repeater i will receive;

n is the decline index, and it generally equals to 4;

D_l is the distance between l^{th} interference source and the repeater i , and in this paper, the interference source is another repeater who transmits signals with the same frequency as repeater i does and is located within the signal transmission range of repeater i .

Generally, the Co-Channel Interference can be ignored if SIR is no smaller than 18dB [14].

We first locate the repeaters without considering the distribution of frequencies, and then we assign frequency to each repeater. Making allowance for the feasibility of the present algorithm, our goal is to minimize the Co-Channel Interference rather than eliminate it when we distribute frequencies after the minimum number and locations of repeaters have been identified.

For the i^{th} repeater, g_i is a 0, 1 variable. When $SIR \geq 18$, $g_i = 0$, otherwise, $g_i = 1$.

$$\text{Our objective function is: } \min \sum_{i=1}^P g_i . \quad (2)$$

Where: P is the number of repeaters.

We adopt the Genetic Algorithm in order to find the optimal solution efficiently.

3 Results and Discussion

Now we consider such a situation where there is a circular flat area with radius of 40 miles. There are 100 radios on the area. The exact locations of those radios are marked by black dots and the radios have been numbered by Arabic numerals as is shown in Fig.4. Only 9 different frequencies (denoted by 1, 2...9) and 5 sub audible tones are available in this area. Based on conventional VHF radio and repeater data, we suppose the signal emission radius of radios and the signal transmission radius of repeaters are 5 miles and 30 miles respectively.

We run our algorithm to get the results. The definite location of each repeater is decided (See Fig.5). The repeaters have also been numbered. The numerals represent the positions of the corresponding repeaters. Black dots are the radios, the same as Fig.4 shows. The radius of each circle is the signal emission range of radios. If the i^{th} radio is landed in the circle that surrounded the j^{th} repeater and the i^{th} radio happens to use the same frequency as the j^{th} repeater does, then the signal emitted by radio i can be recognized by repeater j .

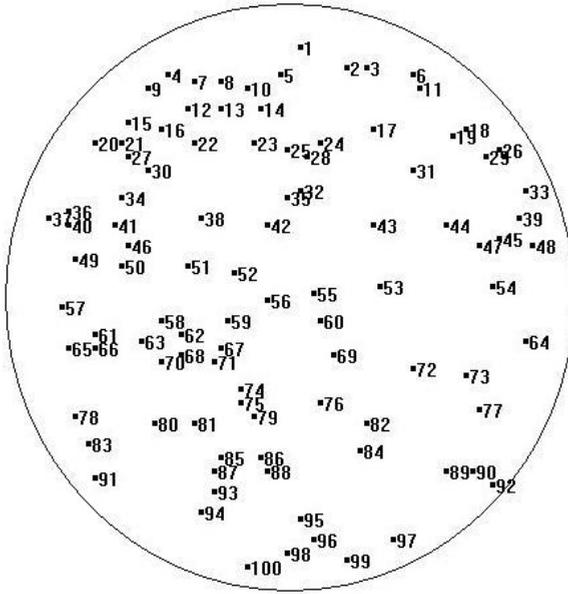


Fig. 4. The locations of the 100 radios

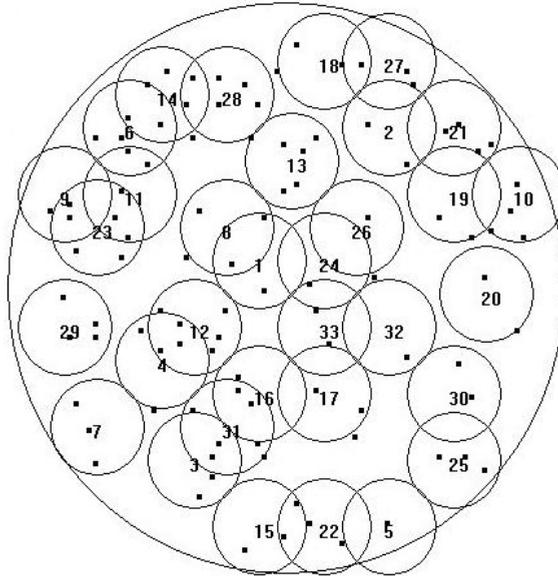


Fig. 5. The locations of repeaters

The minimum number of repeaters we need is 33. The emission frequencies of radios have been distributed (See Tab.1).

Table 1. A partial results of the above case

| radio | the repeater that serve the certain radio | the emission frequency of the radio |
|-------|---|-------------------------------------|
| 1 | 18 | 2 |
| 2 | 18 | 2 |
| 3 | 27 | 1 |
| ... | ... | ... |
| 98 | 15 | 3 |
| 99 | 22 | 8 |
| 100 | 15 | 3 |

Our study also has shortcomings. We take no notice of the geographical situation of the area. We simply suppose the land is flat with no obstacles, thus signal emission ranges of all radios in the area are the same, so as the signal transmission ranges of all repeaters.

Actually, a transmitted signal, modulating a carrier wave, is attenuated and distorted, depending on the specific properties of the transmission circumstances [15]. Many scholars have studied about how the radio signals fade in a specific environment. In [16], the author gives a simplified version of the analytical model to calculate the total path loss L in decibels, which can be applied to urban and suburban environments. As to obstacles such as mountains or trees, the path-loss a signal suffers can be estimated according to blade diffraction model [17].

So when we adopt our models to work out how to serve the radios in one area with the minimum number of repeaters in a real case, the location of each radio should also be taken into consideration. When we choose the candidates as is described above, the determination of S (see Fig.2)—the signal emission range of radios, also equals to the radius of the circles to cover all the radios in the area—can be rather troubled. Our model should be further adjusted on this aspect.

4 Conclusion

Given a flat area with many VHF radios on there, in order to make sure that any two radios can communicate at any time, we establish a complete set of arithmetic to calculate the minimum number of repeaters this area needs, determine their locations and distribute signal transmission frequencies to both radios and repeaters. Then we adopt our arithmetic to a specific case, the outcome of which is rather reasonable. The whole set of algorithm can also be extended to other distribution problems. However, if the geometric environment of this area is complex, the signal emission range of radios will be different from each other, so does the signal transmission range of repeaters. This should be taken into consideration in the future study.

Acknowledgement. This work is supported by the Zhejiang Provincial Natural Science Foundation of China (No. Y1080627).

References

1. Andrusenko, J., Miller, R.L., Abrahamson, J.A., Emanuelli, N.M.M., Pattay, R.S., Shuford, R.M.: VHF General Urban Path Loss Model for Short Range Ground-to-Ground Communications. *Transaction on Antennas and Propagation* 56(10), 3302–3310 (2008)
2. Kemp, C.A.: Extending Mobile Radio Range by VHF Repeaters. *Vehicular Communications* 6, 33 (2006)
3. Pezzini, E., Lovisolo, G.: Narrow Band FM, RF Repeater for Mobile Radio Communications. *Vehicular Technology* 17, 84 (2006)
4. Information on, <http://www.repeater-builder.com/rbtip/repeaterlinking.html>
5. Rabczuk, D.: The Micro-controller System for Remote Control of the Functions of Radio-communication System with CTCSS and Cell-call Signaling, Modern Problems of Radio Engineering. *Telecommunications and Computer Science*, pp. 505–506 (2004)
6. Isberg, R.A.: Radio Communication in Subways and Mines through Repeater Amplifiers and Leaky Transmission Lines. In: *Vehicular Technology Conference*, vol. 28, pp. 248–254 (1978)
7. Jog, S.A.D., Bhagwat, S.S.: A Modern Communication System for Maharashtra Police. In: *Fourth IEEE Region 10 International Conference Tencon 1989*, pp. 538–541 (1989)
8. Dixon, J.S.: Network Topologies for Rural Subscriber Radio Links. *Rural Telecommunications*, 67–72 (1990)
9. Oguchi, K., Hadamada, Y., Minowa, J.: Optical Design and Performance of Wavelength-division-multiplexed Optical Repeater for Fiber-optic Passive Star Networks Connection. *Journal of Lightwave Technology* LT-4(6), 665–671 (1986)
10. Liu, B., Quan, B., Yu, H.: The Distribution of Source and Regional Allocation in Ad_Hoc Network. *The Mathematic Practice And Recognition* (2007)
11. Information on, <http://www-cs-faculty.stanford.edu/~uno/papers/dancing.ps.gz>
12. Information on, <http://www.devarticles.com/c/a/Development-Cycles/Greedy-Strategy-as-an-Algorithm-Technique/1/>
13. Information on, <http://www-cs-faculty.stanford.edu/~uno/papers/dancing.ps.gz>
14. Information on, http://teach.jwc.bupt.cn:4213/mobile/wlkj/C5/C5_chapters/c5.2.2.htm
15. Braun, W.R.: A Physical Mobile Radio Channel Model, *Vehicular Technology Conference, Gateway to the Technology in Motion*, pp. 472–481 (1991)
16. Howard, H.: A Simplified Analytical Model for Predicting Path Loss in Urban and Suburban Environments. *Vehicular Technology*, 1040–1046 (1997)
17. Park, D.-C.: Classification of Audio Signals Using Fuzzy C-Means with Divergence-based Kernel. *Pattern Recognition Letters*, 794–798 (2009)

Research on a Defogging Method of Fog-Degrade Image Based on Depth Region Segmentation

Yi Yang

Department of Computer Application
Henan Business College
Zhengzhou, 450044, Peoples R China
yiyalazihei@hotmail.com

Abstract. According to the effect of fog from images, a new defogging method for image based on depth region segmentation is proposed in this paper. This method reasonably applies the fuzzy cluster analysis thinking, Construct the edge characteristic vector based on multiresolution wavelet. It has achieved the depth region segmentation image based on closed boundary of the target object in the scene, thus attaining the effectively enhanced contrast by the equalizing transform of each histogram in the same depth. Experiment results show that this algorithm can efficiently improve the degradation of image and enhance the clearness of image.

Keywords: Defogging, Fuzzy Cluster Analysis, Edge Detection, Depth Region Segmentation.

1 Introduction

Fog is a common natural phenomenon. It can change or degenerate the contrast and color of the image which is shot outdoor, thus in order to guarantee the video monitor can work normally, it is necessary to do fog processing to the video image.

However, removing the fog is a challenging task, because the depth information that the fog in the atmosphere relays on is unknown to us. And when only one image is available, and it can be analysed, it is still restricted to solve the problem. Therefore, many methods using multiple images or more other auxiliary information are put forward. In [1] and [2] it use many images from angle of polarized light at different degree to remove the interference of the fog. In [3, 4, 5] it is by getting more contrast information from the same scene in different weather situation to remove the fog. In [6, 7] the method based on depth requires rough depth information from the corresponding pictures or known 3D models.

Recently the fog removing method based on a single image has made great progress, more research is the prior single image fog removing algorithm based on the dark primary color [8, 9], this algorithm establishes the fog removing model through a prior, and it achieves images of high quality by estimating the fog concentration directly, but this algorithm needs benefitting from a powerful prior or assumption, because dark unbleached prior is a statistical rule, it might doesn't work well to some special images,

for example, dark unbleached theory is invalid when the scenery is close to the atmosphere and not covered with shadow.

2 The Realization of Mist Removing Method

This paper probes into fog removing method of the image in the weather of mist based on deep regional segmentation. This method achieves its goal by edge detection of the image, deep regional segmentation and enhanced processing of clarifying. In the experiment, good visual effect is achieved.

The Edge Detection for Image of Dropping Quality. Since different scales have different edge characteristics to a drop quality image, it is required to find an edge detection method that can not only reflects the edge of different images adaptively, but also has strong antimonies ability. The clustering theory gives us a feasible solution to the problem. Essentially, the edge detection for image is to find the point set constituted by edge points in the image according to characteristics of edge points.

Supposing a two-dimensional image is $f(x, y)$, and two-dimensional corresponding scale function is $\gamma(x, y)$, then the corresponding two-dimensional wavelet function is:

$$\psi^1(x, y) = \partial[\gamma(x, y)] / \partial x \tag{1}$$

$$\psi^2(x, y) = \partial[\gamma(x, y)] / \partial y \tag{2}$$

Then the corresponding wavelet transformation respectively in the direction of x and y of the scale i of function $f(x, y)$ is

$$W_{2^i}^1 f(x, y) = f(x, y) * \psi_{2^i}^1(x, y) \tag{3}$$

$$W_{2^i}^2 f(x, y) = f(x, y) * \psi_{2^i}^2(x, y) \tag{4}$$

According to the feature that wavelet transformation modulus maxims of the signal and the noise in different scales signal and noise have distinct characteristics of propagation, it is available to make wavelet coefficients of adjacent scales multiply to enhance the signal and reduce the noise.

If $Cor_i(x, y) = W_{2^i} f(x, y) \cdot W_{2^{i+1}} f(x, y)$ is the correlation coefficient in the point (x, y) of adjacent scales i and $i + 1$, and supposing $i = 1, i = 2, i = 3$, then $Cor_1(x, y), Cor_2(x, y), Cor_3(x, y)$ form the eigenvector $\{Cor_1(x, y), Cor_2(x, y), Cor_3(x, y)\}$ of the edge point.

Taking every pixel in the image as a data sample, because grey values of edge points and non -edge points present a bigger difference after the multi-scale wavelet transformation, so a three -dimensional characteristic vector $k(x, y)$ is constituted by three characteristic vectors:

$$Cor_1(x, y), Cor_2(x, y), Cor_3(x, y) \quad (5)$$

Here, $k(x, y) = \{Cor_1(x, y), Cor_2(x, y), Cor_3(x, y)\}$. Then, use genetic fuzzy clustering algorithm to class the data set, what one kind of them have described is just the set of edge points. Like this, we can detect the edge points of the image adaptively to achieve the goal of extracting the edge.

The Regional Segmentation of Image of Dropping Quality. Next, we should do deep regional segmentation to the drop quality image according to the detection of edge points. Because the conversation between the goals and objectives or the target and the background in the image might cause the shift of the image depth, so the task to divide the drop quality image into different regions at the same depth can be achieved by segment the image.

Assuming the pixels in every target of the drop quality image are always at the same depth, and the background region is also considered as a target, like this, we can divide the drop quality image into several targets, several different depth regions. Then the depth regional segmentation of the image is achieved.

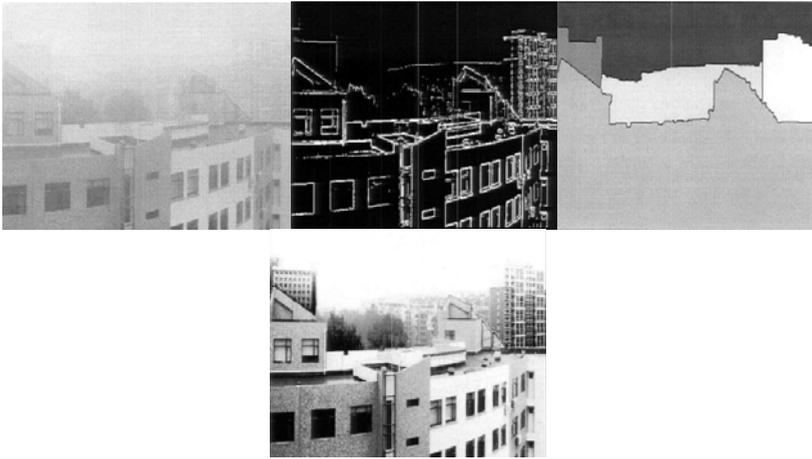
After the edge point's detection of the drop quality image, it is necessary to connect the edge points to form a target boundary. What we should pay attention to is when multiple small closed boundaries appear in the connection process, in order to avoid the phenomenon of too many divided regions, we can merge some small closed regions which have similar regional characteristics. The final task is to create mark matrix, mark the deep regions according to the target boundaries in the marked matrix, so you can achieve the depth regional segmentation of the drop quality image.

