DATA-DRIVEN OPERATIONS MANAGEMENT

FOR MULTICHANNEL CUSTOMER SUPPORT SERVICES

by

Rodrigo Caporali de Andrade

A DISSERTATION

Submitted to the Faculty of the Stevens Institute of Technology
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Rodrigo Caporali de Andrade, Candidate

ADVISORY COMMITTEE

Dr. Somayeh Moazeni, Chairman — Date

Dr. Paul T. Grogan, Co-Chair — Date

Dr. Jose Ramirez-Marquez — Date

Dr. Ricardo A. Collado — Date

Dr. Mo Mansouri — Date

STEVENS INSTITUTE OF TECHNOLOGY
Castle Point on Hudson
Hoboken, NJ 07030
2020

DATA-DRIVEN OPERATIONS MANAGEMENT
FOR MULTICHANNEL CUSTOMER SUPPORT SERVICES

ABSTRACT

Emerging digital technology and the availability of big data collected from multiple channels can create opportunities for businesses to more accurately predict future interactions with their customers and to devise improved operations management strategies. This dissertation focuses on various prediction and operations management problems in multichannel customer support services. The dissertation first develops a system architecture for a generic multichannel customer support center, capturing various phases of customer interactions. This architecture enables the development of structural models to measure service performance and particularly the customer service reliability for further process control and redesign. In addition, the redesign and optimization of process strategies for contact centers require accurate predictions of customer service requests. This research builds on machine learning techniques and datasets on contact transactions from an insurance company to identify features driving different customer interaction patterns. This analysis allows us to formulate various managerial insights for an enhanced service quality. A data-driven approach is developed to predict an individual customer's call arrival pattern, as a key input of the contact centers' operations management. This model leverages the data collected from multiple contact channels and incorporates information related to the past Web contacts of an individual customer to predict the customer's future telephone queries. Finally, the dissertation studies several data-driven operations strategies for the customer support service and assesses the cost and service quality by a detailed simulation model.

Author: Rodrigo Caporali de Andrade

Advisor: Dr. Somayeh Moazeni

Date: July 2020

Department: School of Systems & Enterprises

Degree: Doctor of Philosophy

To my wife Sílvia.

Acknowledgments

My journey as a Ph.D. student approaches to the end and this dissertation embraces the hard work and learning experience, I gained throughout the past four years. This research is the result of countless discussions with my advisors, teachers, colleagues, and friends. It is also the result of feedback from conferences and reviewers from the newspapers to which we publish and share our studies.

First, I would like to thank Dr. Moazeni for offering me the opportunity to pursue my Ph.D. at Stevens, and Dr. Grogan for his constant support, encouragement, and patience.

In addition, I would like to thank the members of my doctoral advisory committee for their invaluable insights to this research. Dr. Ramirez-Marquez, Dr. Collado, and Dr. Mansouri's confidence in my contributions and constructive advice helped enrich this work in the past four years.

My appreciation also goes to Dr. Abel for her incredible encouragement during the years I worked as her TA, and the SSE staff for their extensive assistance since the beginning of my doctoral program.

To my friends and colleagues Pedro, Sergio, and Thiago, my gratitude for sharing their knowledge during our numerous brainstorming and coffee break sessions. I also appreciate the immeasurable support from our group of friends in NJ that became family.

I extend my deepest recognition and thankfulness to my parents for always believing and supporting me in every possible way.

Finally, to my wife, Sílvia, your love and understanding helped me through the good and difficult times. You kept me going on, and this work would not have been possible without you. You earned this degree right along with me.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Organizations around the world can benefit from integrating digital technologies into their operations management strategies to improve operational efficiency, support robust decision-making, and increase agility to respond to market trends, demands, and innovations. These technologies enable generating and collecting business event data, creating business value from that data by redesigning their business processes in various areas such as customer engagement, expanding employee capabilities, cultivating new product and service offerings (Berman, 2012; Hanifa et al., 2014; de la Boutetière et al., 2018). Organizations taking advantage of data-enable technologies and opportunities are not limited to specific domains, but shared across different industries. Examples of the digitization phenomena include the manufacturing industry with the Industry 4.0 (or smart manufacturing) (Wahlster et al., 2013; Lasi et al., 2014; Geissbauer et al., 2016), the military (DoD, 2018), the financial industry (Gabor and Brooks, 2017), and healthcare (Agarwal et al., 2010).

Digital transformation requires that organizations integrate data-driven technologies optimally into their business processes and operations management strategies in order to accurately assess their value and then decide whether to pursue this transformation or not. This has been identified to be challenging in some areas (Berman, 2012; Kurti and Haftor, 2015; Matt et al., 2015; von Leipzig et al., 2017; Geissbauer et al., 2016; Kiel et al., 2017). Identify use cases that create value for the business is key to getting everyone in the organization aligned behind and committed to the transformation journey. To achieve this goal, it is essential to have a deep understanding of the enterprise mission, products, services, operational processes, people, and how these elements interact with each other. A holistic understanding of the business allows the organizations to be more flexible to incorporate new technologies and market trends to remain competitive in the domain of operation.

The digital technologies organizations incorporate to their business processes generates data that allows for an in-depth and real time control from product design to customer support services processes (Zhang et al., 2017; Mehrbod et al., 2018). Data-driven operations management is therefore the adoption of data-driven decision making by operations managers. It is a field of operations management focused on the adoption of digital technologies and the use of data analytics to support decisions of optimal use of resources to improve business performance (Choi et al., 2018; Brynjolfsson and McElheran, 2016). The role of data analytics is to help operations managers make decisions not hypothesis-drive, but rather based on patterns learned from data (Provost and Fawcett, 2013).

In the service industry, customer relationship management (CRM) is an area that can benefit from technology platforms and data analytics. CRM is a broad field with extensive research literature developed in different business segments such as the banking industry (Nagar and Rajan, 2005), insurance industry (Moon and Russell, 2008), and e-commerce (Venkatesh and Agarwal, 2006). It is defined as a "*management approach that seeks to create, develop, and enhance relationships with carefully targeted customers to maximize customer value, corporate profitability, and thus, shareholder value*" Payne and Frow (2004). The availability of big data collected by companies about their customers creates opportunities for businesses to more accurately predict future interactions with their customers.

Contact centers provide firms with the opportunity to collect rich customer interaction data from multiple channels (Erevelles et al., 2016; Moazeni and Andrade, 2018). Analyzing such big datasets enables companies to better forecast customers' needs and to improve their business processes by providing customized services and more efficient operations. Accurate predictive models for customer behavior are essential to design and optimize business processes (Kelleher et al., 2015). In particular, call forecasting is considered as one of the three fundamental challenges in the management of call centers (Aksin et al., 2007). The complexity of understanding customer behavior is further compounded in the management of contact centers with data recorded from multiple channels (Neslin et al., 2006).

Identifying patterns in interactions of the customer with customer service platforms assists organizations to predict future behavior of customers, and consequently better design various service processes. Contact centers, for instance, are generally part of an enterprise's overall CRM and are valuable sources of customer-related information (De Ruyter and Wetzels, 2000). They typically consist of one or more online call centers, but also include other medium of contact, such as e-mail newsletters, postal mail catalogs, website inquiries, online chat, mobile apps, social media and the collection of in-store purchasing information. That being said, contact centers fall under the concept of "multichannel customer management" (MCM) (Neslin et al., 2006), meaning that different channels, through which the customer can use to interact with the company, are offered.

The contact center industry has a great global economic importance accounting for US$310 billion in 2015. The United States maintains the largest global market share with US$9.4 billion in revenues, and nearly 2.6 million employees (Mularoni and Barnett, 2017). The contact/call center industry also plays significant roles in the economy of countries such as Philippines and India (Friginal, 2009). In general, about 60-80% of the overall operating budget of call centers accounts for workforce

(Stolletz, 2003; Brown et al., 2005; Aksin et al., 2007; Khudyakov et al., 2010), which makes it a constant object of research.

The literature on call centers typically focuses on forecasting call arrival rate based on time series models for different purposes, including customer service representatives (CSR) allocation and other operational planning to meet specific performance metrics (Avramidis et al., 2004; Brown et al., 2005; Weinberg et al., 2007; Shen and Huang, 2008a,b; Taylor, 2008; Shen, 2010; Ibrahim and L'Ecuyer, 2013; Robbins, 2016; Veiga, 2016). To reduce the required workforce in call centers and, consequently, the associated operational costs, businesses resort to self-service channels such as Interactive Voice Response (IVR) systems and web-based portals (Jerath et al., 2015). It is thus essential for businesses to ensure these channels can efficiently handle customers' inquiries in the sense that they are successfully addressed. This not only decreases the likelihood of a customer requesting assistance from a customer services representative (CSR) but also reduces the risk of ineffective connection with potential customers and losing them. There are studies that investigated problems directly related to IVR systems (Khudyakov et al., 2010; Suhm and Peterson, 2002; Soujanya and Kumar, 2010; Ndwe and Dlodlo, 2015; Yu et al., 2016), and others that proposed models to understand and predict customer behavior in multichannel customer support services (Jerath et al., 2015; Xue et al., 2007; Kumar and Telang, 2012).

Although numerous studies were carried out on customer support operations management, there is a lack of research considering large amount of data on customer contacts to make predictions at the individual-level and design data-driven operation strategies. Prior research typically does not consider the customers' heterogeneity, apart from basic customer segmentation based on customer's estimated patience or service priority (Yu et al., 2017). A survey conducted by the Boston Consulting

Group in 2018 (Ringel et al., 2018) shows that strong innovators indicate that technology platforms and big data analytics are the two areas of innovation and product development that will have the greatest impacts in their industry in the next three to five years. This research seeks to explore how data analytics can be used to provide better support for strategic decision-making in customer support services.

1.2 Problem Statement

1.2.1 Research Questions

The extensive literature on customer service operations management continuously advances and creates techniques to bring innovative solutions to traditional problems and also to the new challenges arising in the industry. The previous sections reviewed topics related to the increasing use of data for decision making in operational processes in general, and how researchers tackle the main daily problems in customer service management. However, there is a lack of research that contemplates the use of a large volume and detailed customer data to develop predictive models of customers' behavior at the individual level. The predictive models presented in the literature do not consider detailed customer information. Furthermore, the existing literature pays little attention to connect the implementation of data-driven technologies leveraging the collected data to strategic-level decisions. This research seeks to close the knowledge gaps in the field driven by a central question:

How can data analytics improve decision-making processes in customer support services operations management?

To guide this exploration, we pose three secondary questions.

1. What are the forces that influence contact center behavior and how can system reliability measurement achieve consistent desired behavior?

2. How can a feature-based model leverage multichannel transaction level customer support data to predict customer behavior?

3. How can operations managers assess the trade-offs between investing in data-driven technologies that improve customer matching and classification and their operational cost savings?

The research development approach is roughly modeled as a systems engineering process (Forsberg and Mooz, 1991). Figure 1.1 presents a v-shape diagram that summarizes how these research questions connect and address the main question. From a top-level of abstraction, the first question provides a holistic understanding of the problem under inquiry and defines the boundaries of this research. The properties of the customer support service system reliability measurement enable assessing the system behavior as a whole when inferring improvements to the parts. A critical interpretation of the relationship between the elements of the system allows envisioning the enhanced global performance and avoid local solutions.

The second question details technical components of the implementation of data-driven methods in operations processes with a low level of abstraction. This dissertation seeks to address this questions in two parts approaching distinct problems in customer service operations. Part I seeks to identify opportunities to act proactively in an attempt to solve potential problems of those customers who are more likely to call back in the short term. The purpose is to enhance resource planning based on big data collected from customers' interactions in contact centers. Furthermore, Part II focuses on improving the automation of real-time decisions in the customer support service based on data gathered before and during the customer interaction process, at the individual and aggregate levels.

Finally, Question 3 forward validates the infusion of these data-driven technol-

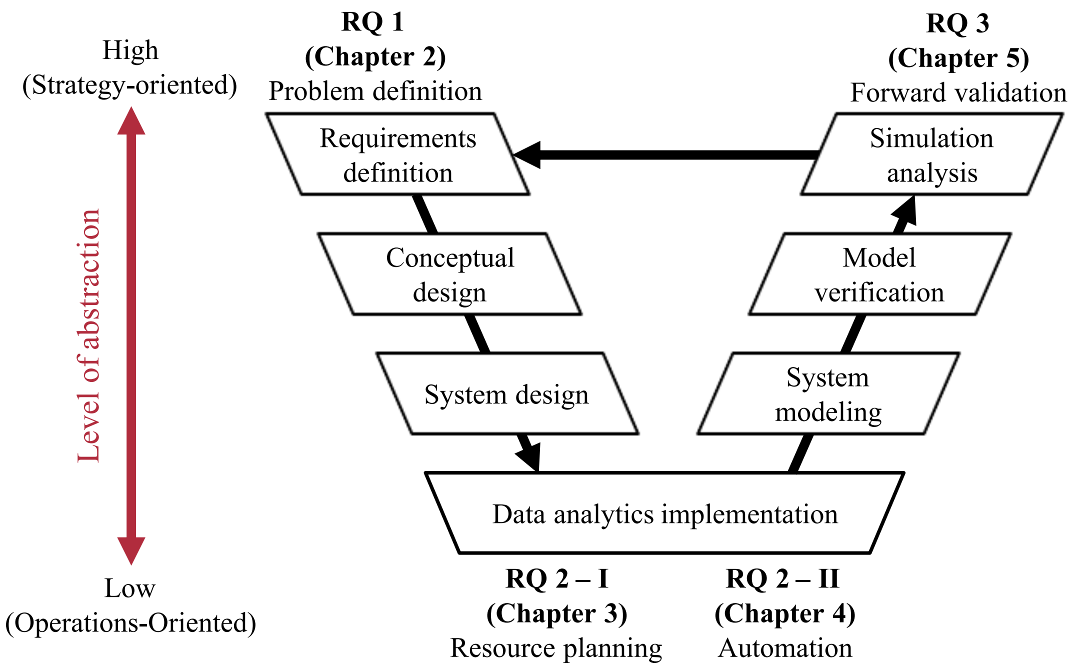


Figure 1.1: Research development model.

ogy components and their influence in the system previously defined in Question 1. This question explores the anticipated effects of these changes to the system from a higher-level perspective.

1.2.2 Research Approach

We address the research questions through the use of a set of quantitative and qualitative methods. Our approach builds upon three pillars: systems thinking, big data analytics, and modeling and simulation (see Figure 1.2).

1. Systems Thinking

   We adopt a systems thinking approach to help decision-makers have a holistic view of the customer support service process. We apply two systems thinking techniques, the systemigram and conceptagon, to provide an in-depth interpretation

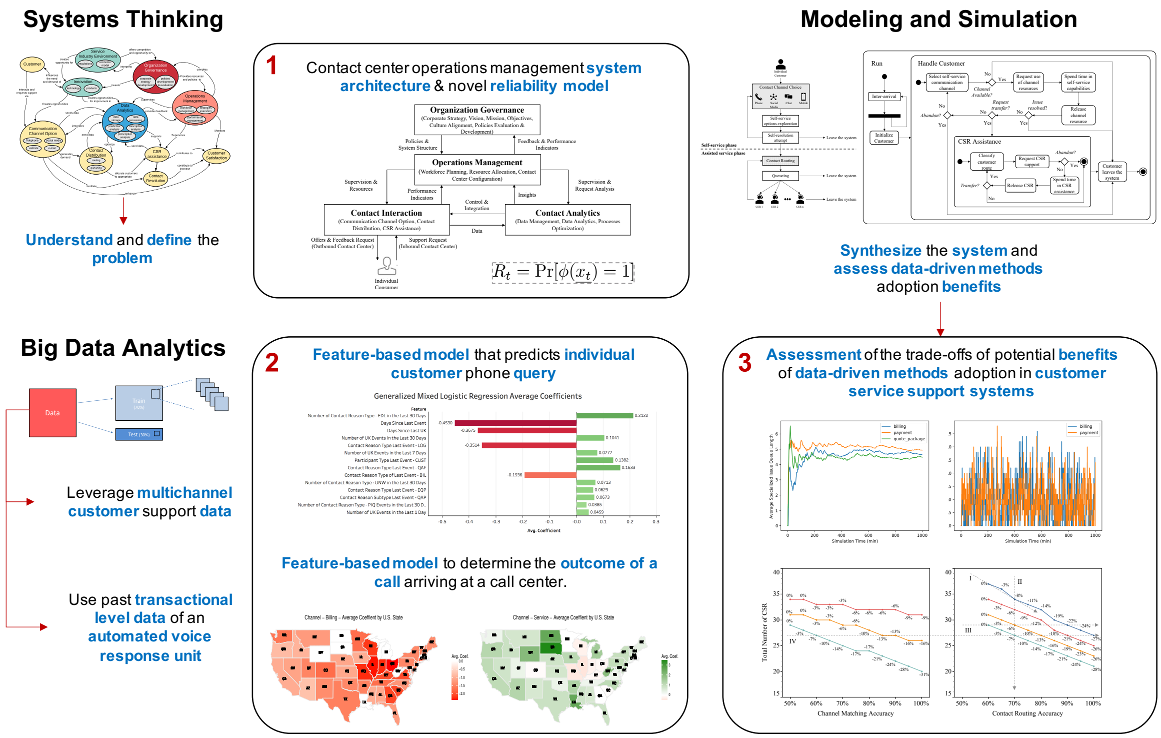


Figure 1.2: Research approaches: (1) systems thinking, (2) big data analytics, and (3) modeling and simulation.

of contact centers as systems, to identify the driving forces influencing their behavior, and to design a contact center system architecture. The reliability model builds on the systems approach that shows that the contact center system reliability is a function of the ability of its components to properly perform their tasks.