Biaxial Stretching Enhancement. Using global histogram equalization to realize the contrast enhancement processing of each depth region, after segmentation processing of depth the only thing we need to do is the enhanced processing of the contrast with respect to regions at same depth. Histogram equalization, as an effective method to enhance the global contrast, is to do smooth processing to image histograms to maximize primogeniture. It have the advantages of simple operation, good real-time performance and the good treatment effect. In the experiment, we realize enhanced processing to every depth area by this method. Since histogram equalization stretches the grey range of the image to $[0,255]$, when the region is sky, we directly assume it as 255, instead of enhanced processing.

3 Experimental Results and Analyses

In order to verify the effectiveness of the algorithm proposed in this paper, in MATLAB platform we have done a large number of experiments.

Figure 1 (a) is image of dropping quality filmed in fog weather. Because of the influence of the mist, the contrast of image declines, especially for the corresponding in images distant area, and edge profile has turned already very fuzzy. Through the edge detection of drop quality image, we got figure 1 (b), it can be seen from the graph that the edge of the drop quality image is clearer, especially for the outline, as makes



(a) Fog-degraded images (b) Edge detection (c) Depth regional segmentation (d) Stretching effect

Fig. 1. Fog removing experiment results based on depth segmentation



(a) Fog-degraded images (b) Histogram equalization (c) Retinex algorithm (d) The new algorithm

Fig. 2. Fog removing effect of different algorithms in misty condition

good preparation for the figure 1 (c) drop quality image depth regional segmentation. Figure 1 (d) is an enhanced biaxial stretching fog-removing image, from mist-removing effect, image contrast of close range and distant prospect got obvious improvement compared with the primary image, and visual effect becomes better.

In order to further validate the validity and superiority of the method proposed in this paper, respectively we use histogram equalization algorithm and single yardstick Retinex algorithm and use this algorithm to do the mist-removing process of an image

in dense fog, and we can compare them in figure 2. Visibility, the image processed through histogram equalization in figure 2 (b), nearby scene gray level becomes lower, and contrast drops, which is obviously not what we expected. Figure 2 (c) is the treatment effect of Retinex algorithm, this algorithm always does well in dealing with illumination, but it often has no obvious advantages when handling with mist image, although the boundary of processed image has obvious enhancement, but it can't eliminate the feeling of "the mist". The image processed with this algorithm just shown as figure 2 (d), the nearby and distant area both got different enhancement, and the scenery is trenchant, especially that the outlines of the trees in distant are clear, therefore it presents a good effect on the mist-removing.

4 Conclusion

From the experiment which uses improved algorithm to remove fog of the image in the mist, we can see that the new method doesn't need accurate information of the weather and scene, and it can effectively realize the fog removing processing to a image of drop quality, and can enhance contrast degree of degraded images. Images after the processing have the characteristics of high definition, low color distortion and simple operation. The proposed algorithm can effectively enhance details, and remove the fog naturally and integrally; meanwhile, local details are obvious. In a word, this is an effective method to remove the mist.

References

1. Schechner, Y.Y., Narasimhan, S.G., Nayar, S.K.: Instant dehazing of images using polarization. CVPR 1, 325 (2001)
2. Shwartz, S., Namer, E., Schechner, Y.Y.: Blind haze separation. CVPR 2, 1984–1991 (2006)
3. Narasimhan, S.G., Nayar, S.K.: Chromatic framework for vision in bad weather. CVPR, 598–605 (January 2000)
4. Narasimhan, S.G., Nayar, S.K.: Contrast restoration of weather degraded images. PAMI 25, 713–724 (2003)
5. Nayar, S.K., Narasimhan, S.G.: Vision in bad weather. In: ICCV, p. 820 (January 1999)
6. Kopf, J., Neubert, B., Chen, B., Lischinski, D.: Deep photo Model-based photograph enhancement and viewing. In: SIGGRAPH, Asia (January 2008)
7. Narasimhan, S.G., Nayar, S.K.: Interactive dewatering of an image using physical models. In: Workshop on Color and Photometric Methods in Computer Vision (January 2003)
8. Tan, R.: Visibility in bad weather from a single image. CVPR (January 2008)
9. Fattal, R.: Single image dehazing. In: SIGGRAPH, pp. 1–9 (January 2008)

An Algorithm of Incremental Construction Voronoi Region

Pengfei Xu^{1,2}, Zhigang Chen¹, and Ming Chen²

¹ College of Information and Engineering, Central South University, ChangSha, China

² College of Mathematics and Computer Science, Hunan Normal University, ChangSha, China
pkrysl@yahoo.cn

Abstract. This paper proposes an algorithm for the construction of Voronoi region. First, the efficient clipping rules are presented to incrementally construct Voronoi region. Then the rules are modified and restricted in order to deal with co-circular nodes and collinear nodes.

Keywords: Computational Geometry, Voronoi Region, Incremental, co-circular, collinear.

1 Introduction

The Voronoi Tessellation and the Delaunay Triangulation are dual graphs each other in computational geometry [1], some efficient algorithms were proposed to construct the Voronoi Tessellation or the Delaunay triangulation [2-4]. However, some distributed applications in Ad hoc networks only need to construct the Voronoi region or Delaunay triangles incident on one node[5-8]. Therefore, some algorithms were proposed to construct the Voronoi region [5,6]. Wang 5 could not deal with collinear nodes. Zhang 6 just could construct the closed Voronoi region. In addition, Wang 5 and Zhang 6 did not consider co-circular nodes and collinear nodes. This paper proposes a simple algorithm to incrementally construct the Voronoi region, which can also deal with co-circular nodes and collinear nodes.

2 Preliminaries

Assume that S is a set of nodes in the plane R^2 where no two nodes are overlapping. Let $L(u,k)$ be the perpendicular bisector line of the node u and k , and $H(u,k)$ be the half-plane bounded by the line $L(u,k)$ and containing the node u .

The Voronoi Tessellation for S partitions the plane R^2 into some convex regions. Each convex region, called the Voronoi region, contains only one node of S . The Voronoi region incident on the node $u \in S$, denoted by $V(S,u)$, is the intersection of $n-1$ half-planes,

$$V(S,u) = \bigcap_{k \in S, k \neq u} H(u,k) \quad (1)$$

The boundary of the Voronoi region is called the *Vor edge*. The intersection point of two *Vor edges* is called the *Vor vertex*. The node u and k are *Vor neighbors* each other if and only if $V(S,u)$ and $V(S,k)$ share the common *Vor edge*. The *Vor edge* of $V(S,u)$ shared with $V(S,k)$, denoted by $V(S,u)(k)$, may be a line-segment, half-line or infinite-line of the line $L(u,k)$.

For convenience, the infinity on the half-line or the infinite-line are abstracted as virtual points. Then any *Vor edge* can be viewed as a line-segment between two *Vor vertices*. Traversing the *Vor edges* of $V(S,u)$ in anticlockwise order, let k_1 and k_2 be the *first Vor vertex* and the *second Vor vertex* of the *Vor edge* $V(S,u)(k)$ respectively, then the *Vor edge* $V(S,u)(k)$ is also denoted by $V(S,u)(k)(k_2 \leftarrow k_1)$, as Fig.1(a) shows.

3 Incremental Construction Voronoi Region

The main idea of Incremental Construction Voronoi Region is, given the Voronoi region $V(S,u)$, to build the Voronoi region $V(S/z,u)$ after adding a new node z into the plane R^2 , where $u \in S, z \notin S$ and $S/z = S + \{z\}$. From Eq.1, $V(S/z,u)$ can also be rewritten as the following,

$$V(S,u) \cap k \in (S + \{z\}), k \neq u \cap H(u,k) = \cap k \in S, k \neq u \cap H(u,k) \cap H(u,z) \tag{2}$$

Eq.2 implies that $V(S/z,u)$ is the intersection between the region $V(S,u)$ and the half-plane $H(u,z)$, as the shadow of Fig.1 shows.

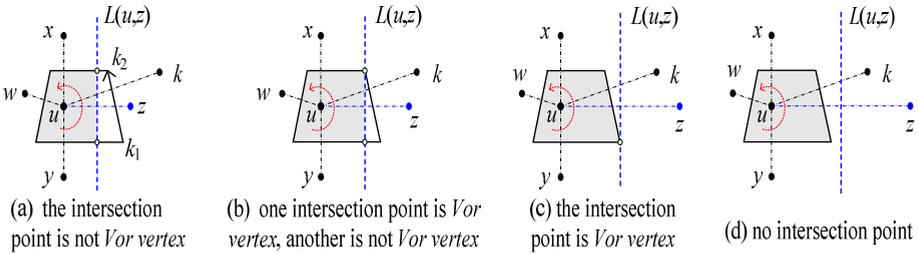


Fig. 1. Voronoi Region and Half-plane

When the Voronoi region is characterized by its *Vor edges* and *Vor neighbors*, building the Voronoi region $V(S/z,u)$ from the Voronoi region $V(S,u)$ can be simplified to discard all segments on the *Vor edges* of $V(S,u)$ which are not in the half-plane $H(u,z)$, instead of them is one segment of the line $L(u,z)$ which is in the region $V(S,u)$, where the line $L(u,z)$ is the common boundary between the half-plane $H(u,z)$ and $H(z,u)$.

4 The Basic Clipping Rules

When no four nodes of S/z are co-circular and the nodes of S/z are not collinear, the line $L(u,z)$ either goes through the Voronoi region $V(S,u)$ intersecting with at least one of its



Vor edge but not on its *Vor vertex*(as Fig.1(a) shows), or is outside the Voronoi region $V(S,u)$ without intersecting with any of its *Vor edge*(as Fig.1(d) shows). Consequently, the *Vor edge* $V(S,u)(k)(k_2 \leftarrow k_1)$ and the line $L(u,z)$ would be clipped by one of the following basic rules.

Rule I $k_1 \in H(z,u)$ and $k_2 \in H(u,z)$: the *Vor edge* $V(S,u)(k)(k_2 \leftarrow k_1)$ intersects with the line $L(u,z)$ on the point p , as Fig.2(a) shows. The segment pk_1 which is not in the half-plane $H(u,z)$ should be discarded, that is, $V(S/z,u)$ has the *Vor edge* $V(S/z,u)(k)(k_2 \leftarrow p)$, and it is be sure that the point p is the *second Vor vertex* of the new *Vor edge* $V(S/z,u)(z)$.

Rule II $k_1 \in H(u,z)$ and $k_2 \in H(z,u)$: the *Vor edge* $V(S,u)(k)(k_2 \leftarrow k_1)$ intersects with the line $L(u,z)$ on the point p , as Fig.2(b) shows. The segment k_2p which is not in the half-plane $H(u,z)$ should be discarded, that is, $V(S/z,u)$ has the *Vor edge* $V(S/z,u)(k)(p \leftarrow k_1)$, and it is be sure that the point p is the *first Vor vertex* of the new *Vor edge* $V(S/z,u)(z)$.

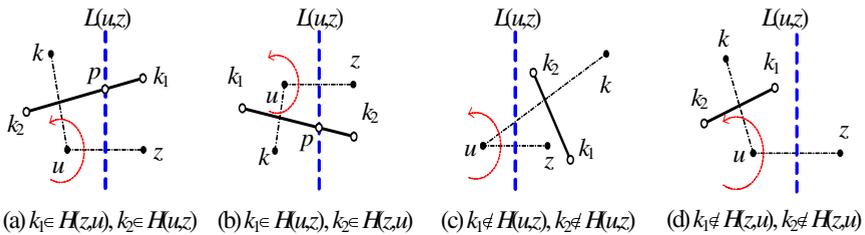


Fig. 2. The *Vor edge* and the line $L(u,z)$

Rule III $k_1 \notin H(u,z)$ and $k_2 \notin H(u,z)$: the *Vor edge* $V(S,u)(k)(k_2 \leftarrow k_1)$ is not in the half-plane $H(u,z)$, as Fig.2(c) shows. The segment k_2k_1 should be discarded, that is, $V(S/z,u)$ has not the *Vor edge* $V(S/z,u)(k)$, but it is be sure that $V(S/z,u)$ will have the new *Vor edge* $V(S/z,u)(z)$.

Rule IV $k_1 \notin H(z,u)$ and $k_2 \notin H(z,u)$: the *Vor edge* $V(S,u)(k)(k_2 \leftarrow k_1)$ is in the half-plane $H(u,z)$, as Fig.2(d) shows. Thus, $V(S/z,u)$ has the *Vor edge* $V(S/z,u)(k)(k_2 \leftarrow k_1)$, but it is not sure whether $V(S/z,u)$ will have the new *Vor edge* $V(S/z,u)(z)$.

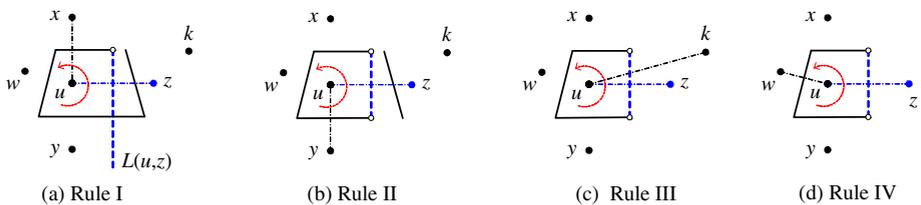


Fig. 3. The clipping process



According to the followed Basic Rules, **Rule I**, **Rule II** and **Rule III** make sure that $V(S/z,u)$ will have the new *Vor edge* $V(S/z,u)(z)$, but its *Vor vertex* will be decided by all *Vor edges* of $V(S,u)$ satisfying **Rule I** and **Rule II**. For the Voronoi region $V(S,u)$ showed in Fig.1(a), Fig.3 gives its detailed clipping process.

On the other hand, if all *Vor edges* of $V(S,u)$ satisfy **Rule IV**, then $V(S/z,u)=V(S,u)$, that is, $V(S/z,u)$ and $V(S,u)$ have uniform *Vor edges* and *Vor neighbors*. For example, as the Voronoi region showed in Fig.1(d), all *Vor edges* of its satisfy **Rule IV**, then $V(S/z,u)=V(S,u)$.

5 Co-circular Nodes

From Fig.1(b) and Fig.1(c), the line $L(u,z)$ can also intersect with the *Vor edge* of $V(S,u)$ on its *Vor vertex*. Thus, no less than four nodes of S/z are co-circular. On the other hand, the nodes of S/z are not collinear.

Four co-circular nodes. Assume that the *Vor edge* $V(S,u)(k)(k2←p)$ and $V(S,u)(m)(p←m1)$ intersect on the *Vor vertex* p , that is, p is the intersection point of the line $L(u,k)$ and $L(u,m)$. Adding the new node z , assume that the point p is on the line $L(u,z)$, that is, the nodes $\{z,u,m,k\}$ are co-circular. If the *Vor edge* $V(S,u)(k)(k2←p)$ and $V(S,u)(m)(p←m1)$ are on the half-plane $H(z,u)$, then the convex region formed by the segment $pk2$ and $pm1$ can not contain the node u , which is a contradiction to the definition of the Voronoi region. Consequently, the *Vor edge* $V(S,u)(k)(k2←p)$ and $V(S,u)(m)(p←m1)$ will be one of the following cases.

Co-circular case 1: Two *Vor edges* are on both sides of the line $L(u,z)$, as Fig.4(a) shows. Because all *Vor edges* of $V(S,u)$ form a continuous convex polygon when all nodes of S are not collinear1, the *Vor edge* $V(S,u)(k)(k2←p)$ and $V(S,u)(m)(p←m1)$ which intersect with the line $L(u,z)$ satisfy **Rule I**, other *Vor edges* of $V(S,u)$ which intersect with the line $L(u,z)$ satisfy **Rule II**, and any *Vor edge* of $V(S,u)$ which does not intersect with the line $L(u,z)$ satisfies either **Rule III** or **Rule IV**. Thus, $V(S/z,u)$ has the *Vor edge* $V(S/z,u)(k)(k2←p)$ and $V(S/z,u)(m)(p←p)$, the point p is also the *second Vor vertex* of the new *Vor edge* $V(S/z,u)(z)$. Obviously, the clipping results can accord with Eq.2. Hereafter, the *Vor edge* which is one point is also called the *point-Vor edge*, such as $V(S/z,u)(m)(p←p)$.

Co-circular case 2: Two *Vor edges* are on the half-plane $H(u,z)$, as Fig.4(b) shows. Because $V(S,u)$ is a convex region, the *Vor edge* $V(S,u)(k)(k2←p)$ which intersect with the line $L(u,z)$ satisfies **Rule I**, the *Vor edge* $V(S,u)(m)(p←m1)$ which intersect with the line $L(u,z)$ satisfies **Rule II**, other *Vor edges* of $V(S,u)$ must not intersect with the line $L(u,z)$ and satisfy **Rule IV**. Thus, $V(S/z,u)$ has the *Vor edge* $V(S/z,u)(k)(k2←p)$, $V(S/z,u)(m)(p←m1)$, and the new *point-Vor edge* $(S,u)(z)(p←p)$. Obviously, the clipping results can accord with Eq.2.

According to the followed analyses, the Basic Rules can correctly deal with four co-circular nodes. Observe that the clipping results in Fig.4(a) and Fig.4(b) both have one *point-Vor edge*, that is, the clipping results of four co-circular nodes is similar.



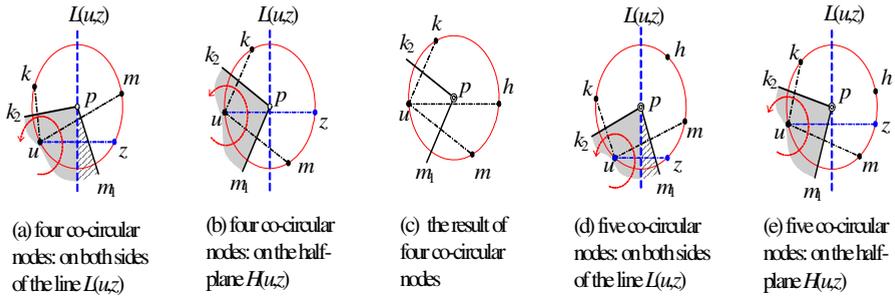


Fig. 4. Co-circular nodes

For simplicity, denote the nodes set S/z and the node z in Fig.4(b) as the nodes set S and the node h respectively. That is, $V(S,u)$ has Vor edges $V(S,u)(k)(k1 \leftarrow p)$, $V(S,u)(h)(p \leftarrow p)$ and $V(S,u)(m)(p \leftarrow m1)$, as Fig.4(c) shows.

Five co-circular nodes. Adding another new node z to Fig.4(c), assume that the point p is on the line $L(u,z)$, that is, the nodes $\{z,u,m,k,h\}$ are co-circular. The point-Vor edge $V(S,u)(h)(p \leftarrow p)$ on the line $L(u,z)$ satisfies **Rule I** and **Rule II**. Then the Vor edge $V(S,u)(k)(k1 \leftarrow p)$ and $V(S,u)(m)(p \leftarrow m1)$ will be one of the following cases.

Co-circular case 3: Two Vor edges are on both sides of the line $L(u,z)$, as the Fig.4(d) shows. Without considering the point-Vor edge $V(S,u)(h)(p \leftarrow p)$, this case is similar to the four co-circular case 1. Whether **Rule I** or **Rule II** clips the Vor edge $V(S,u)(h)(p \leftarrow p)$, $V(S/z,u)$ will have the point-Vor edge $V(S/z,u)(h)(p \leftarrow p)$, but **Rule II** will mistake the point p as the first Vor vertex of the new Vor edge $V(S/z,u)(z)$.

Co-circular case 4: Two Vor edges are on the half-plane $H(u,z)$, as Fig.4(e) shows. Without considering the point-Vor edge $V(S,u)(h)(p \leftarrow p)$, this case is similar to the four co-circular case 2. Whether **Rule I** or **Rule II** clips the Vor edge $V(S,u)(h)(p \leftarrow p)$, $V(S/z,u)$ will have the point-Vor edge $V(S/z,u)(h)(p \leftarrow p)$, and the line $L(u,z)$ can also be correctly clipped to the point p .

Restricted the Basic Clipping Rules. Similar to the followed analyses, if the node u , its Vor neighbors and the new node z , are $m(\geq 4)$ co-circular nodes, then $m-4$ point-Vor edges of $V(S,u)$ are the same point on the line $L(u,z)$. Without considering the point-Vor edges on the line $L(u,z)$, the Basic Rules can correctly clip all other Vor edges of $V(S,u)$ and the line $L(u,z)$. Although **Rule I** and **Rule II** can both take the point-Vor edge on the line $L(u,z)$ as the point-Vor edge of $V(S/z,u)$, **Rule II** will incorrectly clip the line $L(u,z)$. In order to deal with co-circular nodes, all point-Vor edges of $V(S,u)$ on the line $L(u,z)$ should be clipped by **Rule I** but not **Rule II**.

6 Collinear Nodes

When S has only the node u , the Voronoi region $V(S,u)$ is the plane R^2 and has no Vor edge. After adding the new node z , the node u and z are collinear, $V(S/z,u)$ is the

half-plane $H(u,z)$ whose *Vor edge* is the line $L(u,z)$. This belongs to the initialization. The following assumes that the nodes of S/z are collinear and its number is no less than two. When the nodes of S are collinear, the *Vor edge* of $V(S,u)$ is infinite-line or two parallel infinite-lines1.

Only Vor edge $V(S,u)(k)(k1 \leftarrow k2)$. $V(S,u)$ is the half-plane $H(u,k)$, that is, $V(S,u)(k)(k1 \leftarrow k2) = L(u,k)$. Thus, the node u, k and z will be one of the following cases:

Collinear case 1 $z \in uk$: as Fig.5(a) shows, from Eq.2, $V(S/z,u)$ is the half-plane $H(u,z)$ whose *Vor edge* is the line $L(u,z)$. The virtual *Vor vertex* $k1$ and $k2$ satisfy $k1 \notin H(u,z)$ and $k2 \notin H(u,z)$. From **Rule III**, the *Vor edge* $V(S,u)(k)(k1 \leftarrow k2)$ will be discarded, and the line $L(u,z)$ becomes the only *Vor edge* of $V(S/z,u)$. Obviously, the clipping results are in accord with Eq.2.

Collinear case 2 $k \in uz$: as Fig.5(b) shows, from Eq.2, $V(S/z,u)$ is the half-plane $H(u,k)$ whose *Vor edge* is the line $L(u,k)$. The virtual *Vor vertex* $k1$ and $k2$ satisfy $k1 \notin H(z,u)$ and $k2 \notin H(z,u)$. From **Rule IV**, $V(S/z,u)$ has the *Vor edge* $V(S/z,u)(k)(k1 \leftarrow k2)$, i.e. the line $L(u,k)$. Obviously, the clipping results are in accord with Eq.2.

Collinear case 3 $u \in kz$: as Fig.5(c) shows, from Eq.2, $V(S/z,u)$ is the region between the line $L(u,k)$ and $L(u,z)$. The virtual *Vor vertex* $k1$ and $k2$ satisfy $k1 \notin H(z,u)$ and $k2 \notin H(z,u)$. From **Rule IV**, $V(S/z,u)$ has the *Vor edge* $V(S/z,u)(k)(k1 \leftarrow k2)$, i.e. the line $L(u,k)$; but it can not be sure that the line $L(u,z)$ is the new *Vor edge* of $V(S/z,u)$. Obviously, the clipping results can not be in accord with Eq.2.

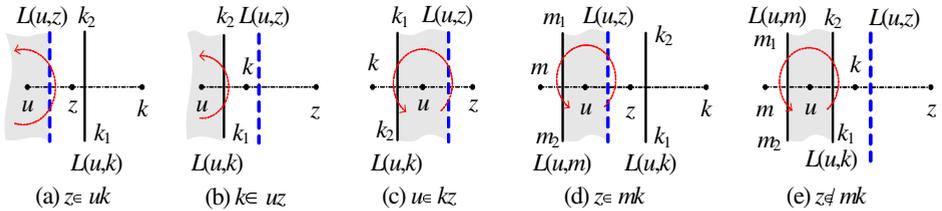


Fig. 5. Collinear Nodes

Two parallel Vor edges $V(S,u)(k)(k1 \leftarrow k2)$ and $V(S,u)(m)(m1 \leftarrow m2)$. $V(S,u)$ is the region between the line $L(u,m)$ and $L(u,k)$, that is, $V(S,u)(m)(m1 \leftarrow m2) = L(u,m)$ and $V(S,u)(k)(k1 \leftarrow k2) = L(u,k)$. Thus, the node z, m and k will be one of the following cases.

Collinear case 4 $z \in mk$: as Fig.5(d) shows, from Eq.2, $V(S/z,u)$ is the region between the line $L(u,m)$ and $L(u,z)$. The virtual *Vor vertex* $m1$ and $m2$ satisfy $m1 \notin H(z,u)$ and $m2 \notin H(z,u)$. From **Rule IV**, $V(S/z,u)$ has the *Vor edge* $V(S/z,u)(m)(m1 \leftarrow m2)$, i.e. the line $L(u,m)$. The virtual *Vor vertex* $k1$ and $k2$ satisfy $k1 \notin H(u,z)$ and $k2 \notin H(u,z)$. From **Rule III**, the *Vor edge* $V(S,u)(k)(k1 \leftarrow k2)$ will be discarded, and the line $L(u,z)$ becomes one new *Vor edge* of $V(S/z,u)$. Obviously, the clipping results are in accord with Eq.2.



Collinear case 5 $z \notin mk$: as Fig.5(e) shows, from Eq.2, $V(S/z,u)$ is the region between the line $L(u,m)$ and $L(u,k)$, i.e. $V(S/z,u)=V(S,u)$. The virtual Vor vertex $m1$ and $m2$ satisfy $m1 \notin H(z,u)$ and $m2 \notin H(z,u)$, the virtual Vor vertex $k1$ and $k2$ satisfy $k1 \notin H(z,u)$ and $k2 \notin H(z,u)$, that is, all Vor edges of $V(S,u)$ satisfy **Rule IV**, then $V(S/z,u)=V(S,u)$. Obviously, the clipping results are in accord with Eq.2.

Modifying the Basic Clipping Rules. From the followed cases that the nodes of S/z are collinear, the Basic Rules can correctly deal with all cases other than the collinear case 3, i.e. Fig. 5(c). In Fig.5(c), the node u is on the segment kz (i.e. $u \in kz$), and the Vor edge $V(S,u)(k)(k1 \leftarrow k2)$ satisfies $k1 \notin H(z,u)$ and $k2 \notin H(z,u)$. This case can be viewed as the special instance of **Rule IV**. Consequently, **Rule IV** can be rewritten as the following.

Rule IV* $k1 \notin H(z,u)$ and $k2 \notin H(z,u)$: (1) If $u \in kz$ and $V(S,u)$ has only one Vor edge, then $V(S/z,u)$ has two parallel Vor edges $V(S/z,u)(k)(k1 \leftarrow k2)$ and the line $L(u,z)$; (2) otherwise, $V(S/z,u)$ has the Vor edge $V(S/z,u)(k)(k1 \leftarrow k2)$, but it is not sure whether $V(S/z,u)$ will have the new Vor edge $V(S/z,u)(z)$.

When the nodes of S/z are collinear, Fig.5(b), Fig.5(d) and Fig.5(e) also have the Vor edge satisfying **Rule IV**, but Fig.5(b) satisfies $u \notin kz$, Fig.5(d) and Fig.5(e) have two Vor edges. On the other hand, when the nodes of S/z are not collinear, the nodes $\{u,k,z\}$ satisfy $u \notin kz$ even if $V(S,u)$ has only one Vor edge and it satisfies **Rule IV**. Therefore, **Rule IV*** will not affect other clipping cases.

7 Conclusions

In this paper, the half-plane intersection set of the Voronoi region is transformed into the relationship between Vor vertex and half-plane, and the efficient clipping Rules are presented to incrementally construct Voronoi region. Then, the clipping Rules are modified in order to deal with co-circular nodes. Finally, the clipping Rules is restricted in order to deal with collinear nodes.

Acknowledgments. This work is partially supported by the National Science Foundation of China under Grant 60873082, and the Youth Scientific Research Fund of Hunan Normal University of China under grand number 60901.

References

1. de Bertg, M., van Kreveld, M.: Computational geometry algorithms and applications. Springer, New York (2000)
2. Evazi, M., Mahani, H.: Transport in Porous Media 83, 541 (2009)
3. Borut, A.: Computer Aided Design 37, 1027 (2005)
4. Jean-Daniel, B., Olivier, D., Samuel, H.: Proceedings of the 25th Annual Symposium on Computational Geometry, June 8-10, vol. 1, p. 208. Aarhus University, Denmark (2009)

5. Wang, J., Sirisha, M.: Proceeding of Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, New York, USA, October 8-10, vol. 1, p. 61 (2007)
6. Zhang, C., Zhang, Y.C.: Wireless Networks 15, 3 (2009)
7. Lee, D.-Y., Lam, S.S.: Proceeding of 27th International Conference on Distributed Computing Systems, Toronto, Canada, June 25-27, vol. 1, p. 26 (2007)
8. Satyanarayana, D., Rao, S.V.: Proceeding of First international conference on Emerging Trends in Engineering and Technology, Nagpur, Maharashtra, July 16-18, vol. 1, p. 282 (2008)

A Brief Analysis of Geocoding

Lijing Zhang¹ and Jing Yi²

¹ Information & Network Management Center North China Electric Power University
Baoding, China

² Department of Computer Science North China Electric Power University Baoding, China
lijing166@sina.cn

Abstract. With the rapid development and widely application of the Geographic Information System (GIS), and spatial location information plays a more and more important role, and spatial information service has been one of the focuses of social development. The paper describes geocoding technologies and its process, and it is emphatically introduced about several major geocoding methods. It also analyzes China's current geocoding situation and introduces some methods to improve geocoding.

Keywords: Geocoding, Geocoding Methods, Address Matching.