2. Big Data Analytics

We leverage multichannel customer support service data to predict customers' behavior and provide managerial insights. The approach includes data exploration to uncover patterns in past Web and Telephone transactional data, feature engineering to aggregate data at individual level and translate insights into predictive model input, feature selection to reduce dimensionality and identify most important factors, and development of machine learning models for prediction.

3. Modeling and Simulation

We take an engineering approach to assess the impacts at the system level of the adoption of new digital technology and data-driven methods on contact center operational processes using simulation models. We investigate potential cost savings from introducing more accurate classification methods to direct customers to more efficient self-service communication channels, and to improve customer-agent matching based on the query type and the agent skill set. Simultaneously, at the higher level, we simulate four customer support system configurations. The study enables a multi-dimensional trade-off analysis between lower-level and system-level design strategies and costs savings from reduced workforce levels.

1.2.3 Research Outcomes and Contributions

The main contributions of this research are:

1. Develop a contact center operations management system architecture that supports a novel mathematical model of the system reliability.

2. Develop feature-based models to leverage multichannel customer support service transaction-level data to predict customers' behavior at the individual level for resource planning and real-time automation.

3. Demonstrate how to assess the impacts, at the system level, of changing the system design when adopting data-driven technologies for contact center operational processes

1.3 Organization of the Dissertation

This dissertation is organized in the following manner. Chapter 2 reviews related literature in data-driven operations management and analyze previous studies on customer support service operations problems. Chapter 3 discusses the characteristics of multichannel customer support systems employing a systems thinking approach to provide a holistic understanding of the customer service process. It introduces a contact center operations management system architecture and proposes a new formulation for the system reliability. Chapter 4 presents a data-driven modeling approach to forecast the likelihood of a call arrival by an individual customer within the next thirty days. Chapter 5 studies the effective features for determining the outcome of a call arriving at the self-service Interactive Voice Response system. Chapter 6 conducts a discrete event simulation to analyze the infusion of data-driven technologies into customer service operational processes and better inform the decision to pursue digital transformation. Finally, Chapter 7 summarizes the findings of this dissertation and discusses avenues for future research.

Chapter 2
Related Literature

This Chapter reviews the literature related to our research in two steps. Section 2.1 discusses the main aspects of data-driven operations management, highlights differences between studies in the manufacturing and service industries, and summarizes the implications of infusing digital technology to operations processes. Finally, section 2.2 explores several problems in customer support service operations management problems including contact forecasting, personnel planning, and contact routing.

2.1 Data-Driven Operations Management

Operations management (OM) involves the design of operation strategies and control of business processes. Operations are the set of transformation activities that create value in the form of goods and services from input resources (e.g., labor, capital, raw materials). Goods are tangible outcomes of operational processes such as appliances and cars, while services are intangible products such as banking, insurance, entertainment, and education. Nevertheless, organizations often produce a mixture of a service and a tangible product (Heizer and Render, 2014).

The service industry is the largest economic sector in the U.S. and comprises over 80% of the labor force (U.S. Bureau of Labor Statistics, 2019). Although the activities involved in the production of goods and services share similar characteristics, there are important differences. In addition to intangibility, two outstanding aspects are customization and consumer contact. The manufacturing industry produces generally standardized consumer goods, which facilitates automated and serial production, and has limited interaction with the consumer during the production process. In contrast, services are generally customized, and present a high level of customer interaction. Customers buy and consume products at the same time (Heizer and Render, 2014).

As a major function of any organization, the operations management involves a wide variety of challenges to guarantee the health of a business, while monitoring the processes performance and promoting improvements. Examples of challenges are rapid product development, mass customization, and productivity (Heizer and Render, 2014). Rapid product development is necessary to adapt to new technology, meet customer needs, and create demands. Organizations should also rapidly respond with product design and flexible production to cater individual needs. Finally, productivity is the ratio between outputs (goods and services) and inputs (labor, capital, management). To improve productivity means to improve efficiency.

Improving efficiency in the service sector is particularly challenging due to its labor-intensity, difficulty to mechanize and automate processes, and to evaluate for quality (Sahay, 2005). Furthermore, measuring productivity in service organizations in terms of service volume is problematic due to lack of service storability (Blois, 1984). The strategic decision is to generate enough capacity to meet demand, which inclines customers to value service quality over quantity (Sahay, 2005). Management, therefore, is key to ensure the effective use of labor and capital. The increase in productivity is the result of improvements in the use of knowledge and the application of technology (Heizer and Render, 2014).

Organizations benefit from the use of digital technologies (e.g., artificial intelligence, big data analytics, social media, mobile devices) that leverages the big data to improve efficiency of operations, support robust decision-making processes, integrate internal institutional processes, and increase agility to response to market trends, demands, and innovations (Davies, 2015; Posada et al., 2015). The evolution in the capacity of data generation and processing, as well as digital technology affordance in recent years, enable the institutions to create business value from that data, and in some cases, redesign their business models to take advantage of uncovered opportunities (Yoo et al., 2012; Opresnik and Taisch, 2015). The continuous and disruptive process of adoption and adaptation to digital technologies by enterprises to enable major business improvements is called *digital transformation* (Roblek et al., 2016; Wahlster et al., 2013; Geissbauer et al., 2016). The infusion of digital technologies in operations management brings additional benefits to the automation of repetitive processes (Oks et al., 2016). It also integrates business technologies, streamlines operations, and provides in-depth and real-time control of business processes from product design to customer support service Hanifa et al. (2014). Simultaneously, companies seek new approaches to engage customers, expand employee capabilities, and cultivate new product and service offerings.

Operations managers can leverage data technologies to support three strategies driving competitive advantage for the business: differentiation (better), cost leadership (cheaper), and response (faster). These strategies involve operations decisions in several areas including product (goods and services) design, quality management, process and capacity design, location strategy, supply chain and inventory management, scheduling, and maintenance (Heizer and Render, 2014). Extensive literature reviews in Schoenherr and Speier-pero (2015); Addo-Tenkorang and Helo (2016); Choi et al. (2018); Nguyen et al. (2018); Wang et al. (2018); Tiwari et al. (2018) describe Big Data Analytics (BDA) techniques and explore applications, methods, and strategies in different areas of operations management. Table 2.1 below organizes and provides a brief description of examples of researches related to distinct strategic decisions areas.

The manufacturing and service industries share several applications of digital

Table 2.1: Examples of studies with Big Data Analytics applications in operations management.

Decision Type	Application	References	Study Description
Design of goods and services	Predictive product lifecycle	Ma et al. (2014)	Introduces a new demand modeling technique, Demand Trend Mining, for Predictive Life Cycle Design
	Product lifecycle	Zhang et al. (2017)	Proposes a framework for Big Data driven Policy Lifecycle Management
	Servitization	Opresnik and Taisch (2015)	Proposes a concept of a Big Data Strategy framework in servitization.
	Product development	Jin et al. (2016)	Identifies product features and predicts trends using customer opinion data from websites
		Tao et al. (2018)	Develops a method for digital twin-driven product design to converge physical and virtual products.
	Product recall prediction	Mukherjee and Sinha (2018)	Investigates judgment bias in product recall decisions based on user-generated adverse event reports.
Production management	Quality management	Wang and Zhang (2016)	Uses BDA to predict cycle time, a KPI of production efficiency of a semi-conductor wafer fabrication system.
	Process and capacity design	Zhong et al. (2015)	Develops big data analytics to create an RFID-enabled intelligent shop floor environment.
		Katchasuwanmanee et al. (2016)	Uses BDA to build a smart system to improve production efficiency and reduce carbon emission.
	Maintenance	Zhang et al. (2019)	Classify the industrial applications based on machine learning models.
Supply-chain management	Location strategy	Wang et al. (2018)	Use BDA to identify the optimal distribution centers network system.
		Zheng et al. (2020)	Use BDA to investigate logistics distribution of major e-commerce enterprise.
	Inventory control	Delen et al. (2011)	Applies operations research, data mining and geographic information-systems-based analytics to blood supply chain management.
		Stefanovic (2015)	Develops a business intelligence semantic model to predict out of stock inventory.
	Demand forecasting	Huang and Van Mieghem (2014)	Investigates the use of clickstream data for operational forecasting.
		Liu et al. (2016)	Use online platforms unstructured data to predict consumers behavior
		Bertsimas et al. (2016)	Solves an optimization problem for inventory prescription based on inventory data sets from a network of retailers.
		Chong et al. (2017)	Investigates online promotional marketing and reviews as predictors of consumer product demands
		Sagaert et al. (2018)	Develops a forecasting framework to predict sales from a massive set macroeconomic indicators

technologies in the management of operations, but focus on different value creation. Both domains use Big Data for predictive and prescriptive analysis and in-depth and real-time operational processes control (Chen et al., 2016). However, the value the manufacturing industry creates follows a traditional concept of unit price minus cost, while the service industry generates value as customer satisfaction (Rouse, 2005a). Therefore, digital technologies often contribute to transform production processes in manufacturing, while in the service domain, it supports changes to customer-centric delivery processes.

For instance, the manufacturing industry technology evolved the processes of information automation and data exchange, creating a notion of "smart factory", key feature of Industry 4.0 (Wahlster et al., 2013). The concept of smart factory consists of the combination of cyber and physical technology (Cyber-Physical Systems) in the factory to enable the integration of previously independent systems resulting in a production line with more complex and precise automation (Hanifa et al., 2014). See Roblek et al. (2016); Wahlster et al. (2013); Geissbauer et al. (2016); Davies (2015); Posada et al. (2015); Lee et al. (2014); Baines et al. (2009) for further reading in Industry 4.0.

In contrast, digital technologies enables innovation in customer services by integrating multiple communication channels and leveraging customer related data to provide a consistent, individualized, and seamless customer service (Lehrer et al., 2018). Companies in multiple industrial categories (e.g., retail, financial services, food and beverage, and technology) employ machine learning, social media analytics, mobile analytics, and artificial intelligence to generate insights on customer behavior, enhance customer experience, improve revenue management, identify optimal pricing policies, and develop marketing strategies (Choi et al., 2018). In the past decade, big data partially drove the emergence of on-demand ride-hailing platforms that significantly changed the transportation industry (Cohen, 2018).

In addition to studies that focus on revealing and developing potential applications of Big Data Analytics and digital technology in the industry , researchers investigate the characteristics of digital transformation processes and draw the implications for operations management (Verhoef et al., 2019; Gölzer and Fritzsche, 2017). A comprehensive review of the industry's digitization process in Gölzer and Fritzsche (2017) discusses the consequences for operations management in terms of organizational changes from a data processing perspective. The result of the extensive literature review is a compilation of 27 data processing requirements for Industry 4.0. Table 2.2 describes the requirements classified into seven main categories, as introduced in Gölzer and Fritzsche (2017).

Table 2.2: Data processing requirements for data transformation of the industry Gölzer and Fritzsche (2017).

Requirements Group	**Category**	**Description**
Data	Data Model	Shows requirements for characteristics of data, structure and sources to integrate in the context of the digital transformation of industry.
	Data Integration	Refers to different perspectives of data integration within an enterprise and beyond.
	Data Content	Shows requirements for necessary data to be processed
Processing	Decision Processing	Refers to requirements for autonomous, de-centralised self-control and self-optimisation performed in CPS networks.
	Knowledge Processing	Refers to requirements for processing of actual and past data to generate knowledge and additional value for decision-making.
	Real-time Processing	Focuses on requirements for processing real-time operations performance.
	Safety and Protection	Contains requirements for IT-security within overarching value networks.

The data processing categories summarize common challenges faced to insert Big Data into operations in different sectors of the industry. The generation of business opportunities with new digital technologies contributes to increasing the importance of information technology and their influence on the organizational structure of industrial activities. Therefore, companies should adapt existing operations management concepts to embrace data processing opportunities (Gölzer and Fritzsche, 2017).

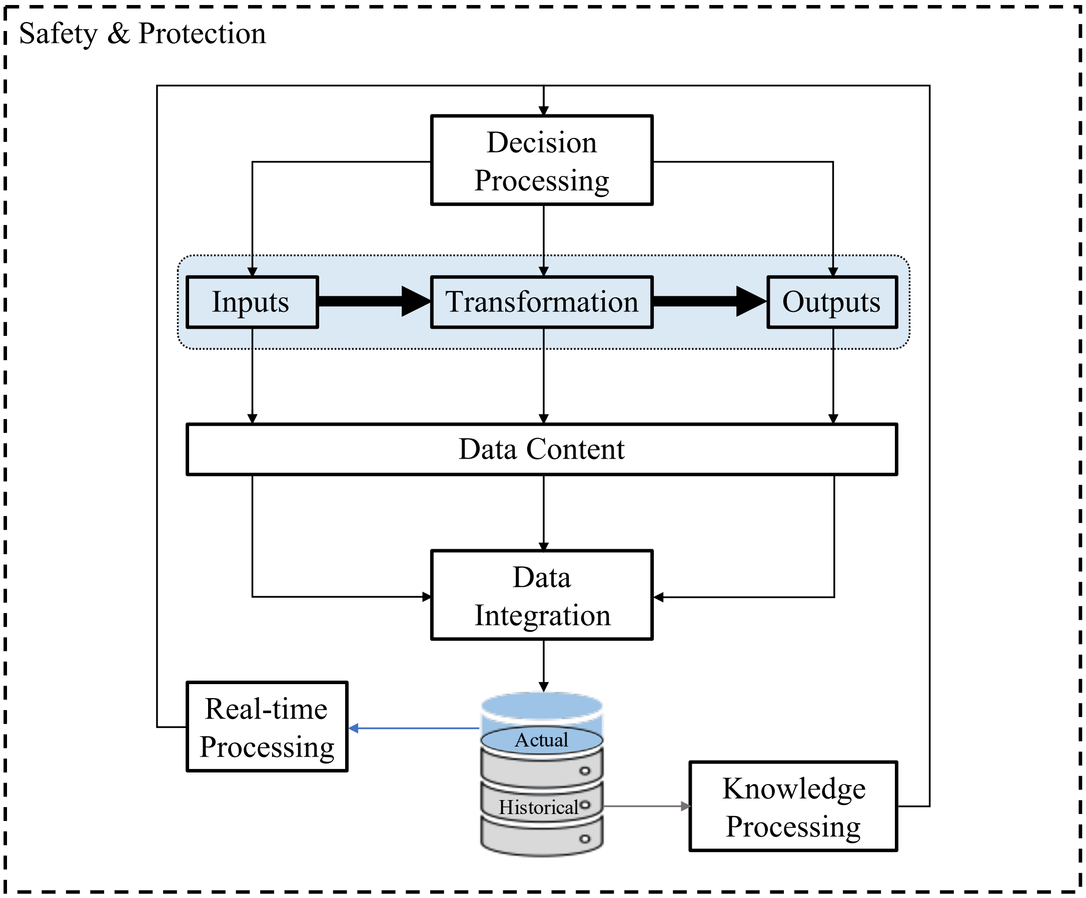


Figure 2.1: Data-driven decision-making in operations. The data processing requirements of the Industry 4.0 augments the economic system that creates value by transforming inputs to outputs (highlighted rectangle) Gölzer and Fritzsche (2017).