1 Introduction

It can be found by analyzing existing information systems: non-spatial data resources have specific occurred, which is a key link between non-spatial data resources and spatial data. However, X and Y coordinates are necessary to display the data in GIS, because the system can know where the data should be displayed on the map [1]. How are geo-spatial coordinates given to these non-spatial data resources? Geocoding technology can achieve this process; it can identify the location of data resources in the geospatial reference range so that it establishes the link between non-spatial information and spatial information, and therefore it realizes information integration in various address space range (i.e., administrative area, census area, street) [2].

1.1 The Concept of Geocoding

The concept of geocoding is different from the general coding definitions, and it does not use numbers or letters to represent a certain surface features, but is the process of mapping the address data into geographical coordinates. In other words, it establishes the link between address data containing spatial location information and the geographical coordinates so as to make these data into the coordinates which can be used in GIS. Geocoding, also known as address coding, is a kind of coding method based on positioning technology, which looks at ways to change the natural language description of spatial location information into geographical coordinates (Fig.1), such as the daily address, address numbers, building names and company names. It can also be considered that geocoding refer to establish the link between spatial information and spatial location information. Geo-coding makes it possible to determine the location of

geographical entities which are represented by these address data, and it is a bridge of the realization of analysis of social economic information based on the spatialization as well as visualization of GIS.

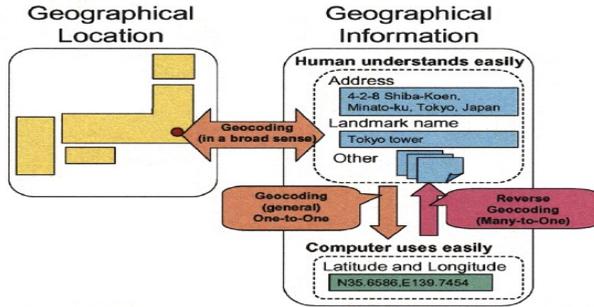


Fig. 1. Overview of Geocoding

1.2 The Process of Geocoding.

The principles of the process (Fig.2) is: split and standardize the strings of the attribute records which need encoding, then use these key values to match with the corresponding attribute values of entities in spatial reference data. If it succeeds, give the geographical coordinates of these corresponding entities to them, and the encoding process is completed. Therefore, the process typically contains address standardization and address matching [2].

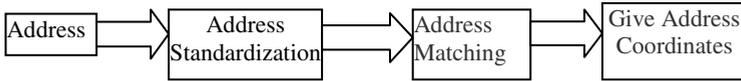


Fig. 2. The Process of Geocoding

1.2.1 Address Standardization

Address standardization makes the general form of address strings into structured word groups the computer can understand. In short, it is the normalization before encoding the address [4]. It is necessary for address string conversion to make use of address dictionary. Chinese address dictionary can be received from “Chinese place with a thesaurus”. The composition form of phrase structure is address model. Different countries and regions have different the expression of address (address model). It usually describes mostly addresses according to the address hierarchy, so we can establish the address model based on the address hierarchy. Address segmentation is the key to address standardization. It segments the addresses into different word groups according to address strings, address dictionary and address model [6]. The decomposition of English address can follow space and punctuation, while Chinese addresses are decomposed with the help of address dictionary.



1.2.2 Address Matching

Address Matching is the core technology of geocoding. It just matches addresses after address Standardization with records of address database according to certain rules so as to result in the process of matching. In the process of geocoding, it needs to match two types of data: one is address data (such as street addresses, zip code, administrative division), which only contain geographical entities, and do not have any location information (geographical coordinate). Other is the data which include the geo-referenced data (it contains street data, zip code) with map position information (geographical coordinate), and the data collection or database is the role of spatial reference in matching process, and is the most central part of geocoding.

At present, address matching and positioning software related to geocoding is ArcInfo geocoding modules and MapInfo Map-Marker. The definition of geocoding in ArcInfo is that it adds address attributes into geographic features to identify a location after entering the address. In MapInfo, it is the process of giving the geographical coordinate to records. By address matching, it can connect the address database to map database, and give map location information to address data in address database.

2 Application of Geocoding

If it lacks the spatial reference information the computer cannot understand and display it on the map. Geocoding just gives the certain entry the X and Y coordinates to identify where the entries are in the earth surface. Therefore, it improves the readability of the data, and enables addresses to be displayed on the map so that it can be easy to dispose and use for researchers, planners, developers and analysts, and integrate information in a variety of spatial scales.

It distinguishes the geocoding accuracy in accordance with the range and precision expressed by address data. The method of high-accuracy is by encoding the data which belong to the level of the street address, and this method is generally applicable to systems which need high accuracy results, such as GPS; the method of low-accuracy uses the data which belong to the level of the zip code to match and map. The more detailed the address level is, the higher accuracy result of geocoding it gets. For example, the geocoding results of district-wide demographic range of the population have higher accuracy result than the results of zip code range.

2.1 Geocoding by Street Address

According to street number information described by records, geocoding by street address changes records into points to position to both sides of street. Firstly, match the street name with the records in database. If it works, examine whether the specified street number belongs to the number range of the street. If it succeeds, give the address the geographical coordinates by using data approximation method. For example (Table 1), the relationship between the record and a street known as People Avenue is described in the following table. There are two records about this street in the table, and one's street number range in the left side is from 1 to 11, the range in the right side is from 2 to 20; other's street number range in the left side is from 13 to 21, the range

in the right side is from 22 to 28. If the checked address is 20 People Avenue, the value of the record ID of the corresponding data in the table is 15.

Attribute Description: ID refers to the sequence number of records in the table; FullName refers to the name of the full name; L_F_ADD refers to the left first street number; L_T_ADD refers to the left last street number; R_F_ADD refers to the right first street number; R_T_ADD refers to the right last street number.

Table 1. The Basic Data

| ID | FullName | L_F_ADD | L_T_ADD | R_F_ADD | R_T_ADD |
|----|---------------|---------|---------|---------|---------|
| 15 | Renmin Street | 1 | 11 | 2 | 20 |
| 16 | Renmin Street | 13 | 21 | 22 | 28 |

2.2 Geocoding by Postal Code

Geocoding by postal code matches the zip code information of each record with the zip code entities the database, and then it changes records into points to position to the center of the zip code regional entities to achieve encoding. In Canada, the Postal Code Conversion File (PCCF) also provides a correspondence between the postal code and Statistics Canada's standard geographic areas for which census data are produced. Through the link between postal codes and standard geographic areas, the PCCF permits the integration of data from various sources. According to the Canada Post Corporation, the postal code consists of six characters (e.g.B0S 1M0) of which the first three are known as Forward Sortation Areas (FSAs). Forward Sortation Areas are large polygons with distinct boundaries, whereas postal codes do not have boundaries even though they represent areas.

Geocoding based on a postal code produces radically different results in urban and rural areas. Urban postal codes represent very small areas as they approximate a block face-one side of the street between two intersections. The results of geocoding are therefore fairly accurate. The question is: who would use data sets geocoded this way? For some businesses it is the only available information about location of customers, gathered most likely from customer satisfaction surveys that include a question about postal code. Knowing who the people are who patronize their businesses (using the link between postal codes and census standard geographic areas) and where (what part of the city, town, or county) they are located helps in planning marketing campaigns and targeting new customers. For some research projects, due to data confidentiality, customer data is often aggregated by postal code (in both urban and rural areas), then geocoded and analyzed.

2.3 Geocoding by Area

Geocoding by area matches the regional information of each record with the regional entities with the appropriate attributes in the database, and then it changes records into points to position to the center of the regional entities to achieve encoding. The attribute data of each regional entity have the description information about addresses, such as Xiyuan District BeiShiOu, Caizhen District BeiShiOu. It cannot be covered among the same regional level in the address database, and it must be the seamless

connectivity between the administrative regions. The accuracy resulted in this method depends on the spatial extent of the regional entities in the address database, and the smaller the regional entities describe the higher the accuracy is. It is obviously that the accuracy of the residents of the community is higher than the accuracy of the administrative regions.

3 Geocoding in China

3.1 Status of Geocoding

Because of China's long history, vast territory and variety of local customs, it brings the following difficulties for geocoding:

(1) Poor address named regularity, and complex format. Due to the historical reason, the way of naming address often has humanities color, and it is the human habits and it is without any scientific design, any available authoritative rules and obvious address sign word collection. The same word may identify both a street and a cell, so it means the geocoding workload is heavy and the process is complex.

(2) Chaos of urban planning. Many street numbers are named randomly, and it is serious for additions and deletions, and it depends on manual work to collect large amounts of data.

(3) Poor level. It is difficult to measure the interlocking mesh of administrative areas.

(4) Continuous writing. Chinese word segmentation problem is a major difficulty of geo-coding.

However, a more consummated Chinese geocoding system should have the following characteristics:

(1) Semantic parsing. Understanding the internal address and address string semantic information, such as identifying equivalent redundant part of the address, automatically trying to resolve the most accurate address elements.

(2) Controllable address matching. It is able to use the information resolved the address, and it allows users to specify a different strategies to match the address to meet certain conditions, such as various precision matching.

(3) Certain degree of ambiguity and fault tolerance. With the help of address dictionary, a good address system should be able to check the address specification to improve the system availability and the users' experience.

3.2 Geocoding in China

Geocoding is a positioning technology based on the address, and the usual way to achieve by three methods above. These traditional matching ways is obvious not to meet the need of the current situation in China, and it is the main reason which the use of geocoding is not high. So is necessary to research the geocoding system suitable for China's national conditions.

It is necessary for Chinese-style geo-coding to research and analysis the address situations. Transform the old cities, expand the new cities and the city is ever-changing. Due to these, they make the address complex. Through comprehensive

analysis of the actual address data, it can summary that the address mode is complex hierarchical model. Address is composed of geographical entities, such as administrative division name, road name, district, maker name and house number. Different orders identify different meanings. According to this law, it can establish the models of a building type, a building grade class, a marker class and a slice class.

4 Proplems and Improving Methods

4.1 Problems

When encoding the non-standard address data, geocoding software is often not well executed. If the description of an address is obscure and it can be matched with several street addresses, so the user cannot select the specific location of this address, and it can only output these fuzzy, respectively results. Therefore, it is very important to fully understand the data in the database and how to use these data before geocoding these addresses. Generally consider three issues:

(1) Understand the specific type of the address data stored in the database. Determine whether the range these data express is the level of street address, the town level, or the Post level. And try to eliminate the ambiguity information contained in the address, such as the same name of streets. But it cannot geocode by street address in different geographical areas, it need to add the limit of the space area or range to the address data, otherwise it is difficult to distinguish them and give the only coordinates to them.

(2) Select the geo-referenced type used in the process of the geocoding.

(3) Consider the address data accuracy of a database of geographic. If it tries to locate the cable, the crime scene or fire hydrants, it needs to require a higher accuracy, and it should be at street level address to geocode (the most basic unit [5]).

4.2 Improving Methods

Geocoding can not be successfully applied reasons: incomplete data or the address is ambiguous nature; address data contains some of the characters or the format of address data cannot be processed as normal by geocoding system; address data which meet the requirements as geo-reference street map data is not complete or did not update; address data is correct, but it has changed in the boundary of the area. Usually there are some geocoding methods to improve:

(1) Improve the quality of address data. It is vital to standardize address data and clear the geo-coding elements which do not meet the requirements of geocoding.

(2) Select a better geo-referenced database system (the database contains the street map data, zip code map data, map data and other administrative divisions).

(3) Specially train some people to collect data. The workers should arrange people acquainted with the specific area to collect and deal with data by following standardization of the address, and avoid the fault of the address data which may contain a number of non-standard factors influencing geo-coding and the quality.

(4) Geocoding based on a variety of geo-referenced database, and it can also be combined with older versions of data files, and use geocoding to understand the specific changes of the regions in time and space to analysis.

(5) Set parameters to geocode so that it can accept "non-standard" of the address data.

(6) According to certain geocoding sequence to geocode. Generally execute geocoding which requires a higher accuracy, and then adjust the parameters, and next to execute geocoding which requires a lower accuracy.

5 Conclusion

Geocoding can use the address design as a basis, and it meets the application requirements of all work, and it also establishes the link between them. Therefore, in the same geographical framework it can fuse frameworks and models between spatial information and non-spatial information and between non-spatial information and non-spatial information, and based on this framework and model it can meet the direct, real-time data sharing requirements to achieve different data types, interoperability between different systems. For Digital City's information resources integration, geocoding is a critical technology. The domestic complex address model and the address status lack of uniform standards bring a lot of difficulties and problems to geocode, we must find ways to address these issues to make geo-coding system promoted and applied. This article describes the views of some reference, hoping to provide researchers with a little help.

References

1. Hu, Q., Xu, J., Wang, Z.: Study on the Method of Address Automatically Matching in GIS Database. *Geomatics & Spatial Information Technology*, 50–52 (2008)
2. Jiang, Z., Li, Q.: Research on the Applications of Geocoding. *Geography and Geo-Information Science*, 22–25 (2003)
3. <http://www.geocode.com>
4. Clodoveu, A., Davis, J., Fonseca, F.T.: Assessing the Certainty of Locations Produced by an Address Geocoding System. *Geoinformatica*, 103–129 (2007)
5. Schlosser, J.: Geocoding Issues, Alternatives & Techniques. Schlosser Geographic Systems. NWMMapInfo Users Group (1998)
6. Gou, H., Song, G., Ma, L., Wang, S.: Design and Implementation of Address Geocoding System. *Computer Engineering*, 250–252 (2009)

Study the Method of the Optimum Path Based on GIS Features

Lijing Zhang¹ and Jing Yi²

¹ Information & Network Management Center North China Electric Power University
Baoding, China

² Department of Computer Science North China Electric Power University Baoding, China
lijing166@sina.cn

Abstract. Any kind of mature in a particular field is applied to other areas, and it is necessary to analyze the particularity of the field in order to change some parts of the mature for adapting to its application. Once the shortest path algorithm mainly applied to the Computer and Graph theory field, and there are many research achievements. Optimum shortest path algorithms in GIS mostly study from the theoretical basis and operating approaches of the Computer and Graph theory field, but it often ignores the features of GIS. Therefore, it cannot break through the limits of optimizing the shortest path algorithms. In this paper, the features of the GIS data, the characters of the road network and the needs of users are analyzed, and it introduces several methods of the optimum path based on GIS features.

Keywords: GIS, Optimizing, The Shortest Path Algorithm, Optimal Path.

1 Introduction

Currently the theoretical basis of the path algorithm is Dijkstra of the graph theory field, and it is mainly based on the adjacency matrix representation and traversal algorithm, and the result of operation is the shortest path of a vertex to all other vertices. Programming of this algorithm is simple, and versatile is relatively strong. But in the specific implementation, there are still some problems in the operating efficiency and storage space. Although the majority of researchers around the world research and improve the Dijkstra algorithm, but most methods are still based on graph theory and other areas of computer. However, data GIS processes has its own special structure, and in practical applications (focusing on the urban road network), the road network itself also has its own characteristics, therefore it makes the shortest path of GIS not only consider to make use of the characteristics of data as well add road network information.

1.1 Characteristics of GIS Data

At present, data involved in GIS systems is mainly spatial data, and it is the operation of GIS, and it stands for the mapping between the geographical entities or the phenomenon and the world of the information, and it reflects on the basic information

of natural society and the cultural landscape according to its spatial location in the Earth's surface. Spatial data can be graphics, images, texts, tables, figures and others. It has three basic features:

1) Spatial Location

Spatial data contain spatial coordinate of geographic entities which is also called spatial orientation. It also implies spatial distribution which is the position relationship or the spatial relationship between geographical entities, such as the distance, adjacent, conjoint, include. It means that it should analysis the spatial distribution of spatial data when organizing spatial data.