Figure 2.1 adapts the diagram introduced in Gölzer and Fritzsche (2017) to incorporate the data processing requirements to the dynamics of the data-driven decision-making process in operations. In Industry 4.0, data from all stages of the economic system that creates value by transforming inputs into outputs generates information that heavily supports the decision-making cycle. The first requirement is to have a data model that allows the exchange of information between a heterogeneous group of CPS generating data with different characteristics. For example, quality data and specification of inputs and outputs have different properties than data collected by sensors in machinery during the transformation process. The data set heterogeneity requires an integration capacity of the information along the value chain (Tao et al., 2011). The integrated data support the decision process in two ways. Information processed in real-time allows monitoring of operations in real-time, enabling ad-hoc reactions (Lee et al., 2014). Furthermore, data analytics insights gained from historical data enable models and rules development to optimize processes aiming to satisfy overall system goals from lower-level system decisions. Examples of lower-level system decisions include self-awareness and self-maintenance of individual machines (Lee et al., 2014). Finally, communication and data exchange networks require protection technology in terms of physical security and availability of IT systems, software security, data security, and reliability (Holtewert et al., 2013). According to Gölzer and Fritzsche (2017), there are four main aspects of the implications of digital transformation for operations management:

- *Adapted decision processes*: decisions in operations management need to accommodate real-time insights gained from operational processes. Allows for rapid reactions to events in the shop floor and also based on models that leverage historical data. Close activity monitoring enables decentralized decisions with assessment of overall systems performance values.

- *Extended repertoire of data*: in addition to quality, material, and product data, traditionally collected in operations management, and with the advance of digital technology and sensorial data, new types of data with different characteristics, structures, and frequencies are generated. The new data gain importance by providing a in-depth view of value creation in the production process.

- *Expanded data management*: extended repertoire of data requires expanded data management to coordinate updates, assure consistency, and integrate information.

- *Big Data treatment*: to cope with high volume, velocity, and variety of data two approaches are necessary: *knowledge processing* and *entity access*. The first one consists of data analytics, mining, and prognosis. The second one comprehends real-time data access and operative decisions.

2.2 Customer Support Service Operations Management

Extensive research literature has been developed in customer service management and for distinct industries segments. For background information and basic concepts see Payne and Frow (2004); Parvatiyar and Sheth (2001); Kumar (2010); Boulding et al. (2005). Aksin et al. (2007) and Neslin et al. (2006) present a collection of bibliography and discussions about challenges and opportunities in call centers and multichannel customer management. While Neslin et al. (2006) organizes the challenges into five categories (data integration, understanding customer behavior, channel evaluation, allocation of resources across channels, coordination of channel strategies), Aksin et al. (2007) describes call centers problems related to call forecasting, personnel planning (resource acquisition, staffing, scheduling, and routing), performance. In addition to literature surveys papers, numerous studies conducted empirical focus on understanding and predicting customer behavior in multichannel customer support systems (Kumar and Telang, 2012; Jerath et al., 2015).

Typical operational problems in customer support systems formulate and solve as optimization problems related to contact forecasting, contact routing processes, personnel planning, and customer behavior in multiple contact channels. In this section, we discuss studies exploring these topics. Extensive reviews on the literature on customer service call centers are provided in Aksin et al. (2007); Mandelbaum (2004); Gans et al. (2003). A comprehensive description of call center operational data and analytical methods can be found in Shen (2010). For an overview call center optimization and description of the common problem managers face in call center operations refer to Koole (2013). The author introduces performance measures, addresses the goals and nature of forecasting demand models and workforce optimization, and discusses the aspects of multi-skill and multi-channel environments.

2.2.1 Contact Forecasting

Demand Estimation

Numerous studies have examined the call center staff scheduling and allocation problems by developing forecast models for call arrival rate. The majority of prior research proposed the use of time series models (Avramidis et al., 2004; Weinberg et al., 2007; Shen and Huang, 2008a; Taylor, 2008, 2012), queuing theory (Brown et al., 2005; Robbins, 2016), stochastic models, (Tezcan and Behzad, 2012), and linear mixed-effect models (Aldor-Noiman et al., 2009; Ibrahim and L'Ecuyer, 2013).

The methodologies for determining call arrival rates in call centers have evolved over the years. Avramidis et al. (2004) builds a doubly stochastic Poisson model permitting the estimation of the expected value of performance measures for a given daily staffing. The experimental model accounts for overdispersion, time-varying rate, and intraday dependence. Results provide evidence that within-the-day planning could benefit from short-term predictions. Brown et al. (2005) suggests that using Erlang-A queuing model for staff planning instead of Erlang-C, that does not consider customer abandonment, could improve operational performance.

Weinberg et al. (2007) proposes a two-way multiplicative Bayesian model for call arrivals using time-inhomogeneous Poisson process with a state dependent arrival rate. In this model the arrival rate depends on the day-of-week and time-of-day states. Soyer and Tarimcilar (2008) uses a similar model to assess the effectiveness of advertising strategies measured by the increase in call volumes. The authors use a proportional rate model to formulate underlying arrival rate functions as the product of a baseline rate function and an advertisement effect consisting of exogenous variables: media dollars, cost for the advertisement, print media type (weekly or monthly) and the offer type.

Shen and Huang (2008a) and Shen and Huang (2008b) integrate exponential smoothing and Autoregressive Integrated Moving Average (ARIMA) models with dimension reduction (using Singular Value Decomposition) to the square-rooted intraday call volume profiles, treated as a high dimensional vector time series. These studies show how intraday updating can improve the accuracy of call center staffing. The model does not considers independent features, and includes only model parameters and times series of arrivals count on specific time intervals.

Taylor (2008) compares the intraday call arrivals forecasting performance of five univariate time series methods including seasonal ARIMA, dynamic harmonic regression, periodic autoregressive model, exponential smoothing model with two seasonal cycles, and robust exponential smoothing model. The author concludes that exponential smoothing for double seasonality and seasonal ARIMA performs well for forecast periods of up to three days, while “simplistic historical average methods is difficult to beat” for longer horizons. Taylor (2012) expands the previous research by building a Holt–Winters exponential smoothing Poisson models (HWT) with gamma distributed stochastic arrival rates. Taylor (2012) compares the forecast performances of methods using HWT with stationary levels and ARMA models, and shows the prediction improved beyond two days ahead. Although the authors does not consider covariates, they acknowledge the research potential to include other variables such as those related to marketing.

In contrast to previous studies, that rely on one-day-ahead or intraday forecasts, Aldor-Noiman et al. (2009) introduces an arrival rate model based on a mixed Poisson process using a Gaussian linear mixed model to provide weekly forecasts. The model incorporates 14 binary features to the model: 6 variables for days of the week, and 8 variables related to billing cycle (4 delivery period indicators, and 4 billing period indicators). Day-of-week and billing cycle information are inputed as fixed-effects, while day and period effects (such as half-hour buckets) are treated as random effects. Ibrahim and L'Ecuyer (2013) continues the use of models with more than one variable proposing a bivariate linear mixed-effects model for the joint distribution of arrival counts to two separate queues for different call types. The study shows the importance of accounting for different correlation structures in the data since it led to more accurate forecasts than univariate models.

In summary, although research has illuminated accurate methodology to determine call arrival rates in different circumstances, for example intraday or weekly forecasts, they have almost exclusively focused on the number of people calling and no study has examined detailed characteristics of the customers using the IVR system or other mediums of contact to make predictions based on the customers profile and history of communications. Whereas there are studies that investigated call centers with IVR systems such as Khudyakov et al. (2010), that focused on the design of a call center with an IVR system by proposing a Markovian model, and Soujanya and Kumar (2010), which described the implementation of Personalized IVR system in Contact Centers, these studies serve for different purposes than the present research.

Additional references on IVR systems related researches in Suhm and Peterson (2002); Ndwe and Dlodlo (2015); Yu et al. (2016).

Existing literature in contact centers explore different businesses: banks (Brown et al., 2005; Weinberg et al., 2007; Shen and Huang, 2008a,b; Taylor, 2008; Moro et al., 2014), telecommunications companies (Avramidis et al., 2004; Aldor-Noiman et al., 2009; Ibrahim and L'Ecuyer, 2013), and the national services (Taylor, 2012). On the other hand, authors are usually limited to using or describing attributes contained in the data that are related to the day and time of the calls (or their volume) and, at most, assess possible differences in CSRs skills, which is an important factor in solving the problem of staff allocation. Thus, the characteristics of those consumers who call have been barely explored. Data containing details about customers, as the data gathered for this study, have not been widely explored in the literature.

The "data-rich environment" (Gans et al., 2010) of call centers can be useful for understanding customer behavior and identifying possible limitations of the self-service platform that fails to provide efficient capabilities for customer support. Although different studies have been conducted in other business segments to analyze and predict the behavior of customers, potential customers, or other target population (Moro et al., 2014; Ryzhov et al., 2015; van der Aa et al., 2015), this problem is still insufficiently explored in the call center context. Moro et al. (2014) investigates data of a Portuguese retail bank outbound call center and modeled the success of a telemarketing campaign, whereas Ryzhov et al. (2015) proposes an empirical model to analyze the effectiveness of several design approaches for direct e-mail used to cultivate donors of the Red Cross.

Table 2.3: Related work on contact arrival rate prediction.

References	Data Source	Model	Feature Types
Brown et al. (2005)	Call Center of Israel's Bank. Jan-Dec 1999	Doubly Stochastic Poisson Process	Numerical arrival counts for time intervals, and model parameters
Avramidis et al. (2004)	Bell Canada Call Center. Less than a year of data.	Doubly Stochastic Poisson Process Model Estimation via ML	
Weinberg et al. (2007)	North American commercial bank. Mar-Oct 2003.	Doubly Stochastic Poisson Process. Model Estimation via MCMC	
Shen and Huang (2008a)	Northestern U.S. bank call center. Jan-Oct 2003	Doubly Stochastic Poisson Process. Expo-nentia smoothing ARIMA.	
Taylor (2008)	Retail bank in the UK. Jan-Sep 2004.	Seasonal ARIMA; Periodic Autoregressive HWT with two seasonal cycles; Robust Exponential; Dynamic Harmonic Regression	
Ibrahim and L'Ecuyer (2013)	Canadian Telecom. company. Oct. 2009 to Nov. 2010	Doubly Stochastic Poisson Process. To evaluate accuracy: Fixed-effect, mixed-effect, two bivariate mixed-effects	
Channouf et al. (2007)	Calgary, Canada, EMS system. Jan. 2000 to Mar. 2004	Five autoregressive models eliminating trend, seaso-nality, and special-day effects. Two Doubly-seasonal models with special-day effects.	
Taylor (2012)	NHS and credit card company in the UK.	HWT and ARMA.	
Ye et al. (2019)	U.S. Telecom company Feb-Aug 2011 Israli Telecom company Jun. 2004 to Apr. 2005	Hidden Markov Models - Single-stream model; Multi-stream model	
Barrow and Kourentzes (2018)	Europe entertainment company. Jun. 2012 to Jun. 2014	Artificial Neural Network	Numerical and Binary
Shen and Huang (2008b)	Northeastern U.S. financial firm. Jan-Oct 2003	Combine Support Vector Machine and AR(1).	Numerical. Arrival count profiles as dimen-sional vector time series.
Soyer and Tarimcilar (2008)	Consumer electronics producer.	Two-way multiplicative Bayesian model.	5 Numerical and Binary features.
Aldor-Noiman et al. (2009)	Israel Telecom company. Feb to Dec 2004.	Mixed Poisson Process. Generalized linear . mixed model mixed model	14 Binary feat.: 6 for weekdays, 8 for billing . cycle.
Moro et al. (2014)	Portuguese retail bank. May 2008 to Jun 2013.	Logistic Regression Decision Trees Artificial Neural Network Support Vector Machine	150 Binary features

Customer Behavior Prediction

Customer behavior analysis and prediction also benefit from data analytics. Customer transactions including products and services purchase, social media content, search history, product usage, and communication channel interactions generate large amount of detailed data companies can use understand individual customer preferences, predict when and what product the customer is more likely to buy, or the choice of channel for the next inquiry.

Jerath et al. (2015) and Kumar and Telang (2012) investigate how customers choose the channel to contact companies. Jerath et al. (2015) develops a stochastic model at the individual-level to predict customer query frequency and choice of the medium of contact to resolve future queries. Although the model provides accurate predictions of customer behavior, it does not contemplate customer-level attributes, which can further enrich the model. Kumar and Telang (2012) analyzes factors that influence the customer's use of different channels in a multichannel service delivery system using a linear regression model. The author concluded that Web-based self-service usage leads to a 14% increase in telephone calls. In a more recent study, Jerath et al. (2015) proposes a novel approach, based on a stochastic function for information stock model, to predict which channel of contact the consumer will opt to use to solve the next query.

Customer interactions strategies in online and offline channels is a discussion topic in the literature (Gallino and Moreno, 2014; Thirumalai and Sinha, 2013). Customer service interactions customization is explored in Thirumalai and Sinha (2013). The authors investigate the implications of the decision to pursue online personalization strategies by retailers. The study applies regression models and statistical analysis on data collected from 422 retailers and provide evidences that companies improve customer loyalty measures from adopting transaction personalization strategies. In Gallino and Moreno (2014), researchers analyse the complex integration of online and offline strategies. One of main managerial implications of the study recommend practitioners take a holistic perspective when evaluating a multichannel strategy. Although the study indicates companies can benefit from implementing "buy-online, pick-up-in-store (BOPS), regression analysis results indicate that implementing BOPS alternatives may decrease online sales. However, the strategy increases store visits and store sales. In conclusion, the consequences of multichannel strategies are also multichannel and evaluating the impacts in one channel isolated can mislead conclusions.

An additional approach to leverage customer data to tailor customer services consists of using customer behavior predictions to identify triggers and plan proactive service contacts in response to individual customer needs. Besides customizing the services, companies can plan to act proactively to shift demand from peak hours to non-peak hours, thus improve service quality (Jerath et al., 2015). Delana et al. (2020) explores the operational impacts of proactive strategies. The study shows how proactive approaches can improve overall system performance by reducing expected waiting times. However, these strategies are not adopted from an economic perspective as customers tend to over join the system due to incentives offered to proactive services adoption. As a result, the so-called "free-ride customer" that would not need service otherwise are also encouraged to join.

2.2.2 Personnel Planning

Managing the workforce is key to secure high levels of service quality in contact centers. It involves, for example, customer service representatives training and proper staffing scheduling. Invest in training is important to provide effective and efficient service to customers. However, about 60-80% of the overall operating budget of contact centers accounts for workforce (Stolletz, 2003; Brown et al., 2005; Aksin et al., 2007; Khudyakov et al., 2010), involving millions in resources (Shen and Huang, 2005). In this context, the optimization of work schedules that meet the predicted demand plays a fundamental role in contact center operations to balance service quality and costs.

The essential structure of the staffing process consists of the following steps: forecasting contact demand, determine the minimum number of agents needed per period (staffing), select staff shifts that satisfy requirements (shift scheduling), and assign servers to the shifts (rostering) (Atlason et al., 2008). Section 2.2.1 discussed the first step and this section reviews literature in staffing and shift scheduling, which are typically conducted in sequence.

To determine an appropriate staffing and work schedule for an inbound and outbound call center, while satisfying service quality requirements, Deslauriers et al. (2007) develops and compared simulation results of five continuous-time Markov chain (CTMC) models. The first model assumes identical blend CSRs (capable of handling both inbound and outbound calls) and call arrival rate following a Poisson process. The second model extends the first by considering mismatched calls and agents with parallel dialing. Further on, the third model assumes two types of agents with distinct and exponentially distributed service times for inbound and outbound calls. In contrast to the third model, model four considers the same mean service times for both call types. Finally, model five assumes only blend agents with different service times depending on the call type. All models are developed under the time-stationary assumption and extended to accommodate piecewise-constant doubly-stochastic arrival rates. Future research directions recommendations includes using the proposed model for optimization purposes.

Legros et al. (2015) studies a call center with multi-server queuing system handling inbound calls (foreground job) also executing a background job (emails). They propose a scheduling policy that uses server's idle hours to do as many emails as possible with no harm to customer service performance, maintaining a satisfactory level of call waiting times. The call arrival process implemented follows a non-homogeneous Poisson process, and all servers are identical and capable of handling both service types. Call and email service times are different, but the foreground job is prioritized and should be quasi-instantaneous answered, while emails could be delayed for hours. One highlighted advantage of the proposed model is its ability to react to changes in the arrival process.

Researchers also advanced approaches to integrate the problem of staffing and shift scheduling. Kim and Mehrotra (2015) introduces a novel integrated staffing and scheduling model formulated as a two-stage stochastic programming with mixed-integer recourse model. The model allows for adjusting optimal staffing levels and schedules as demand data is updated over time. The authors used three years of real data from nurse scheduling from an U.S. hospital to evaluate the new model performance. The study contributes to the literature advancing the approach of incorporating uncertainty in workforce planning. Finally, Bodur and Luedtke (2017) continue to study the integrated staffing and shift scheduling problem with customer arrival rate uncertainties considering a multiskill multiclass system. Results from an empirical test using a bank call center dataset indicate that the integrated stochastic programming integer model yields significant savings in cost at the same service quality levels or presents higher service quality at the same cost in comparison solving the staffing and shift scheduling problem individually.

2.2.3 Contact Routing

Contact routing consists of processes that distribute individual contacts to different channels or to direct customers to receive assistance from customer service representatives with multiple skill set (Ali III, 2010; Mehrotra et al., 2012). Accurate contact routing is an important step for effective and efficient service (Tang et al., 2003). The task to match customers to agents usually combines routing and queuing processes. Customers allocation can be as simple as a first-in, first-out (FIFO) process, or more complex requiring rules that considering servers and customers heterogeneity (Ahghari and Balcioglu, 2009).

Research in contact centers solves the problem of optimizing routing mechanisms with dynamic programming techniques (Armony and Ward, 2010; Mehrotra et al., 2012; Chan et al., 2014; Ibrahim et al., 2016). Armony and Ward (2010) formulates a dynamic control problem to minimize expected waiting times while maintaining fairness among servers. The model assumes heterogeneous servers and a single customer type with arrival rate following an independent Poisson process. Additionally, in contrast to the longest idle server first (LISF) policy, a non-idling assumption to the optimization problem is considered in Armony and Ward (2010). In other words, there is no fixed proportion of desired idleness between different servers groups. Experimental results indicated that the proposed policy outperformed the typical used LISF policy. The authors do not explicitly address the idleness fraction parameters determination problem, but recommend further research to better define server fairness and adapt the problem formulation for heterogeneous servers and customers.

Novel contact routing strategies were proposed in Mehrotra et al. (2012) and Chan et al. (2014). The first considers the problem of routing multiple call types and heterogeneous CSRs in a financial service center. They propose a dynamic policy, based on the CSRs attributes and the current state of the system, to maximize the call resolution rate subjected to the customer waiting time constraints. Different call arrival rates distributions, not following Poisson process, are assumed for each call type, and multiple service rates and call resolution probabilities were considered for the distinct agent groups. The objective function under consideration are subjected to constraints of minimum and maximum utilization rates for the different agent groups, and minimum and maximum effective utilization rates for each call type. Simulation experiments, assuming several call arrival rates processes, service rates, and call resolution probabilities distributions, validate the model effectiveness by resulting in near-optimal call resolution rates with relatively low average speed answer.

While also considering multiple call and server types, dynamic routing policies using an index function to match call type and CSR group type is proposed in Chan et al. (2014). The new policies are expressed in terms of call waiting times and number of idle CSRs. The agents in the same group are assumed to be homogeneous and distinct groups are of the same size. However, no particular call arrival process or service time and patience time distributions is taken into consideration. Several simulations are conducted to evaluate the models. In comparison to other policies commonly used in practice, the proposed model mainly presented the best results or at least competitive with the best. Furthermore, Ibrahim et al. (2016) incorporates agent and call type heterogeneity, and time-dependent agent performance to model inter-dependent service times.

There is a growing body of literature leveraging machine learning and natural language processing techniques to enhance contact routing and queuing processes. For instance, Ali (2011) proposes an online scoring process using supervised machine learning to score available agents against customers waiting to receive assistance. Input data includes customers and agents demographics, costs, and customer satisfaction. In a similar vein, Mehrbod et al. (2018) develops classification models to determine customer satisfaction after calling an Insurance company call center. The authors use customers and agents demographic data and implement four learning algorithms to classify the call outcome as "good" or "bad". Results from both studies aim to improve caller-agent pairing. Finally, Sarikaya et al. (2014) and Rustamov et al. (2018) explore the natural language call routing problem. Their common goal is to better identify callers intent from text data generated from speech-enabled systems and use the information to improve customer-agent matching.

2.3 Summary

Organizations in distinct domains integrate digital technologies into their operations strategies to improve operational efficiency, support robust decision-making, and increase agility to respond to market trends, demands, and innovations. Big Data Analytics help operations managers making decisions not hypothesis-driven, but based on patterns learned from data. Examples of Big Data Analytics applications in operations management include predictive analysis in product development and product lifecycle, production quality management, predictive maintenance, inventory control, and demand forecasting. In this chapter, we reviewed an extensive list of studies in data-driven operations management that leverages data generated and collected from customer reviews, product recalls, production processes, stock inventories, online transactions, marketing, and economic indicators. To embrace digital transformation or expand existing digital capabilities, companies should also understand the requirements and implications for operations management of adopting a data strategy. Key aspects to consider are adapt to the data-driven decision-making process, expand data collection, treatment, and management.

There is a considerable amount of research in the service industry related literature developing new approaches to predict demand, understand customer behavior, engage with customers, and optimize business processes. Multichannel customer support service systems play a relevant role in the industry for establishing a close relationship with customers and being a source of rich customer interaction data. Existing research in customer support services, in particular call and contact centers, often investigates how to forecast customer contacts and arrival rate and to optimize contact routing and staff scheduling. However, typical approaches develop time series and queue theory models to predict aggregate demand and, therefore, do not leverage the big data collected on customers to make predictions at the individual level.

Aggregate demand forecasting remains crucial for operations planning. Nevertheless, understanding and anticipating individual customer behavior allows firms to improve service customization, generate insights to support and evaluate multichannel strategies, enable intelligent self-service capabilities, and optimize contact distribution processes. Despite the potential benefits, few studies have explored and leverage big data in customer services. Furthermore, having a holistic view of how to incorporate digital technologies in operations and understand the tangible benefits is vital to align the different management levels of an organization to commit to the transformation journey and, therefore, better support the decision to make a technology investment.

Chapter 3

A Systems Perspective on Contact Centers and Customer Service Reliability Modeling

3.1 Introduction

Customer satisfaction is influenced by the customer's assessment of the characteristics of the product or service they consume (Gronholdt et al., 2010). Companies try to achieve higher levels of customer satisfaction by establishing a relationship with customers. A satisfied customer brings benefits such as loyalty to the organization, buys with more frequency, and speaks positively about the brand (Gronholdt et al., 2010; Reichheld, 2003; Wallace et al., 2004). Therefore, one of the ways companies use to approach customers is by offering personalized service and multiple channels of communication (Warrington et al., 2007).

Excellence in the services provided and the understanding of the customer's individualities are important points for this approach to be positive. Communication channels put companies and customers in contact, generating a common environment for relationship and interaction. Given the ability to collect rich customer interaction data, they can have strategic importance to the organization by creating opportunities to improve their business process and forecast customers' needs (Moazeni and Andrade, 2018).

Companies should have well-structured communication channels so that customers can communicate effectively, get information, solve problems, and provide feedback. The feedback can be used to identify product or service deficiencies and thus point to opportunities for improvement. Therefore, organizations should be capable of rapidly adapt to meet new customers' demands. It requires a deep understanding of their structure, cross-enterprise functions and processes, and constant monitoring of the environment and operations. The goal of interpretation of the ecosystem is to identify what needs to change and how, while the purpose of monitoring the operations is to ensure the reliability and the quality of customer service, two important factors impacting customer satisfaction (Santouridis and Trivellas, 2010).

Contact centers are structures that offer multichannel customer support services and act as the interface between companies and customers. Consequently, contact centers have been increasingly recognized for their role in ensuring customer satisfaction (Tate and van der Valk, 2008; McNair Andrew & Holmes Richard, 2015). Due to its strategic position in the customer relationship, and also for the high costs involved, contact centers and their precursors call centers (which offer only telephone support), have been extensively studied by different research areas (Gans et al., 2003; Aksin et al., 2007). For example, research has focused on aspects seeking to reduce operational costs. These studies include the development of predictive models for demand, staffing, and scheduling optimization. Other research areas such as psychology, marketing, and management seek to provide understanding into customer behavior and customer relationship, e.g., see Tate and van der Valk (2008) and Tezcan and Behzad (2012). However, studies that consider systems engineering principles, along with business analytics processes, are scarce in the contact center domain.

Given the critical role played by contact centers, they need to adapt to incorporate innovations required to address new customers' needs and aligned with the organization's governance policies. For example, innovations can happen in the communication channels, in the infrastructure to keep contextual customer data shared across different platforms, or procedures to assist customers. In this context, those in the position of making decisions that infer substantial changes in the structure or operation processes, require tools to transmit information of the current and desired state of the business. That being said, enterprises and their components (e.g., contact centers) could benefit if analyzed as systems.

A systems perspective provides an approach to model the cross-enterprise processes and linkages among subsystems and components (Rouse, 2005a; Despeisse et al., 2012). This facilitates developing a holistic understanding of a system under inquiry and introduces flexibility and adaptivity capacities essential in transformation analysis. By relying on systems perspective, one can more accurately evaluate the performance of a system under various scenarios and to better predict system outcomes as the operation processes or input variables change.

Hence, this chapter investigates contact centers from a systems perspective and contributes to the cross-disciplinary operations management literature by developing a contact center system architecture and proposing a bottom-up model for the service system reliability. We adopt two systems thinking techniques, the Systemgriam and the Conceptagon, to understand contact centers as enterprises, to identify the driving forces influencing their behavior, and to develop a contact center operations management system architecture (Giachetti, 2010). The system architecture maps the interfaces among the contact center components. Building on the contact center system architecture, we propose a service system reliability metric defined by key contact center processes and system input data. Since the reliability metric explicitly takes into account the contact center processes, the modeling approach can dynamically adapt to changes in the system processes or their input data. This characteristic makes our modeling approach appropriate to assess the system as a whole, conduct transformation analysis, and discover improvement opportunities.

Enterprise transformations are context-driven, and customer satisfaction is the essential value of the service domain ecosystem (Rouse, 2005b). Reliability is a fundamental dimension of service quality (Parasuraman, Arun and Berry, Leonard L and Zeithaml, 1991), which has been shown to have a relevant influence on customer satisfaction (Santouridis and Trivellas, 2010). The holistic understanding of the contact center operations management system shows that the overall system performance depends on each component's ability to properly execute its function in the customer service process. The reliability of each component is subject to different failure definition and probability of occurrence. Existing performance metrics are deterministic and do not have the necessary capacity to adapt to the different factors impacting the whole system's ability to perform its function. The proposed contact center reliability model is general and adaptable to changes in the configuration of the underline system. Its probabilistic formulation enables the anticipation of the effects of fundamental changes to the system design via simulation.

Finally, we illustrate the application of the system architecture, the reliability metric, and performance control using empirical data from a contact center of a major insurance company. We show that the distributions of the daily reliability of each phase of the support service process and the overall system reliability all present distinct characteristics. We provide evidence that the contact resolution is a function of the individual probabilities of the proper functioning of each phase of the customer service process.

The remaining sections of the chapter are structured as follows. Section 3.2 reviews relevant work in enterprise modeling, and contact center system design. Section 3.3 presents the contact center operations management system architecture and describes its components and functions. Section 3.4 introduces a mathematical formulation for the contact center operations management system reliability. Section 3.5 discusses the visual tool to be used for performance management and control of the system. In section 3.6, empirical data is used to illustrate the applicability of the design, and the system performance evaluation using a control chart. Finally, Section 3.7 contains concluding remarks and directions for future research.

3.2 Contact Centers from a systems perspective

3.2.1 Contact Centers' Background

The term contact center refers to customer relationship and support operations that include the provision of customer service and support via various communication channels such as telephone, SMS, email, webchat, and social networks (L'Ecuyer, 2006; Rijo et al., 2012). There is an extensive literature on call centers and contact centers, e.g., see Moazeni and Andrade (2018); Gans et al. (2003); Aksin et al. (2007); Mandelbaum (2004), and the references therein. The literature has been focused on demand forecasting (Moazeni and Andrade, 2018; Tezcan and Behzad, 2012; Avramidis et al., 2004; Ibrahim and L'Ecuyer, 2013; Taylor, 2012) using particularly queuing models (Koole and Mandelbaum, 2002; Brown et al., 2005), workforce forecasting and capacity planning (Aldor-Noiman et al., 2009), staffing and workforce planning (Borst et al., 2000; Pot et al., 2008; Cezik and L'Ecuyer, 2008), and scheduling (Fukunaga et al., 2002; Liao et al., 2013; Mattia et al., 2017).

Contact centers are important components of companies' overall customer relationship management (CRM) systems. For the literature on CRM, see Kumar (2010); Dyche (2001); van der Aa et al. (2015); Nagar and Rajan (2005). Contact centers are seen as a cost necessity by executives from companies in different industries (Tate and van der Valk, 2008). Due to its close relationship with customers, it plays an important role in the enterprises' customer loyalty strategies. Whether as a simple information platform or as a channel to sell products and services, the contact center's success will depend on the ability to respond to customer needs and market changes.

Therefore, the essential function of contact centers is to connect customers and companies with high-quality services to attain customer satisfaction. Figure 3.1 illustrates the processes involved in the customer support provided by a contact center.

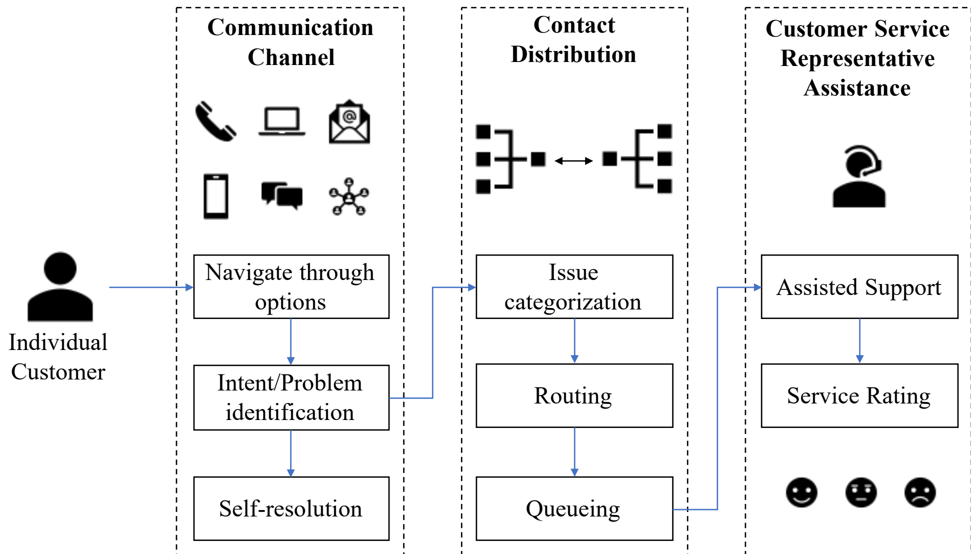


Figure 3.1: Customer service support flow diagram.

Nevertheless, the literature is limited when analyzing contact centers concerning the relationship among its constituent components and the interaction with the environment they operate. A better understanding of the contact center interfaces can support strategic management decisions when inferring transformations. Systems engineering offers mechanisms and principles that guide the identification, definition, and representation of the relationships between entities of contact centers and their interaction with the environment. In turn, it can contribute to transformation's decision-making processes in customer support systems by proving a holistic understanding of the system (Rhodes et al., 2009).

In such a dynamic environment, where new technologies and market opportunities emerge and incite changes in how contact centers offer their services, understanding the big picture and assess how changes made to its parts affect the whole is important to evaluate what-if scenarios. Therefore, given the emergent properties of customer support systems, it seems to be evident that contact centers can benefit from the implementation of a systems engineering perspective.

3.2.2 Motivation for a Systems Approach

Companies constantly seek to adapt to the context (e.g., economic conditions, regulations) in which they operate. The need to change is driven by a recognition of a real or anticipated gap in the value delivered by the organization to its stakeholders, which can be influenced by new market and technologies opportunities, competitor's initiatives, or enterprise crises. These phenomena, commonly addressed as enterprise transformation, tends to be discontinuous and implies fundamental changes rather than continuous improvement of routine processes. The result of fundamental changes includes the development of new products, technology innovation, and new service standards. Therefore, enterprises should have the ability to interpret the ecosystem to increase market share, revenues, profits, and the capacity to quickly adapt and incorporate these changes (Rouse, 2005b).

The way organizations understand the environment highly relies on the nature of their business function. Each domain executes different processes via distinct work practices and, therefore, are subjected to distinct forms of transformation. Companies analyze transformation by gathering actionable insights from the environment that support strategic management decisions and their impact on the overall performance of the business. In this respect, a systems perspective allows a holistic view of an enterprise (Rhodes et al., 2009; Hubbard et al., 2010; Sitton and Reich, 2015).

Enterprises are complex and open systems operating in a dynamic environment and facing increasing competition (Giachetti, 2010; Rouse, 2014). Design enterprises as systems aim to create a model that enables effective decision-making by enhancing the ability to understand system behavior. The systems thinking approach views enterprises as a whole, which combines a critical understanding of cross-enterprise elements relationships and their interaction with the environment. Relying on a systems perspective is key to find improvement opportunities and achieve solutions to the system as a whole and avoid local or sub-optimal solutions (Despeisse et al., 2012). Organizations should be designed as systems to easily adapt due to unexpected changes in their environment or to bring in competitive advantages opportunities (Giachetti, 2010).