2) Attributes

Not only have geographic entities the spatial location, but also specific attributes. Therefore, spatial data must include the special properties which describe geographic entities or phenomena, like name, classification, quantity, etc.

3) Temporal Characteristics

Temporal data is the time or times of Geographic Data Acquisition or Geographic phenomena .In other word, Time-varying geographical entities or phenomena corresponds to the timing or the time.

Three characteristics of geospatial data constitute the three elements of geospatial data (Fig.1).

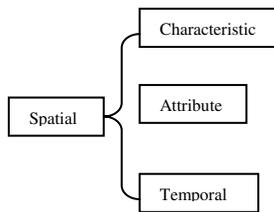


Fig. 1. Spatial Data

1.2 The Characteristics of Road Network

As a special kind of network, road network still has the particularity besides the specific of the general network. It mainly contains the following points:

1) Structure is complex

With the development of the city, urban traffic system becomes more and more complex in the direction of development, and it is more and more common about the traffic characteristics, such as multi-lanes, one-way street, turn restrictions and the overpass system, in addition to more and more complicated new traffic rules, therefore, all of these make the structure of road network becomes very complicated.

2) The characteristics about the grade of road

The purposes of road function classification is to divide road into different levels of times system, and gives each system different mission so as to exert the highest efficiency of the road system. According to the role of the network and the traffic characteristic, road can divided into four grades: expressway, main road, sub-road and branch, and branch road contains level-I branch road and grade level-II branch road.

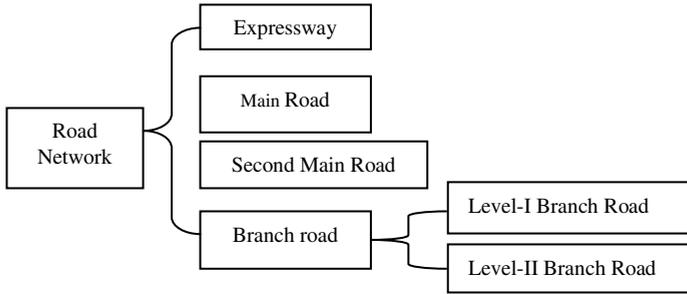


Fig. 2. Road Network

3) Altitude feature.

In order to facilitate to display road, the road is divided into different elevations level road subsets according to the height of the road from the ground.

4) Toll road.

5) The traffic speed is the real-time change

In the actual vehicle process, due to the influences of the traffic management, the traffic flow, the traffic accidents, the weather changes and other factors, the vehicle speed is changing constantly, and it leads that the costs produces in each road is correspondingly changing.

1.3 Needs of the Users

The purpose of the route service is to serve the public, so it needs to meet the different needs of customers. For different users, they have different needs about the paths, and they may have the specific road, not the simple shortest paths. Basically needs of the users have the following points:

- (1) The path with the shortest distance, commonly known as the shortest path.
- (2) The path with the shortest time.
- (3) The most economical path. Just as its name implies, the most economy is that it costs least.
- (4) The path must pass the requirement places..

2 Methods of the Optimal Path Based on GIS Characteristics

(1) Kaiyi Wang and others bring forward their beeline optimizing Dijkstra's shortest path algorithm, and this algorithm makes use of spatial relationship of GIS data, and it also adds the straight-line distance between the current optional and the destination node as reference [1]. The theory of improved algorithm is shown in Fig.3. L_1 is the length from the source node to the temporary marked node P_1 ; L_n is the length from the source node to the temporary marked node P_n , D_1 is the straight-line distance between the temporary marked node P_1 and the as terminal node; D_n is the straight-line distance between the temporary marked node P_n and the as terminal node; the attachment of P_1 and P_n is C . The principle of selecting permanent marked

points in the original Dijkstra algorithm is: if $L1 > Ln$, Pn is selected for the permanent marker node; if $L1 < Ln$, $P1$ is selected as permanent marker node. The principle of selecting permanent marked points in the beeline optimizing Dijkstra's shortest path algorithm is: if $L1 + D1 > Ln + Dn$, Pn is selected for the permanent marker node; if $L1 + D1 < Ln + Dn$, $P1$ is selected for the permanent marker node.

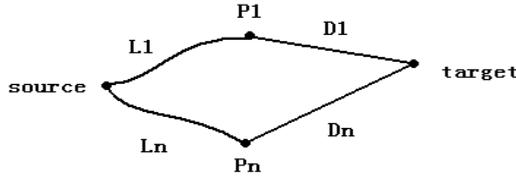


Fig. 3. Beeline Optimizing Dijkstra's Shortest Path

Beeline optimizing Dijkstra's shortest path algorithm is better than the original Dijkstra algorithm Dijkstra, and it obviously reduced the search scope to improve the search speed. In Fig.4, it shows the biggest search scope of the two algorithms. The circle is the search scope of the original Dijkstra algorithm, and the oval is the search scope of the beeline optimizing Dijkstra's shortest path algorithm. The areas of the circle and the oval stand approximately for the number of the traversal nodes of two algorithms, the area of the circle is 4 times than the area of the oval. The scope of the original Dijkstra algorithm is 4 times than the scope of the beeline optimizing Dijkstra's shortest path algorithm.

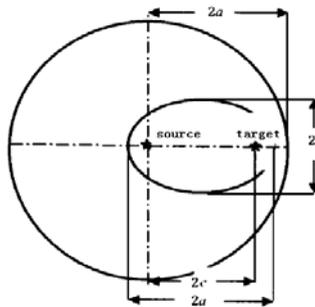


Fig. 4. Both the sizes of the range

(2) Sellen Juergen in Saarlandes and Yan Zhang advance a shortcut search arithmetic basing on Angle of Vector [2]. It makes use of spatial relationship of GIS data, and the first standard of selecting node as the temporary marked node is scalar value of vector angle. Because next adjacency vertex is emerged in the adjacent nodes of the current node, it only compares the adjacent nodes of the current node in each search process. It needs to tectonic two vector parameters: one is that it uses the current node as the starting point and the adjacent node of the current node as the end point, and other that it uses the current node as the starting point and the target as the

end point. The angle between two vectors is the value of the reduction both them. Each adjacent node is corresponding to an angle. It establishes queue and negative queue. The queue contain all the nodes with positive values (including nodes the value of which is zero), and the negative queue contain all the nodes with negative value. The two queues is ordered according to the absolute value of the corresponding node scalar value. It selects the minimum value point from both the queue and the negative queue into vertex set of the shortest path. In other word, it selects the minimum value angle in two directions from both the queue and the negative queue into vertex set of the shortest path. In the process of searching the shortest path, the difference between the original Dijkstra algorithm and this algorithm is modification. Because this improved algorithm has the searching direction from the source to the target in the process of searching and it only considers the neighbor nodes of the current node, the modified time of this algorithm is less than the original Dijkstra algorithm's.

(3) Lingling Hua proposes the shortest path algorithm Dijkstra based on GIS spatial distribution characteristics fully utilizes the GIS spatial distribution characteristics [3]. When the cyberspace is widely distributed, Dijkstra algorithm has low efficiency. In fact, the cyberspace of GIS is in that case. In order to obtain high efficiency, the proposed algorithm mainly finds the shortest path or its approximation as quickly as possible in the early time, so as to avoid searching other branches.

Let's assume that we want to get the shortest path between P-point and Q-point. There are 8-point and 9-point connected to P-point. The angle between vector $IPQI$ and vector $IP9I$ is less than that of vector $IPQI$ and vector $IP8I$, so 9-point is firstly searched. And there are 5-point and 10-point connected to 9-point, simultaneously, 5-point is firstly searched. And there are 4-point,6-point and Q-point connected to 5-point, simultaneously, Q-point is firstly searched. Then we obtain the first search path $P \rightarrow 9 \rightarrow 5 \rightarrow Q$ (actually this is the target path), so it is no need to take on the search from P-point to other paths. The proposed algorithm is efficient and time-saving.

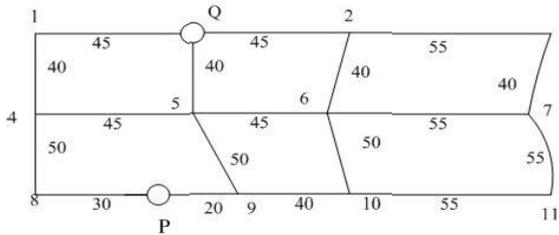


Fig. 5. Optimizing Dijkstra's Shortest Path Algorithm

Based On GIS Special Distribution Feature

(4) LiuSongYe puts forward a method of road optimal path based on ant colony algorithm [4], and its main principle is that when the starting point and the end point are determined, it is the same with the travel tendency of the path. In the other word, once the starting point and the end point are determined, the travel tendency of the path is almost the same to that of these two points. For the sake of high efficiency



and saving search time, according to the information the algorithm higher the concentration of pheromone in the district of possible optimal solution. Therefore, in the search ants lean to choose that area. While the blindness of ants' movement is reduced, the probability of optimal solution is improved.

(5) Yongzhi An discusses the minimum time problem of time-varying network [5]. In the domain of Intelligent Transportation System (IIS) and Vehicle Inducible System of city, traffic system is nonlinear and time-varying, no longer depending on static and invariable information. In the reality of traffic system, the author starts with the statistical distribution of traffic flow and adapts dynamic calculation of travel time on base of time-varying information in order to actually describe the topological feature of road network. The advantage of this algorithm is in the search area it can find a global optimal node sequence in strict accordance within minimum time.

(6) Jinsheng Ren proposes a multi-level search algorithm based on hierarchy of road network [6], which analyses people's travel trait that they trend to choose highest class layer of road and redistribute the layer of road network to increase power weight of high class road that to be firstly chosen. In that way, the search for the shortest path of great significance is accelerated. The core idea of this algorithm is reclassify the layer of road network by the class of the road. For example, the road in the map is only classified into two layers, one is ordinary way, and the other is express way. The ordinary way is connected with the express way at the entrance. In Fig.6, the thick line is the express way, and two thin ones are ordinary ways. To search a path between A-point and B-point, if the road is not classified into layers, all datum of the map from L1 to L2 are needed. If the road is divided into two layers, the datum in the two vertical areas and the express way are needed to solve the problem. In the actual path search it can be solved by three stages. The first stage is from A-point to the entrance of the express way P1-point. The second stage is from P1-point to P2-point. And the third stage is from P2-point to B-point. The datum of the express way is less than that of the all road network, so it is efficient to search path in this way, especially when the distance between A-point and B-point is far, the improvement is in the order of magnitude. This method is called three-stage path search method by Jinsheng Ren, which is based on different road classes. Actually the three-section path search method is not only confined in the three levels, ordinary way and the express way. It can be applied to multi-classes road search. That is the multilevel search thought.

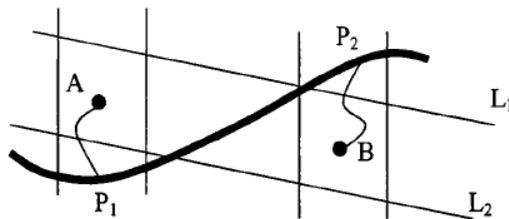


Fig. 6. Multi-level Search Algorithm Based On Hierarchy of Road Network

3 Conclusion

Due to the different applications of the optimal path, it should consider the cost, road congestion and other factors, in the application, except the space distance and direction. Therefore, it should select the appropriate and optimization algorithm to fit the specific situations. This paper describes the characteristics of GIS data and road network, and analyzes the needs of user for road service, and introduces several methods of using GIS data characteristics and road network features. In the real word, it may also take more intelligent factors into the algorithms, and hope this paper can be a view to play a valuable role.

References

1. Wang, K., Zhao, C., Xu, G., Song, X.: A High-efficiency Realization Way of the Shortest Path Search Problem in GIS Field. *Journal of Image and Graphics*, 951–956 (2003)
2. Zhang, Y.: Analysis of Shortcut Base on Angle of Vector. WuHan University (2005)
3. Hua, L.: Dijkstra Shortest Algorithm research based on GIS Special Distribution. College of Computer Science Chongqing University (2007)
4. Liu, S.: The Research of Optimal Path Based on Improved Ant Colony Algorithm. East China Normal University (2009)
5. An, Y.: Research and Implementation of the Optimal Path Algorithm Based on GIS. Nanjing University of Science and Technology (2006)
6. Ren, J.: Multi-level Search Algorithm Based on Hierarehy of Road Network. University of Electronic Science and Technology of China (2008)

Route Search Base on pgRouting

Lijing Zhang¹ and Xuanhui He²

¹ Information & Network Management Center North China Electric Power University,
Baoding, China

² Department of Computer Science North China Electric Power University, Baoding, China
lijing166@sina.cn

Abstract. pgRouting provides many path analysis algorithm. This article absorb pgRouting quick speed, convenient operation features and provide the improved calculating the shortest path method which aims at putting forward the road level and preferred the highway. Geographic Information System is on the basis of database management system (DBMS) analysis and management space object information systems. After 40 years of development, today it has gradually become a door quite mature technology, with its strong geographic information space analysis function. It got extremely extensive application in the path optimization. Path optimization is on the basis of urban road network diagram, planning the best path process. It is the core of the shortest path search. The pgRouting algorithm bag on the basis of PostgreSQL and PostGIS provides path analysis function which including the classics Dijkstra, Shooting A * and Shooting* algorithm. It also provide the traveling salesman problem solution, computing the travel distance and targeted urban problems such as the one-way traffic, traffic lights to in a powerful support. This paper absorbed pgRouting the advantage of quick speed and convenient operation. At the same time introducing highway grade path analysis of priority consideration, it namely always take the highway's situations, this paper proposed the improved algorithm base on the pgRouting Dijkstra algorithm. It makes the calculated route taller with the actual situation needs.

Keywords: Dijkstra, pgRouting, PL/pgSQL, Highway.

1 Introduction

1.1 Introduce of pgRouting

pgRouting is available under the GPLv2 license and is supported by a growing community of individuals, businesses and organizations. pgRouting extends the PostGIS / PostgreSQL geospatial database to provide geospatial routing functionality.

Advantages of the database routing approach are:

Data and attributes can be modified by many clients, like Quantum GIS and uDig through JDBC, ODBC, or directly using PL/pgSQL. The clients can either be PCs or mobile devices.

Data changes can be reflected instantaneously through the routing engine. There is no need for precalculation.

The “cost” parameter can be dynamically calculated through SQL and its value can come from multiple fields or tables.

pgRouting provides functions for:

- Shortest Path Dijkstra: routing algorithm without heuristics;
- Shortest Path A-Star: routing for large datasets (with heuristics);
- Shortest Path Shooting-Star: routing with turn restrictions (with heuristics);
- Traveling Salesperson Problem (TSP);
- Driving Distance Calculation (Isolines).

The corresponding function for dijkstra algorithm is shortest path. This function contains five parameters, such as sql, source_id, target_id, directed and has_reverse_cost. The sql is a SQL query. The source_id is a int4 id of the start point. The target is a int4 id of the end point. The directed is a bool type variable. It is true if the graph is directed. The has_reverse_cost is a Boolean type variable. It is true if the graph has the reverse direct. If it is true, the reverse_cost column of the SQL generated set of rows will be used for the cost of the traversal of the edge in the opposite direction. Function output three fields such as vertex_id, edge_id and cost. The vertex_id is the identifier of source vertex of each edge. There is one more row after the last edge, which contains the vertex identifier of the target path. The edge_id is the identifier of the edge crossed. The cost is the cost associated to the current edge. It is 0 for the row after the last edge. Thus, the path total cost can be computed using a sum of all rows in the cost column.

1.2 PL/pgSQL

PL/pgSQL is a loadable procedural language for the PostgreSQL database system. The design goals of PL/pgSQL were to create a loadable procedural language that

- can be used to create functions and trigger procedures,
- adds control structures to the SQL language,
- can perform complex computations,
- inherits all user-defined types, functions, and operators,
- can be defined to be trusted by the server,
- is easy to use.

SQL is the language PostgreSQL and most other relational databases use as query language. It's portable and easy to learn. But every SQL statement must be executed individually by the database server.

That means that your client application must send each query to the database server, wait for it to be processed, receive and process the results, do some computation, then send further queries to the server. All this incurs interprocess communication and will also incur network overhead if your client is on a different machine than the database server.

With PL/pgSQL you can group a block of computation and a series of queries inside the database server, thus having the power of a procedural language and the ease of use of SQL, but with considerable savings of client/server communication overhead.

- Extra round trips between client and server are eliminated
- Intermediate results that the client does not need do not have to be marshaled or transferred between server and client
- Multiple rounds of query parsing can be avoided

This can result in a considerable performance increase as compared to an application that does not use stored functions. Also, with PL/pgSQL you can use all the data types, operators and functions of SQL.

PL/pgSQL is a block-structured language. The complete text for a function must be a block. A block is defined as:

```
[ <<label>> ]
[ DECLARE
declarations ]
BEGIN
statements
END [ label ];
```

1.3 The Improved Algorithm Based on pgRouting

Algorithm thought describe: The route search is divided into three parts such as the route from the start address to highway entrance, the route from the highway entrance to the highway exit, the route from the highway exit to the end address. Then each part of the route still uses the Dijkstra search algorithm of pgRouting.

This paper presented the road geographic data with the province for the unit. Ordinal reasoning, every geographic data presented all the road of one province.

Computing route algorithm is described below:

Input: route search starting point and purpose points

Output: point id, road section id and road section cost which is gone through by the computing the shortest path.

Begin

Calculate the province Pr1 which the starting point lie in;

Calculate the province Pr2 which the starting point lie in;

If Pr1 is equal to the Pr2

Call route search algorithm 1;

Else

Call route search algorithm 2;

End

Route search algorithm1 are described below:

Begin

Calculate the starting point's corresponding point id in the Pr1 data file;

Calculate the highway entrance's corresponding point id in the Pr1 data file;

Call dijkstra algorithm to calculate the route from the starting point to the highway entrance in the Pr1 data file;

Calculate the highway exit's corresponding point id in the Pr1 data file;

Call dijkstra algorithm to calculate the route from the highway entrance to the highway exit in the Pr1 data file;

Calculate the end point's corresponding point id in the Pr1 data file;

Call dijkstra algorithm to calculate the route from the highway exit to the end point in the Pr1 data file;

End

Route search algorithm2 are described below:

Begin

Calculate the starting point's corresponding point id in the Pr1 data file;

Calculate the highway entrance's corresponding point id in the Pr1 data file;

Call dijkstra algorithm to calculate the route from the starting point to the highway entrance in the Pr1 data file;

Calculate the highway entrance's corresponding point id in the national freeway road data file;

Calculate the highway exit's corresponding point id in the Pr2 data file;

Calculate the highway exit's corresponding point id in the national freeway road data file;

Call dijkstra algorithm to calculate the route from the highway entrance to the highway exit in the national road data file;

Calculate the end point's corresponding point id in the Pr2 data file;

Call dijkstra algorithm to calculate the route from the highway exit to the end point in the Pr2 data file;

End

Due to core of the algorithm is still cited the dijkstra algorithm of pgRouting, It use PL/pgSQL to realize the algorithm. So it can be called by the sql query.

Theorem 1:

For a given starting point and purpose point, computing rote algorithm can calculated the shortest route in limited time.

Prove:

The ideology of the computing rote algorithm is that Route search is divided into three sections, and separately calculated the shortest route. In every part of the route computation process, its core is invoked the dijkstra function.

The dijkstra function can calculate the shortest path in limit time.

It needs to know its corresponding to parameters if it calls the Dijkstra function. First make sure that point id in the corresponding geographic files, this can be known by traversing the scale of each province. Then it can be calculated the corresponding point id in the corresponding data file by traveling the data file. So it can call the dijkstra function to calculate the shortest path. As we all know, either calculating the point id or the file which point lie in can be finished in limit time.

2 Summary

This algorithm can be wrapped by the PL/pgSQL language. PL/pgSQL language is a loadable process language. The algorithm was realized the function by the PL/pgSQL language. It calls the Dijkstra function of pgRouting during the function body. Due to the function was realized by the PL/pgSQL which is an procedural language, it can be called as a sql query. The improving function was operated conveniently.

The core of the improving algorithm was called the Dijkstra function of pgRouting. This algorithm absorbed the pgRouting algorithm's advantages of high efficiency executive and convenient operation.

The execution time of this algorithm is short. Because the pgRouting algorithm bag is the development based on server, it expands many route search function for PostgreSQL. These functions can be run efficiently. Calculating the parameter of Dijkstra function is traveling. It can be finished in short time.

Although the dijkstra function of pgRouting executes high efficiently and operates expediently. But it is only is executed in a geographical file. So it is not realize the limitations of the route involving multiple geographic file research. However the improved algorithm solves the problem well. It linked multiple tables to operate by the PL/pgSQL language. It makes involving multiple geographic file realize the route search possible.

The dijkstra function of the pgRouting has restrictions to the record number of the geographical files. When the record number was too much beyond the function restriction it will unable to run successfully. The new algorithm is a big turn small thoughts. It can make the large files is divided into several small files. Then it is according to the idea of involving geographical files to search route.

The improved algorithm has more practical meaning. Since the route search should be given not only, more should consider from roads and highways always people choice of priority.

References

1. Peng, X.: PostgreSQL object relational database development. Posts and Telecom Press, Beijing (2001)
2. Momjian, B.: PostgreSQL: introduction and concepts. Addison-Wesley (2001)
3. Douglas, K., Douglas, S.: PostgreSQL: a comprehensive guide to building, programming, and administering PostgreSQL databases. Sams Publishing (2003)
4. Stinson, B.: PostgreSQL essential reference. Sams Publishing (2001)
5. Matthew, N., Stones, R.: Beginning databases with PostgreSQL: from novice to professional. Apress (2005)
6. Lisichenko, R.: Using geomedial professional. Thomson Learning (2005)
7. Mitchell, T., Christl, A., Emde, A.: Web-Mapping mit Open Source-GIS-Tools. O'Reilly, Germany (2008)
8. Neteler, M., Mitasova, H.: Open source GIS: a GRASS GIS approach. Springer, Heidelberg (2008)

Research on the Improved Algorithm of Node Split

Lijing Zhang¹, Xuanhui He², and Yulan Ren³

¹ Information & Network Management Center North China Electric Power University, Baoding, China

² Department of Computer Science North China Electric Power University, Baoding, China

³ Science and Education Department Huhhot Power Supply Bureau of Inner Mongolia
lijing166@sina.cn

Abstract. The index dynamic operation includes three operations such as search operation insert operation and delete operation. In the inserting operation it possibly results node overflow. So it needs to use the node splitting algorithm to realize the inserting new node operation. R-tree search performance depends on the cover and overlap two parameters. Overlapping minimize is more critical than forehead minimization. This algorithm put forward in R - tree node apart, according to make the principle of least overlapping. It will divided nodes into two parts. Then it adjusts the position of already insert node and inserts the new node in the new position. The R-tree improved algorithm performances well in the search operation.

Keywords: Division, R - trees, Overlap, Cover,

1 Introduction

1.1 Spatial Index Technology

In GIS field most of the data is belong to the spatial data. The spatial data have traditional data characteristics. It still has characteristics some other data doesn't have:

- **Spatiality:** The spatiality is the main characteristics of spatial data. The spatial data not only describe space objects geographical position, shape and unit of the state in a moment, but also describe the space topological relationships between the space objects.
- **Abstractness:** It realized difficultly and destroy the topological relationship between spatial data easily if deal with the spatial data directly. It must carry on the spatial data abstract processing before we operate the spatial data.
- **Multi-scale and polymorphism:** Through the different scale and different precision the realization of different observing scale. The unified spatial object will have shape difference in the different scale. The form is also not the same.
- **Time-space:** A spatial data available in the same time in multiple of space sequence. The same space object can be in the same space, but different time among the state sequence.
- **Spatial relationship features:** The topological structure data express multiple spatial relationships of spatial relations is recorded by the spatial data. The topological

structure of the data in favor of improving the spatial data query and spatial analysis capability.

In view of the characteristics of spatial data, An ideal space access mechanism should meet the following the most basic request:

- Expandability: Spatial index should be able to adapt to underlying database technology development due to the change.
- Time efficiency: Space access method should be below the inquires the complexity of linear complexity.
- Space efficiency: The space complexity of index should be less than data institute occupies a space.
- The input data and order of irrelevance: Access mechanism should maintain its efficiency no matter how the data distribute and the insert order. When data distributed differently in different dimension this point is particularly important.
- Dynamics: Spatial index should meet the spatial data dynamic. Spatial index can reflect the dynamic changes of spatial data.
- Balance: Spatial index should not only concern the efficiency of a spatial operation, and ignored other operation efficiency. It should not to reduce the complexity of one operation by ignoring the complexity of other operations.
- Minimal impact: Spatial index method integrates with database system to existing systems should have minimum influence to existing system.
- Simplicity: Complex spatial index method often leads to realize mistake. Large-scale applications cannot guarantee sufficient strong.
- Secondary or tertiary memory access seamless: The main memory may not a load whole space database. Spatial index should access secondary index or tertiary index seamlessly.

Space index is stored in a medium of the spatial data on the location information description which is used to provide rapid, selective access spatial data a mechanism, achieve system efficiency for the purpose of acquiring data. Spatial index accords to its way of filtering data set. It will rule out some confounding space information and reduces the range of data which are needed to deal with to improve the efficiency of spatial data into operation. Through the reference to spatial index technology, reduce unnecessary data operation to achieve the purpose of the overall system performance. Spatial index technology is the key technology of spatial database. If it improves the performance of the spatial database it must adopt good spatial index technology. The performance of the spatial index directly influences the quality of spatial database and GIS overall performance. The dynamic operation of the spatial database index includes insert operation, delete operation and query operation.

It may cause the node overflow when inserted the new node into the R-tree. So it needs to split the node to realize the insert operation. This paper puts forward the R - trees index improved algorithm of the node division. Fold in the nodes have under minimum principle, effective control multi-channel inquires, the better the chance to improve the efficiency of the spatial query.

1.2 R-Tree Index

Guttman proposed R - tree which is a highly balance tree. it is a B tree in natural extension of k dimensions. R-tree uses the minimum bounding rectangle to represent object. R-tree has the following several features:

- Each leaf nodes contain m to M index record (including $m \ll m / 2$), unless it is the root node.
- Each index of the leaf nodes records a (I, tuple identifier). It is the minimum bounding rectangle. It contains data object in the k dimensions which is belongs to the tuple in the space.
- Each of the non-leaf nodes has m to M child node, unless it is the root node.
- (I, child node pointer) is an item of a non-leaf node. I is the minimum bounding rectangle which contains its child nodes in the space.
- The root node has two child nodes at least, unless it is a leaf node.
- Each leaf node appears on the same floor.
- The edge of the whole minimum bounding rectangle parallels with a global coordinate axis.

Each node of the tree corresponding to a disk page. A leaf node includes a set of items. It is (I, tuple identifier). It is the minimum bounding rectangle. The tuple identifier is the unique identifiers corresponding to the data which stored as data corresponding to the minimum bounding rectangle in the database. $I=(I_0, \dots, I_{k-1})$ express as a closed interval [a,b] in the direction of I for the I_i . Non-leaf nodes were consisted many item of (I, child node pointer). It is the minimum bounding rectangle of all the rectangle of lower layer nodes which was belonged to the node. As shown in Fig.1 for R - trees schemes.

The evaluation of R - tree algorithm for dynamic operation:

It is a single high balance index tree. The space utilization rate is not high.

It is a fully dynamic index structures, the basic operation can be done simultaneously.

The index directory rectangle overlapped. The searching path is not unique.

Because the R-tree is a single balance tree. The insert operation may cause node splitting into the root in the circumstance of that the corresponding node is already full. It means that the leaf node number has reached the maximum capacity when it finds the proper location to insert the new node. It can't insert the new node. It is a situation of leaf overflow. It needs the split operation. The operation is back to the father node. The father node produces a new child node. Then we can insert the new node into the new child node. But if the capacity of father node is also full, it need to back to the upperstory and run the same operation.

The dynamic operation of R - trees index includes three operations such as insert operation, inquires operation and delete operations. R-tree is a dynamic index structures. Its feature is that it can be performed the operation of spatial data at the same time. For example it can query the spatial data and delete the spatial data at the same time. It will not have the conflict between two operations. It also reflects another feature which is different from other index. Dynamics is very important for the index which the large external storage accesses. It can improve the system response time properly.

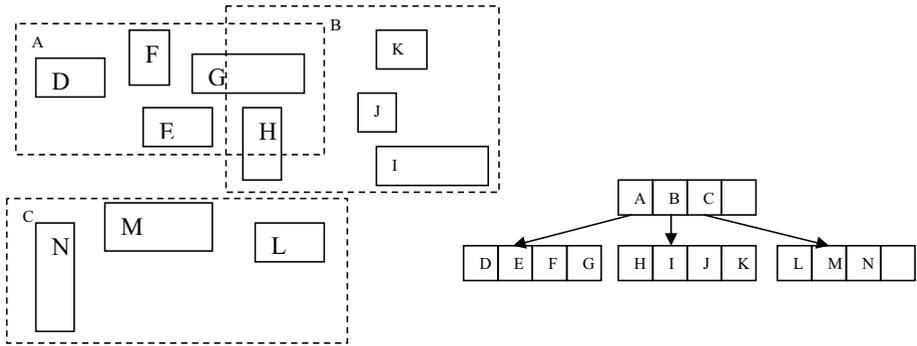


Fig. 1.

1.3 The Improving Algorithm of Spilt Node

R-tree search performance depends on two parameters: cover and overlap. The cover on one layer of R-tree is the whole area which is marked up of the area covers by all the nodes' minimum bounding rectangle. The overlap on one layer of R-tree is the whole area which consist the area overlaps by the multiple rectangle associated to the nodes. Due to the overlap, the query operation must access many node of R-tree. The covering and overlap should be minimum if we want to get an efficient R-tree. Overlapping minimize is more critical than the covered minimization.