Systemic thinking refers to a set of concepts, behaviors, and tools that aid in the understanding of interdependent structures of complex systems. This approach seeks to explain causal relations between a dynamic set of interrelated factors. It is a discipline that allows us to analyze issues in a holistic and integrated way, to envisage dynamic sets of behavior, to identify and understand interconnections that give the system unique characteristics as well as to envisage likely behaviors. In general, the analysis begins with the identification of a problem. Next, we identify the factors and variables that affect this problem. An analysis of the interrelationships between these variables allows us to identify probable behaviors of the system and contributes to the planning of interventions with greater potential to adapt the behavior of the system to the desired objectives.

Previous research has successfully implemented systems thinking approaches to problems of different areas of knowledge. For instance, in the manufacturing domain, a systems perspective was taken for the development of a framework for sustainability assessment system (Moldavska and Welo, 2015) and a conceptual factory ecosystem model to improve activities related to sustainability of production (Despeisse et al., 2012). In addition, systems perspective was applied to review the Defense Acquisition System for the U.S. Department of Defense (DoD) (Cilli et al., 2015), the development of a methodology for address cybersecurity (Bayuk and Horowitz, 2011), and to evaluate business model evolution (Velu, 2017). The researchers advocate that the use of these techniques contributes to cope with the complexity of the systems, instead of neglecting it, ensuring that all important aspects of the systems are brought to light when proposing novel solutions for their respective problems.

Contact centers can be analyzed as enterprises. It fits in the definition provided in The Open Group Architecture Forum (2018) as being a collection of organizations sharing the same set of goals and a single bottom line. The common goal shared by all its components is to provide the necessary and promised support for its customers. Therefore, having a contact center system design can be a valuable basis for the process guiding the evolution of an organization by providing adequate representation of the current and desired states of the enterprise (Aier and Gleichauf, 2010; Aier, 2014).

3.2.3 Contact Center System Design

A formal approach to analyze the relationships among elements of contact centers is through the evaluation of systems' design (Sage and Lynch, 1998). The importance of system design in call centers is discussed in Avramidis and L'Ecuyer (2005). In particular, an optimal design of an agent's skill set and infrastructure on the overall call center performance is investigated in Aksin et al. (2007). In Jaaron and Backhouse (2011), call centers are investigated under a systems perspective and focus on understanding the impacts of service operation design to the working experience in the context of a manufacturing enterprise. The study shows that improvements in the operations service design with the employment of organic structure through a Systems Thinking approach, enhance both the service quality and employees working experience.

According to Rijo et al. (2012), contact center information systems (CCIS) designs are complex, involving a diverse set of actors and interests. In Rijo et al. (2012), a framework for the CCIS design, along with a step-by-step design process, is introduced. The goal is to provide scientifically-based design principles to support the solution of numerous non-deterministic problems faced regularly by contact centers. Despite contributing with a solid conceptual foundation for the development of a contact center system design, a contact center system architecture is not elaborated in Rijo et al. (2012).

A contact center design is proposed in Schoeller and Heffner (2014). The design-building process includes several steps of the framework in Rijo et al. (2012), and concepts of enterprise transformation (Rouse, 2005b). The proposed architecture is composed of six elements: customer channels; contact routing, filtering, and queuing; agent workspace; contact analytics; event collection; and manager workspace. However, while this design covers several aspects of the framework in Rijo et al. (2012), it does not consider features of systems thinking approaches to promote a clear representation of the relational characteristics of the system. For instance, the architecture design introduced in Schoeller and Heffner (2014) fails to include components that interact with the ecosystem in which contact centers operate. Feedback loops are essential to enterprises to maintain stability in dynamic environments (Giachetti, 2010). This strategic role can be exerted by organization governance management.

The organization's governance is responsible for establishing the policies, hence fundamental to the process of development and operation of a contact center. Supported by insights extracted from the domain ecosystem, the actors of this level have the ability to analyze, propose, and implement changes that target the improvement and efficiency of all processes of the customer support system. The governance defines the scope of performance metrics, establish benchmarks, and impose the appropriate corporate culture that aligns the provided services to its core values.

Furthermore, the views of the contact center architecture in Schoeller and Heffner (2014) are not homogeneous. While some views represent processes and functions, others are workspaces. To address this issue, we will take a Systems Thinking approach as a formal method to map the interfaces of interoperability between and among cross-enterprise components. This chapter develops the contact center system design based on the systems engineering bundles. A graphical representation of a system is a useful tool to document the components and processes of the system, as well as describing their functions and interfaces with coherence and consistency.

3.3 Contact Center Operations Management System Architecture

This section describes the contact center operations management system design process, presents the system architecture, and provides details of its components. To guide this development process, we use two system thinking tools: the Systemigram and the Conceptagon framework (Boardman and Sauser, 2013, 2008). The Systemimgram and Conceptagon were chosen because they provide an understanding of the bigger picture, facilitate the identification of information flows between participating components, and serve as reference modes for systems design.

Systemigram is a systems methodology that identifies and represents key elements of systems as a network made up of parts (nodes) and relationships (links). They provide an understandable representation of complex systems and have been successfully implemented in a variety of domains and problem (Cilli et al., 2015; Patrick Eigbe et al., 2009; Sauser et al., 2011).

Figure 3.2 depicts the Systemigram built for the customer support system. Nodes in yellow constitute the mainstay, which articulates the motivation of the model

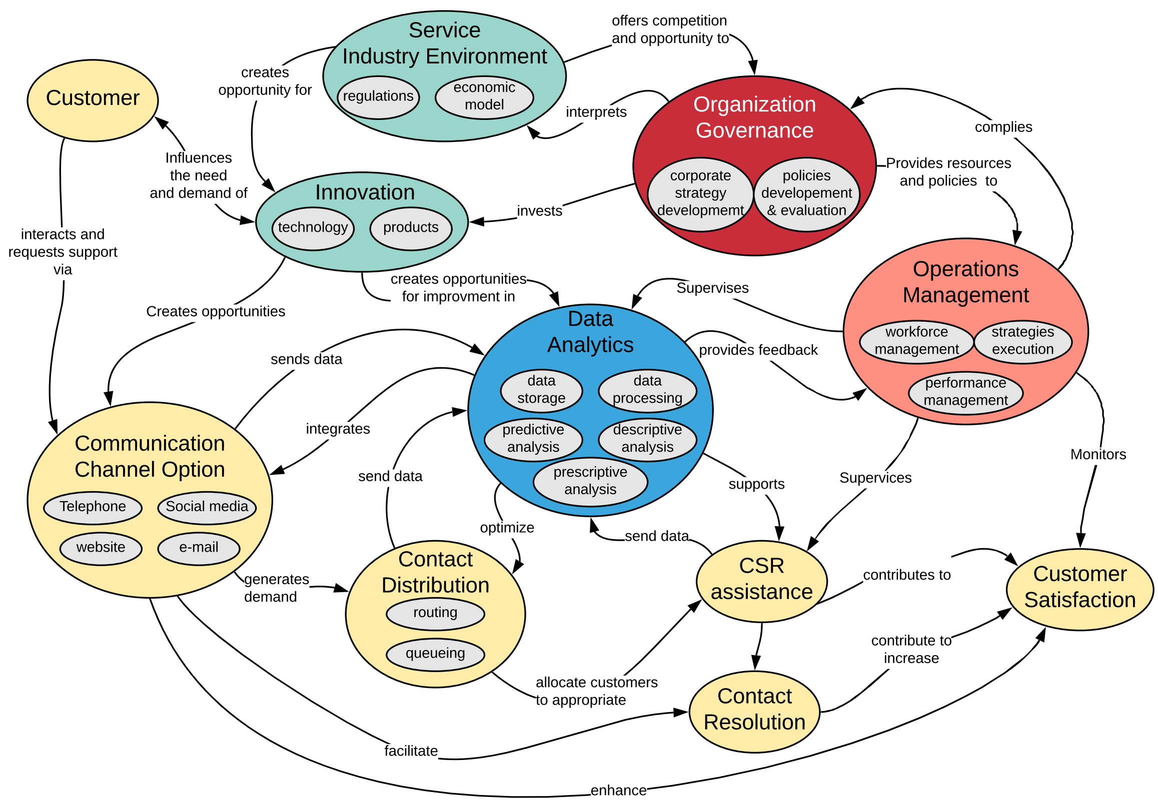


Figure 3.2: Visualizing functions, processes, and factors influencing the customer support system through the use of a Systemigram.

for the strategic intent, its mission, or how it may be accomplished. As discussed, the contact center's purpose is to achieve high customer satisfaction by properly providing the requested support. Then data analytics node plays a central role in the whole customer service process as it concentrates the most number of links, in and out. The Systemigram also provides a vision of how the organization governance interacts with the domain ecosystem and influences the service process through different paths.

While Systemigram helps to articulate the story of our contact center operations system, the Conceptagon is a tool to systematic structure the examination and interpretation of the system of interest. It consists of seven dimensions combining attributes that characterize important properties of the system (see Figure 3.3). Updates and organization of these dimensions into three groups are proposed in Wade

(2016). Table 3.1 provides a brief description of the dimensions as presented in Wade (2016). Next, we describe how each dimension defines the contact center operations system.

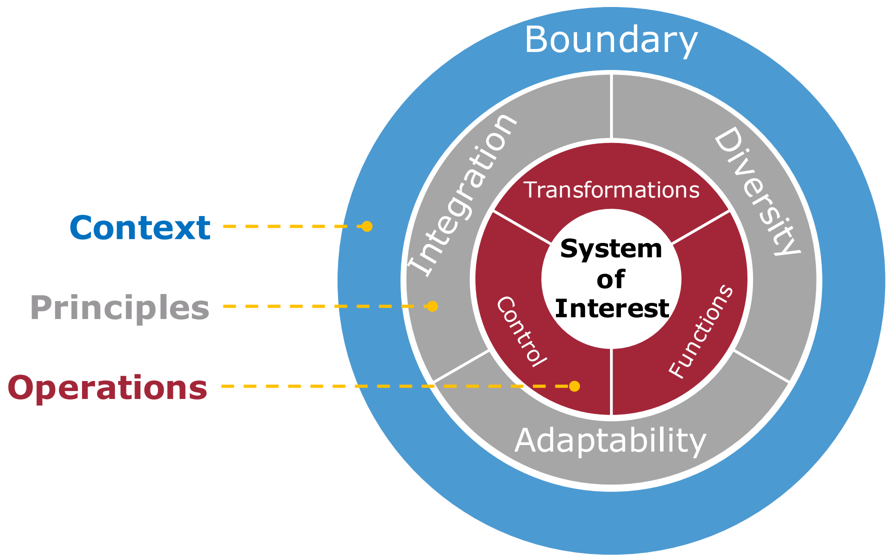


Figure 3.3: The Conceptagon.

Table 3.1: Description of the conceptagon's dimensions (Wade, 2016).

Group	Dimension	Description
Context	Boundary	Determines the system of interest, what is intended to be influenced, and what is accepted as a given.
Principles	Diversity	Balance between multiple views and ability to act coherently.
	Adaptability	The ability to be hierarchical in objectives and net-centric in execution.
	Integration	Defines how the overall system is constructed. Which pieces can stand on their own, which can only function together, and how they are interconnected.
Operation	Control	Determines the actions and behaviors of each element in the system.
	Transformations	The processing of information, energy and material received via the inputs to serve the purposes of the system resulting in desirable outputs.
	Functions	The operations that are required to perform the necessary system transformations.

Boundary: Interior, Exterior. The contact center operation system consists of all elements and processes that influence how the service and support are provided to the customer. That includes communication channels used for interactions, technical and physical infrastructure resources, and employees. External to the contact center is the service industry ecosystem in which it operates, government regulations, the economic model, and competitors.

Diversity: Homogeneous, Heterogeneous. The contact center operations system is a collection of entities performing different functions and processes. Although budget, technology, and human factors can impose constraints on operations, the contact center's main focus is to properly provide support for customers. All components of the system, while executing different functions and having individual goals, still share a common objective to serve its customer meeting quality standards and achieving their satisfaction.

Adaptability: Rigidity, Flexibility. The system is open to the possibility of new resources and functions be incorporated to accommodate requirements of the evolving customer support environment such as new communication channels and embedded technologies, or new products and services offered by the company that may require a different support approach. There is a constant exchange of information, labor, technology, and service within the system or with the ecosystem. Market regulations, social, and political conditions can affect and motivate changes in the contact center operations. Nevertheless, a hierarchical structure of the system articulates the objectives and execution of the service provided.

Integration: Monolithic, Modular. In terms of modularity, some types of communication channels can provide self-contact resolution capabilities. However, the monolithic property is perceived as the customer service representative assistance depends on the communication channel, and contact distribution connect them to customers.

Control: Centralized, Distributed On the one hand, the organization's governance of the contact center is responsible for formulating and impose the directions for the operations seeking to fulfill the strategic goals of the system. On the other hand, decision-making in the execution of these directions and policies are distributed in lower levels of the system. For example, there are decisions on the operations management to determine the ideal workforce schedule or the choice of methods to be used in the optimization of the routing and queuing process.

Transformations: Inputs, Outputs The system receives distinct inputs such as customers' inquiries, market opportunities, competitions, and technologies advancement. Disruptive events, e.g., cyber attack to the website of the database server, may also force abrupt changes to the system. Outputs include the contact resolution and the customer experience and satisfaction with the service provided.

Functions: Structure, Process. The purpose of the contact center is to connect customers and organizations with high-quality services to attain customer satisfaction. There are several processes involved in the customer support service, as shown in Figure 3.1. The Systemigram explores the interconnectivity of these processes and functions.

As a result of the insights gained from the Systemigram and discussion of the Conceptagon dimensions, we propose an architecture of the contact center operations management system (CCOMS) (Figure 3.4). The system represents the cross-enterprise operational and managerial processes and functions necessary to provide the customer support services. The system design is general and adaptable enough to be implemented across industries. It consists of four components: contact interaction, contact analytics, operations execution, and organization governance. These components interact with one another, as indicated by the arrows in Figure 3.4, and perform distinct, but complementary, functions in the customer support service process.

1. **Organization Governance**. It is a set of activities with an emphasis on cross-functional management. The focus is to look outside the company, embracing

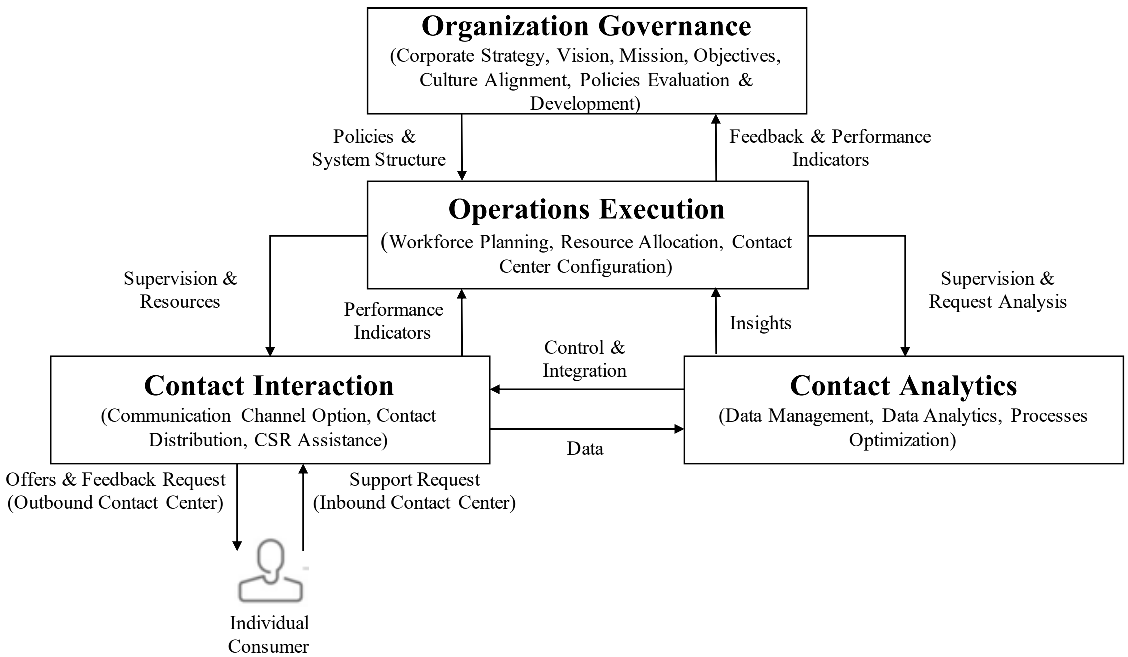


Figure 3.4: Contact center operations management system architecture.

the interpretation of the ecosystem (e.g., economic conditions, market and technology opportunities), and translate into policies aimed to achieve the desired future. The set of practices defined here seeks to maintain and improve the quality of the product or service offered by developing internal and external control mechanisms of business activities. These practices induce a culture to be adopted by the company and its stakeholders. The corporate culture will guide employee behavior and the technologies used, will also make changes to improve the process, describe the scope of performance metrics, and establish benchmarks (Rouse, 2005b; Sharp, 2003).