The R-tree index asks the object whose distance is recant in the same node. It embodies the spatial data clustering features. Clustering is an efficient way to improve the spatial data query processing properties. It can save the logic related object in the same disk page by using the clustering technology. Using the clustering technology can reduce R - trees constraint rectangular overlap regions between and blank area greatly. But it can't minimize the overlap. So in this paper I put forward a improved R-tree node split algorithm. This algorithm split the M childe node and one new node into two classes according to the principle of least overlap in case has found the proper location to insert the new node. Then the R-tree deletes the nodes which in the class the new node lie in. The father node of the location node produces a new child node. Then it inserted the nodes which in the class the new node lie in into the R-tree as the child of the child node.

Algorithm is described as below:

Input: R - trees and a new node

Output: new R – trees after inserting a new node

Step 1: Find the proper location A according to the principle of the space increment minimization

after insert a new node into the catalogue outsourcing rectangular.

Step 2: Judge the child node number of the location node A. If the number is less than M, it runs the. Otherwise it runs the Step 4.

Step 3: Inserting the new node into the R - trees as the child node of A directly. Algorithm ended.



Step 4: The M child node of A and one new node was consisted a set. Selecting one node s 1 into set S' optionally. Deleting the node s 1 form set S. If there is a node si in the set S merge with the set S' produce the minimum bounding rectangle don't overlap with the minimum bounding rectangle of set S exclude the node si it runs the Step 5. Otherwise it runs the Step 7.

Step 5: Inserting the node si into set S'. It deletes the node si from the set S.

Step 6: Repeating the step 4-5 until there is no node merge with the set S' produce the minimum bounding rectangle doesn't overlap with the minimum bounding rectangle of set S exclude the node. It runs the Step 9.

Step 7: Finding the node sj merges the set S' produce the minimum bounding rectangle overlap minimum with the minimum bounding rectangle of set S exclude the node sj.

Step 8: Inserting the node sj into set S'. It Deletes the node sj from the set S.

Step 9: Deleting all the nodes in the set s' exclude the new node from the R-tree. It produces a new child node P as the node A's child node. Inserting all the nodes in the set S' into the R-tree as the child node of the node P.

Step 10: From the leaf node upward the root node adjusting the catalog rectangle.

This algorithm's principle is reducing the overlap. It can avoid multiple queries effectively and constructed a more efficient R-tree. For example, there is three node in the inserting location node such as S1, S2 and S3. There is a new node S4 need to insert into the R-tree.

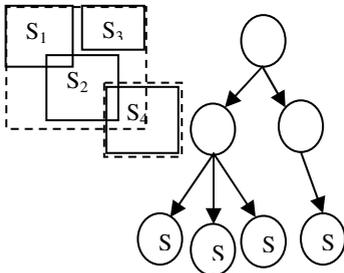


Fig. 2.

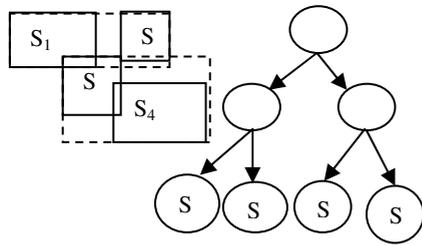


Fig. 3.

From Fig.2 we can see the R-tree node split algorithm is adding the new node as the child node of the proper inserted node location. Then it inserts the new node into R-tree as the child node of the child node. The R-tree node split algorithm through this way to solve the problem of node overflow. But cover and overlap minimize have influenced the R - trees performance. The overlap minimum is more pivotal. From Fig.3 we find that the improved algorithm makes the overlap minimum better to improve the executive performance of R-tree.

2 Summary

This paper analyzes the R-tree node split algorithm. Reference to the factors influence the R-tree performance, this paper gives the improved R-tree node split algorithm. The

improved algorithm with "overlap minimize" for the principle reduce the overlap produce or reduce the overlap to some extent. It improves the R - trees retrieval performance obviously. It control the probability of multi-channel inquires the effectively. It improve the efficiency of the space inquires obviously.

References

1. An, X., Liu, R., Wang, S.: New algorithm about R-tree node split. *Computer Engineering and Applications* 45(3), 158–160 (2009)
2. Huang, J., Bao, G., Xia, B.: Research on spatial database query technology on the basis of R-tree. Central south university, Changsha (2006)
3. Wu, M.: Research on spatial index technology. Jiang su university (2006)
4. Wang, Z.: Research on spatial index technology and the GIS application in electronic power. University of electronic science and technology of China (2009)
5. Zhou, F.: Research on spatial data index technology on the basis of R-tree and application. Harbin university of science and technology (2009)
6. Xie, K., Ma, X., Yang, D.: Spatial database. China machine press, Beijing (2004)
7. Guo, W., Guo, Q., Hu, Z.: Spatial database index technology. Shanghai jiaotong university press (2006)

Modeling Research on Decision-Making Strategies in Human-Computer and Human-Human Ultimatum Game

Dan Xiao, Kerong He, and Jianfeng Hu

Institute of Information Technology, Jiangxi Blue Sky University Nanchang, China
double334_sun@tom.com

Abstract. Game theory provides a solution to the problem of finding a set of optimal decision-making strategies in a group. However, people seldom play such optimal strategies and adjust their strategies based on their experience. In this paper, the game theory research of human-computer and human-human were investigated, and the ultimatum game was used in the experiment. The game theory of human-computer random, human-computer intelligent, human-computer (human) and human-human were investigated in detail. A general learning model was set up to determine whether and how subjects' choices were influenced by their previous selection and gains. The result showed that the subjects tended to choose the same object again after a win trial, and they tended to change their choice after a lose trial. The subjects would have applied different strategies when they faced human and computer respectively.

Keywords: Learning model, Human-human ultimatum game, Human-computer ultimatum game, Game theory.

1 Introduction

At present, monkeys are applied in most experiment in nerve science of game theory, which is less in human to human. Because of the complexity of brain, there are lots of difficulty in game theory between human and human [1-5]. The analysis of decision making in a social group is the topic of game theory [6]. A game is defined by a set of choices or strategies available to each player, and a payoff matrix that specifies the outcome (utility) to each player according to the choices of all players. This is known as Nash equilibrium and defined as a set of strategies from which no players can increase their payoffs by changing their strategies individually [7]. Unfortunately, this important concept has theoretical and practical limitations. In the studies [8-10] of D. Lee et al., they have examined the choices monkeys made during a binary zero-sum game, known as matching pennies. By training monkeys to play such a competitive game against a computer opponent, they showed that the animal's behavior can be modified by the strategies of its opponent. They have analyzed the choice behavior of monkeys during a simple zero-sum game with ternary choices, known as rock-paper-scissors. This was motivated by two considerations. First, rigorous comparative studies of choice behavior in non-human primates can potentially provide important insights into the

evolutionary origins of human decision-making process. Second, such primate models of decision making would also provide important opportunities to understand the neural mechanisms of human decision making.

In this paper, the game theory research of human-computer and human-human were investigated, and the ultimatum game was used in the experiment. The game theory of human-computer random, human-computer intelligent, human-computer (human) and human-human were investigated in detail. A general learning model was set up to determine whether and how subjects' choices were influenced by their previous selection and gains.

2 Materials and Methods

Subjects. Twelve right-handed students (5 females) from Blue Sky University and ranging in age from 18 to 23 years (mean=21.5) participated in the study. All participants were native Chinese speakers with normal vision or normal corrected vision, and with no mental illness. Two participant's data were excluded because he failed to follow the experimental instructions.

Recording. Electroencephalogram (EEG) was recorded using 32-channel NeuroScan system, recordings referenced to electronically link mastoid electrodes. It was recorded using a gain of 1000 and a bandwidth of 0.05 to 100 Hz. Electroencephalogram (EEG) was recorded using 32-channel electrode caps. EEG and EOG activity were recorded using gain of 1000 and a bandwidth of 0.1 to 30 Hz. To eliminate artifacts, trials with EEG voltages exceeding $\pm 70\mu\text{V}$ were rejected from the average. Approximately 10-15% of trials were excluded due to EOG or other artifacts.

Experiments. This investigation adopts ultimatum game, which need to be completed by two subjects, one as the proposer, and another as the respondent. The two participants share 100 points in each trial of the experiment. There are five distribution schemes in the experiment: scheme 1, the proposer gets 20 points, and the respondent gets 80 points; scheme 2, the proposer gets 35 points, and the respondent gets 65 points; scheme 3, the proposers and the respondent get 50 points respectively; scheme 4, the proposer gets 65 points and the respondent gets 35 points; scheme 5, the proposer gets 80 points, and the respondent gets 20 points.

In each trial of experiment, starting with the proposer proposing the distribution scheme, the respondent decided whether to accept its plan, if it is accepted, the parties sharing scores according to the scheme, if the respondent rejects it, neither of them gets scores in this trial. The experimental interface of the proposer and the respondent are shown in figure 1.

There are four blocks in the experiment, for the human as the proposer, the respondent are random computer, intelligent computer, human (in this block, participants would be instructed that the respondent was computer, but actually, the respondent was an another participant) and human, respectively. Therefore, the 4 blocks are simply express: human-computer random, human-computer intelligent, human-computer (human), human-human.

The participants' final gains will be positive relation with their total score, and the ratio of their scores will determine distribution of the gains in the end of experiment.

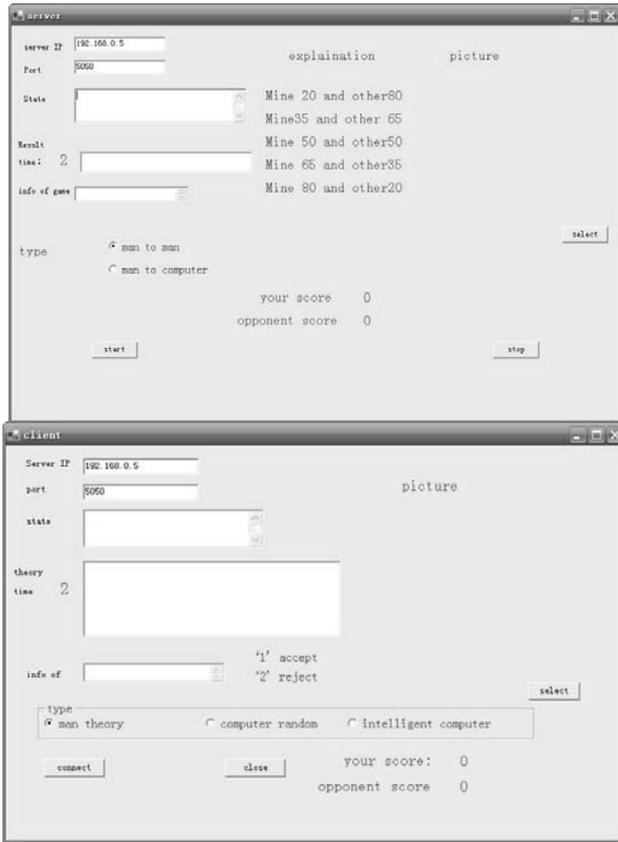


Fig. 1. The experimental interface of proposer and respondent

3 Results

Behavioral results. In the human-computer random block, the proposer was instructed the decision maker is a computer, and the choices of compute are randomly. From the table 1, we can see that, the choice of distribution projects of proposer all concentrated in the third, fourth and fifth project, This is due to the proposer want to gain the high score must choose the distribution projects which in favor of himself (the third, fourth and fifth project were in favor of proposers). Computer as the decision maker, no matter which projects been choose, computers were likely to agree or disagree, the probability of computer to choose agreement is 42.06%, the probability of disagreement is 57.94%, the probability of agreement and disagreement almost.

In the human-computer intelligent block, the proposer was instructed the decision maker was a computer, the choices of compute according to certain standards but not randomly, and the standards need proposer test in the experiment. From the table 1, we can see that, first, the choice of distribution projects of proposer all concentrated in the third, fourth and fifth project which projects were all in favor of proposer (the

Table 1. The results of four blocks

| human-computer random | | | | | | | | | | |
|-----------------------------|-------|----|--------|----|--------|--------|--------|--------|--------|--------|
| projects | 1 | | 2 | | 3 | | 4 | | 5 | |
| proportion | 0 | | 0 | | 20% | | 36.55% | | 43.45% | |
| feedback | 11 | 12 | 21 | 22 | 31 | 32 | 41 | 42 | 51 | 52 |
| proportion | 0 | 0 | 0 | 0 | 10.34% | 9.66% | 12.41% | 24.14% | 19.31% | 24.14% |
| human- computer intelligent | | | | | | | | | | |
| projects | 1 | | 2 | | 3 | | 4 | | 5 | |
| proportion | 0 | | 0 | | 37.41% | | 22.45% | | 40.14% | |
| feedback | 11 | 12 | 21 | 22 | 31 | 32 | 41 | 42 | 51 | 52 |
| proportion | 0 | 0 | 0 | 0 | 37.41% | 0 | 0 | 22.45% | 0 | 40.14% |
| human-computer (human) | | | | | | | | | | |
| projects | 1 | | 2 | | 3 | | 4 | | 5 | |
| proportion | 0 | | 4.81% | | 26.92% | | 25.96% | | 42.31% | |
| feedback | 11 | 12 | 21 | 22 | 31 | 32 | 41 | 42 | 51 | 52 |
| proportion | 0 | 0 | 4.81% | 0 | 7.69% | 19.23% | 2.88% | 23.08% | 0.96% | 41.35% |
| human-human | | | | | | | | | | |
| projects | 1 | | 2 | | 3 | | 4 | | 5 | |
| proportion | 2.4% | | 11.22% | | 27.55% | | 17.35% | | 41.84% | |
| feedback | 11 | 12 | 21 | 22 | 31 | 32 | 41 | 42 | 51 | 52 |
| proportion | 2.04% | 0 | 11.22% | 0 | 4.08% | 23.47% | 0 | 17.35% | 1.05% | 41.84% |

proportion is 100%). Second, the standards of computer are agreement of the third project and disagreement of the fourth and fifth project.

In the human-computer (human) block, the proposer was instructed the decision maker was a computer, but actually the decision maker was a person. Because the proposer and decision maker did not see each other, so the proposer did not know the decision maker was a computer or a person. From the table 1, we can see that, firstly, the choice of distribution projects of proposer almost concentrated in the third, fourth and fifth project which projects were all in favor of proposer (the proportion is 95.29%). Secondly, the decision maker agrees the projects which were in favor of him. It is very interesting in the third distribution project (proposer and decision maker gain 50 respectively); the probability of disagreement is great more than agreement in the choices of decision maker, because the decision maker consider the proposer is computer, and the computer would random choose the project, the decision maker want gain more scores and reduced the scores of proposer, so decision maker did not agree the third project.

In the human-human block, the proposer was instructed the decision maker was a person. From the table 1, we can see that, firstly, the choice of distribution projects of proposer almost concentrated in the third, fourth and fifth project which projects were all in favor of proposer (the proportion is 86.38%). Secondly, the decision maker agrees the projects which were in favor of him.

Learning model[10]. In order to determine whether and how subjects' choices is influenced by their previous selection and gains, a set of mathematical model were set up according to experimental data. There is a variable which is related to subjects' very choice called value function. After each tries, value functions are adjusted according to different model. For example, in reinforcement learning model, value functions are adjusted directly according to subjects' previous choices and gains. In contrast, in belief learning model, value functions are adjusted according to the opponent's choice, and have nothing to do with subjects' previous choices. These two kinds of models represent two different situations, but all of them have defects. A kind of learning model can be set up, in which value functions are adjusted simultaneously according to the choice of both sides, which can be consider as a general learning model.