2. **Operations Execution**. It is a set of activities with an emphasis on functional or operational management. Comprises functions and processes necessary to provide resources and supervise contact interactions and contact analytics. That includes workforce planning, resource allocation, and contact center

configuration. All activities are performed closely aligned with the strategic guidelines established by the company's leadership.

3. **Contact Analytics**. Based on the request and inputs from the operations management, the contact analytics focus on control and integration of the contact interaction processes, by gathering and managing data coming from the customer support process, employing data analytics and processes' optimization. Finally, it provides actionable insights into the operations' execution.

4. **Contact Interaction**. It is a group of activities of the contact center interface with the customers. Consists of processes including the communication channel selection, contact distribution, and customer service representative assistance. The communication channel option refers to the medium of contact chosen by customers to connect with the company. Several channels can be used, e.g., website, mobile app, chat, interactive voice response systems (IVR), and phone. The contact distribution corresponds to contact identification, routing, filtering and queuing. It identifies the customer along with its intent and transaction type. The proper identification of the customer is essential to determine further queuing and routing processes to match customers to skilled CSRs capable of providing the requested service.

Our proposed design is a systematic view of contact centers that bring together operation and governance management elements. We aim to contribute to transformation's decision-making process by providing a systems design for contact centers grounded by enterprise systems engineering (ESE), specifically systems architecture (Sitton and Reich, 2018), and systems thinking concepts (Mcgee and Edson, 2010). The contact center operations architecture provides a clear representation of the system's components, their functions, and interrelationships. The contact center operations management system architecture can work as a tool to help the organization to explore its strengths and identify potential improvement opportunities in the processes involved in the customer service offered.

From the systems engineering perspective, the CCOMS architecture captures ESE attributes that contribute to cross-enterprise connectivity and improves cross-enterprise process functionality and performance, as described in the conceptual model presented in Sitton and Reich (2018). Moreover, from an operational point of view, the system architecture contributes to providing a common situation picture of the contact center, in an integrative holistic view.

The proposed CCOMS architecture is general and can be implemented across different domains. The differences between one company or industry to another can be related, for example, to the budget available, which may affect the number and type of communication channels and technology used; the service type provided and the required CSR skills set; and any adjacent systems that interact with the CCOMS in order to execute the service or delivery the product to the individual. To illustrate the generalization of the proposed design, we briefly explain how it could be applied to an arbitrary retail company and for emergency response services.

For a retail company, the contact center would interact with other systems to deliver a product purchased by a customer. The purchase order, placed through the company website or telephone, is directed to a warehouse or inventory department. The request is then forwarded to a distribution center in charge of the logistics to ship the product to the customer. In the second implementation example, 911 calls made by individuals are typically routed to the closer Public Safety Answering Point (PSAP) based on the caller location information. In contrast to the retail industry scenario, PSAPs need to redirect the contact or forward information to the appropriate emergency service department (e.g., Fire Department, Police Department, Emergency

Medical Services), which may depend on the type of emergency.

In summary, in both examples, the essential function of the contact center remains identical, which is connecting the customers to the corresponding organization, satisfy customer needs (e.g., buy a service or product, or request information and emergency assistance) with high-quality service. In that case, although common performance metrics (e.g., average speed to answer, abandon rate, service level (Gans et al., 2003)) of contact centers can be used to evaluate service quality and operational efficiency in any of the cited implementations, their importance is subject to each industry interests. While the customer experience and customer loyalty can have higher priorities for a retail company, emergency services prioritize providing rapid and effective support.

In the next section, we discuss the performance evaluation of the CCOMS based on system reliability.

3.4 Customer Contact Center Operations Management System Reliability

The concept of contact center system reliability is intrinsically related to the value of the service domain, as discussed in Rouse (2005b), and the quality of the service offered. In this section, we discuss the definition of service value and review the systems' reliability literature. We then conclude by introducing the formulation of the contact center operations management system reliability.

3.4.1 Contact Center Service Value

According to Rouse (2005b), "enterprise transformation is driven by perceived value deficiencies relative to needs and/or expectations" and "involves examining and changing work processes". The concept of value is essential to characterize the current state of the system and to infer what the desired future state is. In the service industry, this value is customer satisfaction with the service received (Tate and van der Valk, 2008; Rouse, 2005b). Recognized as a relevant part of corporate strategy (Fornell et al., 2006), customer satisfaction has transaction-specific and cumulative conceptualizations (Buzzel, 1990; Anderson et al., 1994). The first is an assessment of a choice made after a specific purchase occasion (Hunt, 1977). The second defines customer satisfaction as a cumulative overall assessment based on total purchase and post-consumer experience with a good or service over time (Fornell, 1992; Anderson et al., 2004). In other words, the concept of customer satisfaction is a matter of customer's perception of the expected product or service provided (Oliver, 1981; Parasuraman et al., 1988).

The relationship between satisfaction and quality of service is a well-explored topic (Gronholdt et al., 2010; van der Aa et al., 2015; Chumpitaz Caceres and Paparoidamis, 2007; Luo and Bhattacharya, 2006). Customer satisfaction is typically associated with the form the service is provided and the quality of this service. Although customer standards and requirements may evolve and differentiate from one domain to another, the goal to maintain satisfaction at high levels remains the same: increase acquisition, retention, and customer loyalty, seeking sales growth. Additionally, customer satisfaction has been identified as a key factor in retail customer loyalty (Santouridis and Trivellas, 2010; Cronin et al., 2000). Previous research has indicated the intermediary function of customer satisfaction on the service quality and customer loyalty relationship (Santouridis and Trivellas, 2010; van der Aa et al., 2015). In turn, customer satisfaction is also positively affected by a multichannel environment (Gronholdt et al., 2010; Shankar et al., 2003), such as contact centers. Evidence of the association of this environment with customer satisfaction and sales growth can be found in Wallace et al. (2004) and Thomas and Sullivan (2005), respectively.

A positive relationship between customer satisfaction and quality aspects of contact centers is found in ko de Ruyter and Wetzels (2000) and Feinberg et al. (2000). Customer contact center quality is defined in van Dun et al. (2011) as "the overall evaluation of the customer contact center, as perceived by customers". Following this definition, a conceptual framework is developed in van der Aa et al. (2015) that represents the direct and indirect effects of the quality of the contact center with customer loyalty, mediated by the relationship quality, which is composed of satisfaction, trust, and affective commitment. Finally, the study concludes that customer contact centers' quality has a direct and positive influence on the quality of the customer relationship, which consequently positively affects customer loyalty.

The concept of service quality is a controversial topic due to the subjectivity of quality evaluation (Chumpitaz Caceres and Paparoidamis, 2007), It relies on the comparison between expectations and perceptions of the service provided performance (Parasuraman et al., 1988; Gronroos, 1990). Relying on the definition of quality in terms of expectations and perceptions is an issue since both vary for different individuals. Services are, by nature intangible, which gives them a character of heterogeneity that makes it difficult to evaluate quality. The very trend of service variability can cause the same customer to have different perceptions of the same service on different occasions. Service quality can be defined as the degree to which customer expectations are met exceeded by their perception of the service provided (van Dun et al., 2011; Churchill Jr and Surprenant, 1982).

Several studies have focused on measuring service quality. An instrument called SERVQUAL is introduced in Parasuraman et al. (1988). This tool comprises 22 items categorized in five dimensions: reliability, tangibles, responsiveness, assurance, and empathy. The main criticisms of the instrument are related to the indirect difference score approach (Brown et al., 1993) and measurement of expectations (Teas, 1993), which are susceptible to the reliability, variance, and validity issues (Oh, 1999).

The SERVQUAL dimensions are divided into two groups. The first group consists of Tangibles, Responsiveness, Assurance, and Empathy dimensions, which related to the service process. The second group includes only one dimension, reliability, and is related to the outcome of the service. Moreover, among the five dimensions, reliability is considered the most important in Parasuraman, Arun and Berry, Leonard L and Zeithaml (1991) and is seen by customers as the service "core". Hence, the priority is to receive what has been promised and to have eventual problems resolved. Several studies have reaffirmed such importance (Gronroos, 1990; Johnston, 1995; Ravichandran, 2010; Leong et al., 2015).

The concept of service quality with the focus on contact centers is discussed in van Dun et al. (2011), where other service quality dimensions beyond those considered in SERVQUAL are identified. The proposed scale comprehends seven factors and 46 items. The two scales agree on having reliability and empathy, but the latest includes accessibility, the waiting cost, voice response unit, knowing the customer, and customer focus. Furthermore, the reliability dimension in the customer contact center quality consists of 11 items rather than five, as in the SERVQUAL, as a result of considering factors that have more impact on the service quality perception from the customer point of view. They argue that this need is due to evidence that internal operational measures such as service level and average talk time have little impact on the customer experience (Feinberg et al., 2000; Holland, 2003). This argument is in line with van der Aa et al. (2015) which concludes that customer contact centers are transitioning from a transaction-oriented cost center to a relationship-oriented value center.

Given its importance in measuring the quality performance of the contact center service, it is, therefore, necessary to properly define reliability in contact centers.

3.4.2 System Reliability

The definition and quantification of reliability can depend on the context and type of application, the area of knowledge, and field of research. In the disciplines of Systems Engineering, there are different formal definitions for reliability (see, for example, Nachlas (2017) and Zio (2009)), but comprising similar concepts. The following widely accepted and detailed definition of reliability is provided in Elsayed (2012), based on which we develop our study in this chapter:

> *Reliability is the probability that a product will operate or a service will be provided properly for a specific period of time under the design operating conditions without a failure.*

Based on the above definition, the concept and estimation of reliability depend on the clear characterization of proper performance, specified purpose or function to perform, the period under examination, conditions of operations, and the failure definition. As the above definition indicates, reliability is expressed in terms of probability, as it is often not a question of being or not reliable, but rather the degree of reliability. The level of the product or service ability to properly perform the specific function is commonly measured as the failure rate per unit of time, the average time to failure, average lifetime, and residual half-life (Aven, 2012). Reliability can also be understood as the measurement consistency over time, the measurement stability over different conditions, or the probability of a failure-free performance over the useful life of an item under specific duty-cycle conditions, often expressed as mean time between failures (Nunnally et al., 1967; Drost, 2011).

Within the context of contact centers, reliability is also defined in different ways. For instance, reliability is defined in Parasuraman, Arun and Berry, Leonard L and Zeithaml (1991) as "the ability to perform the promised service dependably and accurately", and Leong et al. (2015) describes it as "the service provider's capability to offer precise and trustworthy services". A reliability dimension incorporating aspects of the contact resolution effectiveness and perceived quality measurements that provide better insights on customer experience is considered in van Dun et al. (2011). Although these definitions carry important concepts of the service offered by contact centers, they do not clear quantify or provide a mathematical representation for reliability. These studies focused on identifying the important dimensions of service quality and attested the reliability's relevance.

The present study aims to close this gap. As discussed in Section 3.4.1, the contact center system reliability is one of the dimensions of the perceived service quality, and it relates to the outcome of the service rather than the process. But how can it be measured and to what extent it can be used to support a transformation decision of the contact center by assessing a future state of the system's ability to deliver its service "core" dimension?

A simulation study in Ali III (2010) compares two call center scenarios, cross-trained agents versus specialized agents, to determine which configuration is most efficient and reliable in terms of customer service. Concepts of Lean Reliability Systems, commonly employed in manufacturing operations, are used to measure the reliability of both models. The Reliable Lean system has four critical resources: materials, equipment, schedule, and personnel (Sawhney et al., 2010). The study considers only the last two. The system reliability is defined in terms of two metrics used in Failure Mode and Effects Analysis (FMEA): Risk Assessment Value (RAV), and Risk Priority Number (RPN). However, these metrics are subject to the subjectivity of the definition of risk assessment (severity of the failure, the probability of occurrence, and the probability of detection). Both RAV and RPN serve as risk prioritization tools, not as reliability metrics

There are several metrics used in contact centers to evaluate customer service performance. Among those of an operational nature, there are, for example, the Average Speed to Answer (ASA), Average Handle Time (AHT), Service Level (SL), Abandon Rate, and Contact Resolution (sometimes focusing on the first contact resolution) (Mandelbaum, 2004; Marr and Parry, 2004). While the first three metrics are related to the service process, the last two are associated with the outcome.

ASA is the average length of time it takes for a customer to reach a CSR when using a telephone communication channel. AHT is the average total time spent by the CSR taking to the customer and the after call work time and used to measure the servers' productivity. In turn, SL is the percentage of customers answered by the CSR within a specified time threshold. Abandon rate is the percentage of inbound interactions that are terminated by the customer before receiving service from a CSR, and contact resolution is the percentage of customers that have their issue solved or request provided. That being said, contact resolution can be used to partially address the question above, as it is an indicator that the core purpose of the contact center service is being met. However, as a deterministic metric, it does not directly incorporate the uncertainties and factors that influence the transaction outcome.

The system thinking approach applied to the contact center operations management system provided us a holistic understanding of the interfaces between the system's components and which can be used to infer different root-causes of the system's failure. Therefore, it is necessary to have a probabilistic formulation of the contact resolution that allows for the simulation of systems accounting for the uncertainties involved to infer about the future behavior of the contact center in different scenarios. These scenarios may involve, for example, the incorporation of new technology into the contact routing platform, or the development of a new communication channel. Thus, the model needs to be adaptable to include the different parameters that can lead to system failure.

For instance, the system failure can be related to a malfunction of the communication channels (e.g., website be offline due to a high volume of transactions or even a cyber-attack), or problems with payments via the website or mobile app capabilities can also make impossible to the customer to perform the desired transaction. Issues to identify customer's inquiries using an online automated web-chat or speech recognition in the IVR system, improper routing to CSRs with the necessary skill set, or insufficient workforce, may increase drop-off rates due to longer waiting times. Furthermore, other issues related to data collection and integration of channels and components of the contact support service can cause delays in the service process, or completely prevent the customer service process to continue. Finally, the efficiency of the processes and tools used to interact with customers is restricted by the resources provided by the operations execution management and the alignment to the organization's governance directions.