In this general learning model, the value function $V_t(x)$ for target x (where $x=a, b, c$, represent three different choices), after trial t was update as

$$V_{t+1}(x) = \alpha V_t(x) + \Delta_t(x), \tag{1}$$

where α is a decay rate, and $\Delta_t(x)$ reflects a change in the value function for target x :

$$\Delta_t(x) = \begin{cases} \Delta_{WL} , & \text{the target } x \text{ could have resulted in a lose in a win trial} \\ \Delta_{WW} , & \text{the target } x \text{ chosen by the subject in a win trial} \end{cases} \tag{2}$$

In this general learning model, the change of the value function after a win trial is different from that after a lose trial. The change of the value function after a lose trial applied are expressed as Δ_{LL} , Δ_{LW} .

The probability of subject choosing a given target is given as

$$p_t(x) = \frac{\exp V_t(x)}{\sum_{\mu \in \{a,b,c\}} \exp V_t(\mu)} \tag{3}$$

In order to investigate quantitatively how the proposer's choice was influenced by previous selections of the opponents, a general learning model according to experimental data was simulated. Every proposer had to carry 4 groups of experiments, and the opponents respectively were random computer, intelligent computer, human (which was tricked as the computer) and human. The proposer had five different schemes to choose, the first scheme and the second scheme were adverse to the proposer, the third scheme was fair to both sides, the fourth and fifth schemes were beneficial to the proposer. In the model research, proposer' choices were divided into three classes: being adverse, being fair, and being beneficial. All the model parameters were estimated using maximum likelihood estimation (MLE).

General learning model parameters of subjects A and B was shown in Table 2. In the first group of experiment, the results of the previous trial had little influence on the next trial. In this experimental model, no matter the subject A or the subject B, the decay



coefficient α was relatively big, which resulting small variation of the value function. This demonstrated that the results of previous trial had little influence on the selection of next trial. At the same time, in the next three groups of experiment, decay coefficient α was relatively small, which resulting relatively large variation of the value function. This demonstrated that the results of previous trial had bigger influence on the selection of next trial. Because in the first group of experiment, the responds of the computer were random, the proposer needn't consider the previous results, just to choose the beneficial allocation schemes to maximize the benefits. The general learning model parameters of the second experiment were approximate to that of the third experiment, which signified that the subjects adopted similar strategies. This demonstrated the trick to subjects had being efficient, both of them thought the opponents in the two groups of experiments were the same, and adopted the similar strategy. The parameters of the fourth experiment had obvious differences with that of the previous two groups of experiments, which demonstrated that the subjects would have applied different schemes when they faced human and computer respectively.

The general learning model parameters showed parameters of win objects Δ_{LW} and Δ_{WW} mainly were bigger than parameters of lose object Δ_{LL} and Δ_{WL} , which was accordant with hypothesis of the belief learning model. After a win trial, the win object' parameters Δ_{WW} were all positive, which showed that the subjects tended to choose the same object again after a win trial. After a lose trial, the win object's parameters were positive too, this demonstrated the subjects tended to change their choice after a lose trial. In human-human game, the subject tended to change object after a win trial, choosing the beneficial allocation scheme.

Table 2. General learning model parameters

| subjects | Experimental style | Δ_{LL} | Δ_{LW} | Δ_{WL} | Δ_{WW} | α |
|----------|----------------------------|---------------|---------------|---------------|---------------|----------|
| A | human-computer random | 0.0362 | 0.0592 | 0.0212 | 0.1073 | 0.9685 |
| | human-computer intelligent | -0.1673 | 0.4842 | -0.3786 | 0.6687 | 0.6541 |
| | human-computer (human) | -0.1348 | 0.4374 | -0.3223 | 0.6375 | 0.6249 |
| | human | 0.2641 | 0.3874 | 0.2347 | 0.5248 | 0.7214 |
| B | human-computer random | 0.0495 | 0.7467 | 0.0067 | 0.1347 | 0.9256 |
| | human-computer intelligent | -0.3424 | 0.4610 | -0.4861 | 0.7065 | 0.5741 |
| | human-computer (human) | -0.2741 | 0.4168 | -0.4297 | 0.6241 | 0.5916 |
| | human | 0.1428 | 0.4935 | 0.3874 | 0.7684 | 0.6487 |

4 Conclusion

Game theory provides a solution to the problem of finding a set of optimal decision-making strategies in a group. However, people seldom play such optimal strategies and adjust their strategies based on their experience. In this research, a general learning model was set up to determine whether and how subjects' choices were influenced by their previous selection and gains. The result showed that the subjects tended to choose the same object again after a win trial, and they tended to change their choice after a lose trial. The subjects would have applied different strategies when they faced human and computer respectively.

Acknowledgements. This work was supported by IT Project of Jiangxi Office of Education [GJJ10274] and Natural Sciences Project of Jiangxi BlueSky University [XYKJ08YB01]. The authors are grateful for the anonymous reviewers who made constructive comments.

References

1. Bhattacharya, A.: *Social Choice and Welfare* 22, 391 (2004)
2. Branzei, R., Dimitrov, D., Tijs, S.: *Mathematical Social Sciences* 47, 313 (2004)
3. Tijs, S., Branzei, R., Muto, S., Ishihara, S., Fukuda, E.: *Fuzzy Sets and System* 146, 271 (2004)
4. Holmström, B., Roberts, J.: *Journal of Economic Perspectives* 12, 73 (1998)
5. Cachon, G., Lariviere, M.: *Management Science* 51, 30 (2005)
6. von Neumann, J., Morgenstern, O.: *The Theory of Games and Economic Behavior*. Princeton Univ. Press, Princeton (1944)
7. Nash, J.F.: *Proc. Natl. Acad. Sci.* 36, 48 (1950)
8. Barraclough, D.J., Conroy, M.L., Lee, D.: *Nat. Neurosci.* 7, 404 (2004)
9. Lee, D., Conroy, M.L., McGreevy, B.P., Barraclough, D.J.: *Reinforcement learning and decision making in monkeys during a competitive game. Cognit. Brain Res.* 22, 45 (2004)
10. Lee, D., McGreevy, B.P., Barraclough, D.J.: *Learning and decision making in monkeys during a rock–paper–scissors game. Cognit. Brain Res.* 25, 416 (2005)

Author Index

- Anding, Pan 439
Assi, Rawad Abu 351
- Bai, Yixuan 627
Bao, WenDong 465
Bataa, Otgonbayar 343
Bin, Chen 181
- Cai, Chaoshi 921
Cai, Zhaoquan 771
Chang, Guofeng 887
Chang, Jing 239
Chang, Sheng 247
Chen, Danying 965
Chen, Ming 979
Chen, Ping 581
Chen, Zhigang 979
Chen, Zuyi 875, 881
Cheng, Fei 41
Cheng, Guilin 33, 49
Cheng, Yun 857, 867
Cui, Yongqiang 737
- Dai, Minli 515
Dai, Xia 557
Dai, Xiliang 901
Daruwala, Rashida Firoz 359
Ding, Jian 415, 423
Dong, Wenyong 123
Dou, Chunhong 149
Du, Jian-guo 649
Duan, Lihua 335, 375
- Fang, Liu 881
Fei, Xia 673
- Gan, Jingyi 41
Gonchigsumlaa, Khishigjargal 343
Gu, Junhua 205
Gu, Qiang 205
Guo, Dan 459
Guo, Kai 737
Guo, Kaijun 25, 33, 41, 49
Guo, Qingran 567
Guo, Tao 139
Guo, Wushi 723
- Han, Ray P.S. 321
Hao, Yitong 257
Hao, Zhi-feng 771
He, Deqiang 627
He, Jian-tong 581
He, Kerong 1015
He, Xuanhui 1003, 1009
He, Yueshun 213
He, Yumeng 271
Hsieh, Meng-Yen 761
Hu, Chen 641
Hu, Jianfeng 1015
Hu, Jingyu 737
Hu, Jun 307
Hu, Zhifen 515
Huang, Han 771
Huang, Simin 521
Huang, Xiangqiong 751
Huang, Yayu 699
Huang, Zhiqiu 307
- Ibrahim, Mohammad Noor 115

- Jia, Dan 25
 Jia, PeiYan 91
 Jia, Shiru 17
 Jiang, Yuelong 929
 Jiefeng, Lv 265
 Jing, Feng 439
 Juan, Li 181
- Kang, Ling 445
 Kaoemancee, Yossawadee 115
 Kasbon, Rozana 99, 115
 Kaur, Manprabhjot 359
 Khan, Faizullah 221, 229
 Khan, Surat 221, 229
 Kim, Young-il 343
- Li, Dong 687, 693
 Li, Fan 673
 Li, Guiyang 139
 Li, Guoqing 909
 Li, Hao 809
 Li, Hua 473
 Li, Jianjun 161
 Li, Jianping 921
 Li, Jing 17
 Li, Jun 745
 Li, Junsheng 155
 Li, Longchang 657
 Li, Quan-guo 445
 Li, Weini 205
 Li, Wen 649
 Li, Wenxing 705
 Li, Xiuwen 593
 Li, Yan 489
 Li, Yanru 617
 Li, Yibing 473
 Li, Yingjiang 155
 Li, Yongyan 415, 423
 Li, Yubao 657
 Li, Yuyi 473
 Li, Zhen-fu 581
 Li, Zhibin 239
 Li, Zhiyong 787, 795, 935
 Lian, Liming 705
 Liao, Fei 657
 Lin, Bin 299
 Lin, Hua Yi 761
 Lin, Jinshan 149
 Lin, Ying 551
 Liu, Chanyuan 177
- Liu, Fenghua 25, 33, 41, 49
 Liu, Hao 841
 Liu, Hong 503
 Liu, Jie 321
 Liu, Jing 779
 Liu, Lei 367
 Liu, Lingxia 1, 9
 Liu, Linyuan 307
 Liu, Qingtang 857, 867
 Liu, Qinrang 529
 Liu, Shuxia 277
 Liu, Tao 33
 Liu, TingXiang 465
 Liu, Wenjuan 929
 Liu, Wentao 131
 Liu, Xinhua 841
 Liu, Yongsheng 617
 Liu, Yongzhi 841
 Lu, An 25, 33
 Lu, Bangjun 901
 Lu, Xiao-xu 827, 833
 Luan, Jingdong 451
 Luan, Weili 25, 49
 Luo, Min 213
 Luo, Yuan 109
 Lv, LingLing 91
- Ma, Lan 551
 Mahamad, Saipunidzam 99, 115
 Mao, Jian 943
 Mazlan, Eliza Mazmee 99
 Meng, JianLiang 915
 Meng, Jintao 833, 935
 Miao, Xiao-ping 189
 Mu, RuiHui 687
 Mu, Shengyong 779
- Oyunbileg, Bat-Enkh 343
- Pan, Hong 693
 Pang, Huancheng 473
 Pang, Jingxia 729
 Piao, Linhua 575
- Ren, Tianzhi 473
 Ren, Yulan 1009
 Ruan, Jing 285, 299
- Shaharom, Sofia 99
 Shen, Maoxing 745

- Shi, Binbin 515, 543
 Song, Xiaoping 41
 Su, Xiaochun 627
 Sun, Jianhong 155
 Sun, Jingcui 567
 Sun, Wei 57, 63
 Sun, Xiaoying 239

 Tan, Yang 771
 Tan, Yanhua 399, 407
 Tang, Shuo 809
 Tao, Wang 951
 Tao, Weidong 383, 391
 Tian, Liang 693
 Tong, Heting 321

 Wan, Yuehua 965
 Wang, Bo 189
 Wang, Gang 557
 Wang, Guitang 929
 Wang, Huichuan 33, 49
 Wang, Jiangpeng 123
 Wang, Li-jun 189
 Wang, Liping 635
 Wang, Meng 161
 Wang, Ning 33, 49
 Wang, Rui-hai 189
 Wang, Shi-zhu 327
 Wang, Xiang 787, 795
 Wang, Xiaopeng 849
 Wang, Xiaorui 41
 Wang, Xing 575
 Wang, Xin-li 599
 Wang, Xuejun 713
 Wang, Yu 321
 Wang, Yunsong 521
 Wang, Zheng 965
 Wang, ZhiAn 489
 Watfa, Mohamed K. 351, 359
 Wei, Ou 307
 Wei, Qiong 599
 Wei, Wei 535
 Wei, Yuanwang 751
 Wu, Dan 197
 Wu, Demin 123
 Wu, Di 257
 Wu, Hao 699
 Wu, Keman 593
 Wu, Meng-yun 649
 Wu, QuanYuan 465

 Xiang, Haifei 299
 Xiao, Dan 1015
 Xiao, Jinjian 521
 Xiao, Ying 921
 Xiaoqun, Liu 641
 Xiazhi, Ge 265
 Xie, Yu 109
 Xu, Feng-liang 543
 Xu, Hong-wei 69
 Xu, Jianqin 41
 Xu, Min 737
 Xu, Pengfei 979
 Xu, Xiaochun 451
 Xu, Xiaoyang 483
 Xu, Yi 557
 Xue, Haitao 415, 423

 Yan, Gu 77
 Yan, Jingfeng 909
 Yan, Lin 17
 Yan, Qing-dong 535
 Yan, Xiaodong 809
 Yang, Che-Yu 761
 Yang, Hongjiang 17
 Yang, Lijuan 893
 Yang, Long 943
 Yang, Xiongwen 627
 Yang, Yaluo 699
 Yang, Yi 973
 Yang, Zengqi 627
 Yao, Bin-bin 611
 Yi, Jing 987, 995
 Yi, Xin 723
 Yin, Cheng 779
 Yin, Haiming 751
 Yin, Maopo 779
 Yong, Hua 875
 Yu, Dongping 965
 Yu, Jin 25, 49
 Yu, Jun 833
 Yu, Quangang 575
 Yu, Tianyi 473
 Yu, Yang 915
 Yu, Yao 257
 Yuan, Bo 529
 Yue, Wang 663

 Zhan, Yulong 257
 Zhang, Bin 221, 229

- Zhang, Changsheng 291
 Zhang, Gaoqiang 801
 Zhang, Guoqing 857, 867
 Zhang, Hongli 85
 Zhang, Huiqing 507
 Zhang, Lei 91
 Zhang, Lijing 987, 995, 1003, 1009
 Zhang, Nuowei 33
 Zhang, Peilin 943
 Zhang, Run-tao 959
 Zhang, Shanwen 315
 Zhang, ShuWen 465
 Zhang, Suqi 205
 Zhang, Wei 827
 Zhang, Yaoping 745
 Zhang, Yong 819
 Zhang, Yu 459, 495
 Zhang, ZhiLai 489
 Zhao, Bo 529
 Zhao, Chengling 857, 867
 Zhao, Hongyu 247
 Zhao, PeiZhang 687
 Zhao, Taixiang 875, 881
 Zhao, Weimin 415, 423
 Zheng, Jun 307
 Zheng, Nan 787, 795
 Zhi, Wenxia 801
 Zhong, Liande 593
 Zhong, Xiaohui 155
 Zhong, Zhishui 679
 Zhou, FuNa 91
 Zhou, Jing 169
 Zhou, Ping 327
 Zhou, Zhengguang 315
 Zhu, Huixia 17
 Zhu, Lei 801
 Zhu, Xiaoliang 857, 867
 Zhu, Xiaoyu 33, 41
 Zhu, Xinhua 431
 Zou, Bo 535
 Zou, Huasheng 291