We then assume that the system components execute distinct functions in the customer service process with different failure definitions and probabilities to occur. Hence, the overall system reliability can be interpreted as a function of its components' reliability. To define the contact center reliability, we rely on the diagram outlined in Figure 3.5. Let x_t^i, $i \in \{1, 2, 3, 4\}$ represent, at time t, the component status variables where:

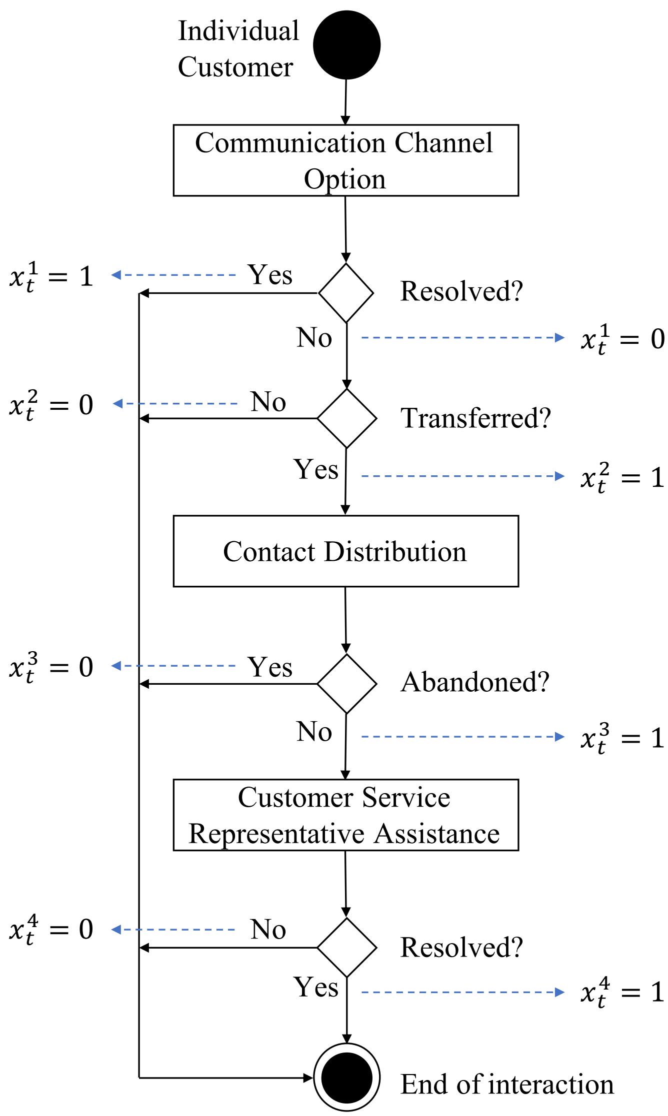


Figure 3.5: Contact flow diagram.

$$x_t^1 = \begin{cases} 1 & \text{if the contact is resolved in the Communication Channel,} \\ 0 & \text{otherwise} \end{cases} \tag{3.4.1}$$

$$x_t^2 = \begin{cases} 1 & \text{if the customer requests transference to an CSR and is directed} \\ & \text{to a contact distribution platform,} \\ 0 & \text{otherwise} \end{cases} \tag{3.4.2}$$

$$x_t^3 = \begin{cases} 1 & \text{if the customer does not abandon before being served by a CSR,} \\ 0 & \text{otherwise} \end{cases} \tag{3.4.3}$$

$$x_t^4 = \begin{cases} 1 & \text{if the customer does not abandon before being served by a CSR,} \\ 0 & \text{otherwise} \end{cases} \tag{3.4.4}$$

Now, assume that each component status variable is a function of other components' status. For instance, the contact can only be resolved within the communication channel if it is available for use (e.g., a website be online) or if it provides the necessary capabilities to resolve the issue (e.g., the IVR system offer the possibility to make a self-service payment). Also, we assume that the contact issue cannot be partially resolved. In a situation where the customer has more than one issue or request, different issues are treated as different transactions and logged separately in the system. Therefore, let each $\underline{x_t^i}$, $i = \{1, 2, 3, 4\}$, be a component status vector, where, $\underline{x_t^i} = \{x_t^{i,1}, x_t^{i,2}, ..., x_t^{i,N}\}$ and N be the total number of component status variables taken into consideration. The structure functions are

$$\phi(\underline{x_t^i}) = \prod_{n=1}^{N} x_t^{i,n}, \tag{3.4.5}$$

and the system structure function is then

$$\phi(\underline{x_t}) = \phi(\underline{x_t^1}) + \prod_{i=2}^{4} \phi(\underline{x_t^i}). \tag{3.4.6}$$

Finally, the system reliability at time t is expressed as

$$R_t = \Pr\left[\phi(\underline{x_t}) = 1\right]. \tag{3.4.7}$$

The proposed metric is general and can be adapted to different scenarios and configurations of a contact center.

We note that the model defined in Eq. (3.4.7) adopts the description of systems reliability based on the full or complete resolution of a customer's issue. However, the model can be extended to assume that these issues can be partially addressed. Let $\underline{x_t^i}$ express the percentage of issue that is addressed in each stage of the customer service process and be defined as a continuous status variable on the interval $[0, 1]$. The system reliability is then given by

$$R_t = \Pr\left[\phi(\underline{x_t}) \geq \eta\right] \tag{3.4.8}$$

for some reliability threshold level $\eta \in [0, 1]$.

3.5 Performance and Control Management

Analogous to the system architecture, it is desirable to have a visual tool to evaluate the system performance and support fast strategic management decision-making processes (Gama Dessavre et al., 2016; Lieu and Sorby, 2009). The use of appropriate visual tools provides a comprehensive and efficient communication of information on the current state of the system, as well as allowing to assess the what-if scenarios and extraction of actionable insights in an interactive environment (Yu et al., 2011). Therefore, the system should have the ability to provide complete and updated information.

One of the popular tools for the assessment of service quality is referred to as the Seven Basic Tools of Quality Control (see Ishikawa (1985)). The seven tools are the Pareto chart, cause and effect diagram, stratification, check-sheet, histogram, scatter diagram, and control chart.

Given that the proposed reliability metric in equation (3.4.7) is a function of time, control charts seem to be the most appropriate tools. These graphs show how a given indicator varies over time with control limits. Control limits specify the natural variability of the process. Control charts allow the evaluation of statistical stability of the process and quick identification of anomalies. Diagnosing these anomalies is very useful to identify the period that one should inspect the process, in the search for improvement opportunities. For this reason, control charts are found to be powerful tools in systems management and organizational governance.

A control chart is typically composed of a horizontal centerline corresponding to the determined statistic (such as the sample average) of the quality characteristic being monitored, a lower control limit (LCL), and an upper control limit (UCL), which represent the maximum and minimum acceptable variation from the mean, and determine the state of statistical control (Montgomery, 2009). Figure 3.6 presents an example of a control chart, with the center line corresponding to the sample mean. It is a common practice to consider $\pm$ 3 standard deviations from the average. Assuming that the data follows a Normal distribution, 99.7% of sampled observations fall within the three standard deviations from the mean interval. Furthermore, based on Chebyshev's inequality, 89% of sample observations fall within the three standard deviation limits, irrespective of the distribution (Wild, 2000).

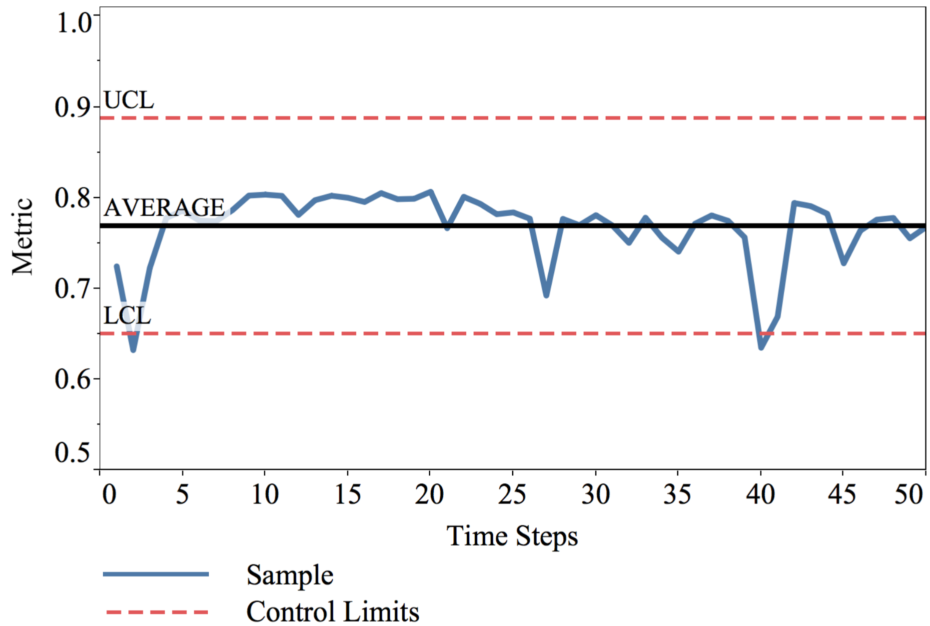


Figure 3.6: Control chart example.

3.6 Example of Insurance Contact Centers

In this section, we use the data from contact centers of a major insurance company to investigate the proposed reliability metric and performance control tool. It serves as an example to discuss the applicability and usefulness of the system architecture and quantitative formulation for contact center reliability. We first describe the data that will be used in this example.

3.6.1 Data Description

The data used for the experimental analysis is from inbound contact centers of a major insurance company in the United States from 1 January 2015 and 31 December 2015. The call centers operate seven days a week and 24 hours per day. The database consists of two layers containing information about customers' interactions with the company.

These layers comprise information recorded by an Interactive Voice Response (IVR) system, and a contact documentation platform (CDP). The IVR system is a self-service platform where customers can solve their demands without speaking with a customer service representative (CSR). The system enables identification and segmentation of callers to be routed to an appropriate CSR depending on their inputs. The CDP comprises information about the queuing process and the service provided by the CSR.

In total, the IVR and CDP platform databases contain 9,951,063 and 12,053,570 data points, respectively. A detailed analysis of this dataset is provided and discussed in Moazeni and Andrade (2018). By combining the attributes of two data sets, it is possible to replicate and gather information from all steps the contact flow diagram presented in Figure 3.5 for each interaction with the contact center.

3.6.2 Empirical Analysis

Let $t \in \{1, 2, 3, \ldots, 252\}$ denote the business days of the year. For the given data, in equations (3.4.1)-(3.4.4) a *contact* refers to a call, and *Communication Channel Option* refers to the IVR system.

The daily system reliability R_t is then calculated using equation (3.4.7). To illustrate how the system performance could be monitored we create a control chart, where the control limits specify ± 3 standard deviations from the mean. Figure 3.7 presents this control chart for the reliability metric over the business days.

Figure 3.7 shows that the reliability corresponding to six business days in 2015 lies outside the control limits. These days are 2 (January 5), 40 (March 2), 89 (May 6), 116 (June 17), 117 (June 18), and 171 (September). A closer investigation of

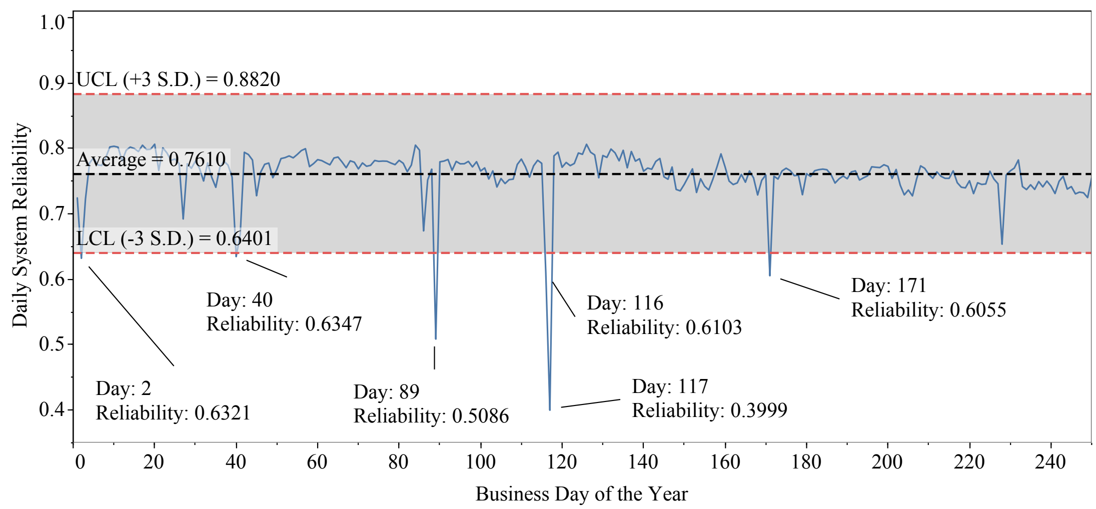


Figure 3.7: Control chart for the insurance contact center system reliability.

the data reveals that the average abandon rate (after being transferred from IVR) was 42% during these days, significantly higher compared to other business days of the year (19%). Furthermore, the volume of transferred calls was, on average, 48% higher than the other days of the year. If the internal demand forecast models of the company were capable of predicting the anomaly, the workforce planning could have been affected. A lower number of agents allocated could not be sufficient to handle the calls with the same quality standards. Consequently, more callers would abandon the system before being served, leading to an unexpected drop in system reliability.

Figure 3.7 also indicates a change in the daily reliability, notably after the 140th business day. While the reliability is prevailing above the average line in the first semester, it is mostly below average through the second term. A closer investigation of the system could look at shorter intervals, such as work shifts, hours, or minutes, or refined by distinguishing the service type, customer type, or contact motivation. A cross-analysis between the reliability and call volume time series can provide insights to improve the current staff scheduling process of the company.

Although different components of the system perform distinct functions and have different failure probabilities, the failure of the parts can potentially compromise the whole system performance. Using data from the insurance company contact centers, Figure 3.8 shows the histograms of the daily probabilities of different steps of the customer support service support to properly execute their functions.

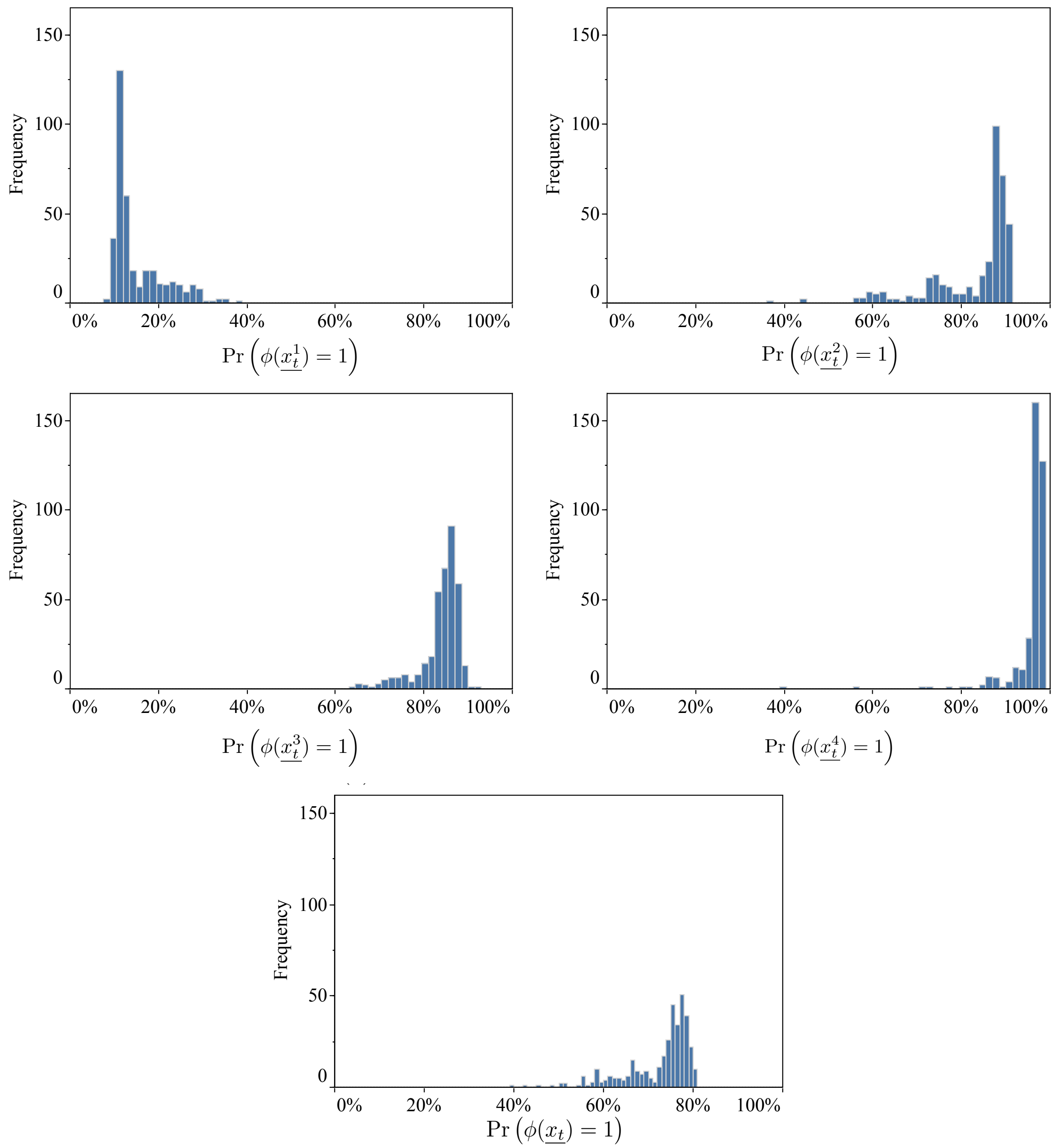


Figure 3.8: Histograms of daily reliability of the parts and the whole contact center operations system.

The distributions of the daily reliability of each phase of the support service process and the overall system reliability present distinct characteristics. For instance, the communication channels have their intrinsic reliability, which can be a function of various factors such as availability, capabilities offered, and effectiveness. These factors can impact on customers' probability to have their problem solved within the contact channel. If the problem is not resolved in the first phase of the service process, different contact distribution processes can still fail to route customers to the appropriate customer service representative. The inefficiency of these processes can also result in the customer's abandonment before receiving proper assistance. Finally, the CSR assistance does not guarantee successful contact resolution. Representative still needs the proper tools and training to effectively solve customers' issues. It is important to have access to customers' data integrated across multiple sources and displayed to agents in a manner to promote an efficient contact resolution. Failures in all these attributes may incur in the inability to solve customers' issues.

The purpose of this example is to illustrate how a visual tool can provide meaningful insights into the entire support service process in contact centers. It also raises important questions regarding current operational processes, such as staff scheduling and demand forecast procedures. Further statistical analyses can be conducted to find out the possible reasons for the downward trend in the level of customer service system reliability as well as for those days where the reliability lies outside of the control limits. Finally, our observations can show potential errors in the service process that have become systematic from a certain point in the year.

3.7 Conclusions and Future Research

Literature from a wide range of disciplines has investigated call and contact centers. It is a multidisciplinary environment that has aroused the interest of communities from operations research, queuing models, marketing, business, and management. This chapter's contribution to the cross-disciplinary operations management literature is twofold: (1) develop a contact center system architecture, and (2) propose a modeling framework for the service system reliability. Using systems thinking tools, the Systemigram and Conceptagon, we map the interfaces among the contact center system components. This comprehensive review of the system enables the development of the contact center operations management system architecture and thereby formulating a model for the service system reliability.

Our study contributes to the cross-disciplinary literature by extending the employment of systems principles to an important sector of the service industry. Furthermore, the developed model for the system reliability opens an opportunity for research on systematic customer service resilience.

In addition, we discuss the importance of having a systemic view of contact centers for organization management, which allows modeling the operation of a complex system. The general architecture of the Contact Center Operations Management System is introduced and followed by a general, adaptable, and time-dependent quantifiable metric for the reliability of the system. The formulation of the proposed reliability model is a result of the systems approach that shows that the contact center system reliability is a function of the ability of its components to properly perform their tasks.

An example using empirical data from a contact center of a major insurance company in the United States is presented to illustrate the applicability of the metric.

The daily reliability of the contact center system was calculated for all days of the year of 2015. Next, we used a control chart as a graphical tool to visually assist the monitoring of processes and their possible variabilities. It allows the evaluation of whether the process is performing as expected and, therefore, contributes to support operational decision-making processes.

Future work is proposed in two folds. First, an investigation to show how the metric can be used to support making better decisions for the system design in a transformation process. For example, a simulation study can be conducted to assess what is the impact of introducing a new cybersecurity system to protect and increase the reliability of the payment system in the company's mobile app on the overall contact center reliability. Other similar analyses can be explored to incorporate different types of uncertainties and decisions. Also, it may be interesting to explore the relationship between the system's reliability and customer satisfaction while considering the financial implications. Second, further research is also needed to fully understand the implications of the management levels in the overall contact center operations management system's reliability.

Chapter 4

A Data-Driven Approach to Predict an Individual Customer's Call Arrival in Multichannel Customer Support Centers

4.1 Introduction

Contact centers provide firms with the opportunity to collect rich customer interaction data from multiple channels. Analyzing such big datasets enables companies to better forecast customers' needs and to improve their business processes by providing customized services and more efficient operations. Accurate predictive models for customer behavior are essential to design and optimize business processes (Kelleher et al., 2015). In particular, call forecasting is considered as one of the three fundamental challenges in the management of call centers (Aksin et al., 2007). The complexity of understanding customer behavior is further compounded in the management of contact centers with data recorded from multiple channels (Neslin et al., 2006).

This chapter is concerned with leveraging multichannel data to predict future telephone queries by an individual customer and to examine the effect of past Web-based contacts by a customer on his future calls to the firm's call centers. Specifically, we develop a feature-based model to predict the likelihood that a customer will call within the next thirty days. The model includes a rich set of features characterizing recency and frequency of past contacts via both telephone and Web channels, contact reason, customer segment, and their interactions.

The context of our analysis is provided by a dataset from call centers of a prominent U.S. insurance firm. The dataset covers about 35 million contact transactions between policyholders in the United States and the insurance firm in 2015. In this chapter, a customer refers to an insurance policyholder. The contact transactions in the database consist of attributes related to the contact date and time, contact reason type and subtype, insurance product type, customer type, and the channel used for the contact. Customers' interactions with the firm via two channels are recorded: Web-based and Telephony. Web-based contacts express communications through the firm's website, while the second class of contact transactions includes inbound calls to the firm's call centers. We translate the transaction-level data into policy-level data and generate new features specifying each policyholder's past contacts.

We first investigate the effect of each attribute in the data on the probability of a policyholder's call arrivals. The averaged hourly volume of calls during a day exhibits different patterns for each contact channel. Similarly, the patterns of the averaged daily volume of calls during a month depends on the contact channel. The highest number of calls by policyholders occur on Mondays while the highest volume of Web contacts occur on Tuesdays. We then present and evaluate a statistical model based on the Lasso method using a rich set of generated features. The out-of-sample performance analysis indicates that the model presents an overall accuracy of 85%, and the test error rate is 15%. The selected features confirms the predictive power of the policyholder's web contacts for his future telephone queries. In addition, frequency and recency of contacts for a set of contact reasons increase the probability of the policyholder's call in the next 30 days. Contacts about Login, Quote Acceptance Form , Billing, eQuote Acceptance Package positively impact the likelihood that the policyholder will call.

An important characteristic of this study is the analysis using the multichannel (both Web and Telephone) data and the nonlinear feature-based model to predict the individual-level customer call arrivals, based on detailed features characterizing the customer's recent contacts such as contact reasons, through any contact channel.

Such a model can provide various practical implications for a firm to improve quality of service through better call arrival estimation.

The rest of the chapter is organized as follows. Section 4.2 summarizes the related literature. Section 4.3 describes our data and a preliminary analysis to support our feature modeling. The model development is explained in Section 4.4. Section 4.5 presents the results and discusses the model performance. The study is concluded in Section 4.6.

4.2 Related Literature

For a review on forecast models in call centers, see Aksin et al. (2007), Gans et al. (2003). The literature on call center predictions primarily focus on estimating the intensity of call arrivals to the call center based on historical telephone queries. For example, in Shen and Huang (2008b), Taylor (2008), Aldor-Noiman et al. (2009), and Ibrahim and L'Ecuyer (2013), linear fixed-effect time-series regression models are developed to predict the intraday call volume. In Antipov and Meade (2002), a linear regression model is developed to predict the daily call volume by loan applicants at a financial services telephone call centre. Our model differs from these studies as we focus on customer-level predictions and includes features characterizing the customer's past contacts via both Web and telephone channels.

The literature on individual-level customer call predictions using data from multiple contact channels is scarce. In Kumar and Telang (2012), linear regression models are developed to identify effective factors influencing the customer's use of different channels. This study shows that Web-based self-service usage leads to a 14% increase in telephone calls. In Jerath et al. (2015), a linear regression model for the total number of queries made by a customer via each channel in a month is investigated. In addition, a Poisson call arrival process whose intensity is defined by an information stock variable is developed. These model do not consider call motivation features and interactions with other attributes. In addition, the model in this chapter relies on the Lasso method using a rich set of features characterizing the customer and the contact. To the best of our knowledge, our model is the first attempt to include the information about previous contacts' motivations by a customer to predict his future telephone queries.

4.3 Contact Center Data

We study a massive dataset recorded from a major U.S. insurance firm between January 1, 2015, and December 31, 2015. The data includes 35, 806, 207 transactions between 7, 463, 600 policyholders and the insurance firm. The transaction-level dataset consists of seven attributes specifying the time and date of the contact, contact reason, policy and account information, channel used for the contact, and the participant involved in the contact. Table 4.1 shows the attributes given in the dataset.

Table 4.1: Description of the original attributes.

Attribute	Description
Event ID	Unique # to identify transactions of the same interaction
Event Time-stamp	Date and time of interaction
Contact Reason Type	First level description of contact motivation
Contact Reason Subtype	Second level description of contact motivation
Channel of Contact	Medium of contact used to interact with the company
Participant Type	Roles/actions to which the contacts are associated with
Policy Number	Unique # to identify customer responsible for the interaction

Each contact event is labeled by a unique Event ID. A contact event may be associated to multiple transactions in the data. This can occur, for example, when the contact consists of several contact reason types or subtypes, each of which leads to a new transaction item in the dataset. Consider a customer who logs into the firm's website regarding several services such as billing services, a policy inquiry, or registration issues. In this case, for each service type, one transaction is recorded in the dataset. For each contact transaction, its date, start time, and end time are specified by the Event Time-stamp attribute.

The Contact Channel attribute specifies the medium or channel used by the participant to interact with the company. Web transactions account for 25,833,965 (72.15%) and telephone transactions constitute 9,972,252 (27.85%) of all transactions in our dataset. Note that the total number of policyholders (unique policy numbers) is 7,463,600.

Table II reports the number of policyholders in our dataset who only used one channel of contact (Web only or Telephone only). The number of policyholders who only used Web is higher than the number of policyholders who only used telephone. Table II also indicates that 21.92% of policyholders in our dataset used both channels at least once in 2015. The number of policyholders per frequency of switch to an alternative contact channel are detailed in Table III for each direction of change.

Table 4.2: Percentage of Policyholders Who Used One Contact Channel.

Contact Channel used	**Number of Policyholders**	[%]
Web Only	3,552,632	47.60%
Telephone Only	2,274,760	30.50%
Web & Telephone	1,636,208	21.90%

It is observed from Table 4.3 that 733, 751 policyholders, who had previously used the Web channel to contact the company, chose the telephone channel for their next contact. However, 1, 466, 620 policyholders changed their contact channel from the telephone channel to the Web channel. Hence, a larger number of policyholders have changed their medium of contact the firm from the telephone channel to Web than from Web to the telephone channel.

Table 4.3: Frequency of Change in the Contact Center Channel by Policyholders.

	Direction of Contact Channel Change			
Change Frequency	**Web to Telephone**	[%]	**Telephone to Web**	[%]
1	580,034	79.05	1,249,635	85.21
2	108,816	14.83	157,390	10.73
3	28,970	3.95	39,665	2.70
4	9,587	1.31	12,135	0.83
5	3,494	0.48	4,392	0.30
6	1,501	0.20	1,862	0.13
7	666	0.09	745	0.05
8	317	0.04	373	0.03
9	179	0.02	189	0.01
10	68	0.01	107	0.01
11	59	0.01	61	0
12	27	0	30	0
13	11	0	11	0
14	2	0	4	0
15	10	0	10	0
16	2	0	3	0
17	4	0	4	0
18	2	0	2	0
19	1	0	1	0
20	0	0	1	0
21	1	0	0	0
Total	**733,751**		**1,466,620**	

Figure 1 depicts the daily volume of transactions in each channel (Web, Telephone) during 2015. This graph shows that the daily Web transaction volumes are consistently higher than the daily Telephone transaction volumes. Fig. 4.1 shows that transaction volumes are typically higher during the weekdays in comparison to weekends. This pattern can be observed for both contact channels. In addition to the day-of-the-week effect in the daily transaction volume in both channels, the volume of daily Web transactions exhibits spikes on the first business day of each month. These days are specified in Fig. 4.1.

Figure 4.2 illustrates the averaged hourly transaction volumes for the two chan-

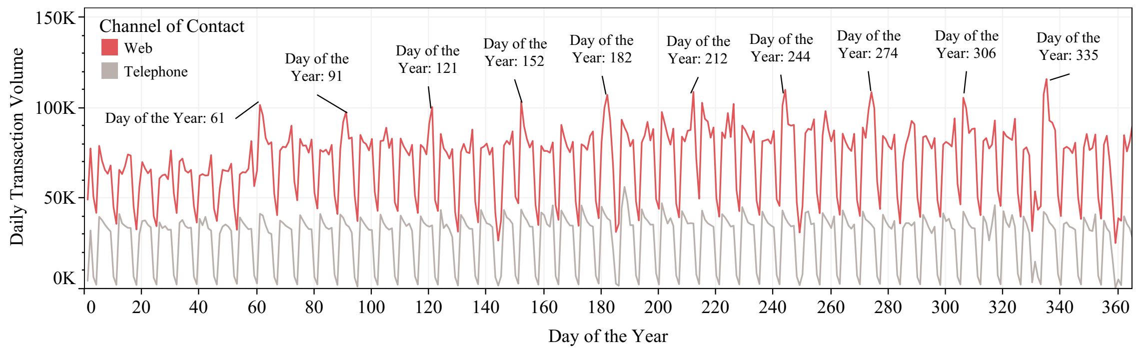


Figure 4.1: Daily transaction volume per channel

nels. This figure indicates that the distribution of the averaged hourly Telephone transaction volume has a convex decreasing tail and significantly drops from 6 pm. In contrast, the Web transaction volume on average does not decrease during evenings.

The Contact Reason Type attribute contains 31 different categories, five of which makes 81.3% of all transactions. These five contact reason categories include Billing (37.9%), Login (17.7%), Policy Inquiry (11.2%), Electronic Message Delivery (10.2%), Policy Change (4.2%). Web transactions are associated to 12 contact reasons. The four contact reasons, Billing (42.1%), Login (24.6%), Electronic Message Delivery (14.2%), Policy Inquiry (9.8%), constitute the reason for 90.7% of the Web transactions.

Figure 4.3 depicts the averaged number of transactions per each business day in the month for each channel (Web and Telephone). The left panel considered transactions with any contact reason and the right panel includes only those transactions with Billing (BIL) as the contact reason. The effect of billing days is prominent in Fig. 4.3 for transactions via the Web channel and with the contact reason Billing.

Telephone transactions are associated to 28 different contact reasons, including Billing (27.0%), Policy Change (15.1%), Policy Inquiry (14.9%), Underwriting (9.1%), Transfer (7.0%), Insurance Document (5.1%). More detailed contact reasons, referred

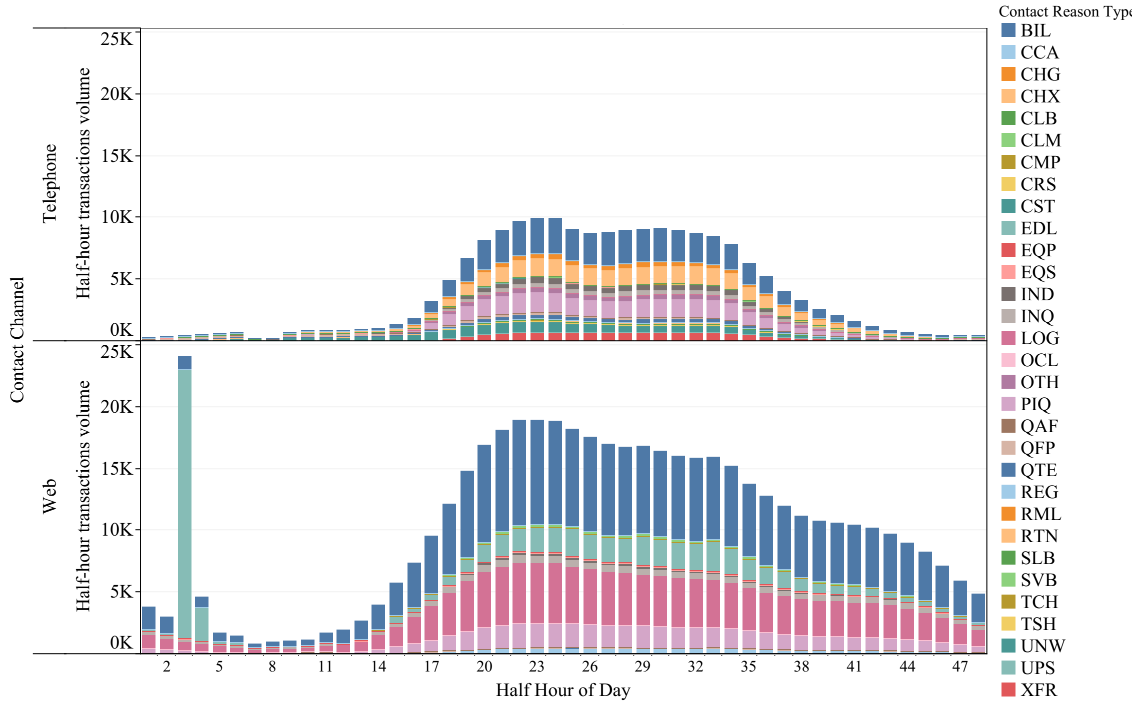


Figure 4.2: Averaged hourly volume of transactions over a day per channel

to as *Contact Reason Subtype*, are also reported with 81 categories. The most frequent contact reason subtypes are Billing Inquiry (10.5%), Underwriting (9.1%), Inquiry Review (6.8%) for Telephone transactions, and Login (24.6%), Web-Billing Payment (13.8%), and Web-Billing Summary (13.3%) for Web transactions.

The *Participant Type* attribute characterizes the participant involved in the interaction. Participant types forWeb transactions are Firm Website Account (69.2%), Internet (24.5%), and Others (6.3%). Participant types for Telephone transactions are Customer (45.4%), Agent (18.2%), and Others (36.4%).

The *Policy Attributes* aim to uniquely specify a policyholder. In our dataset, these attributes, such as policy number, policy form, and billing account number, are masked due to confidentiality reasons. Nevertheless, the uniquely replaced labels and a logic provided by the company allow us to distinguish policyholders. Hence, these attributes enable extracting all contacts made by an individual policyholder from the

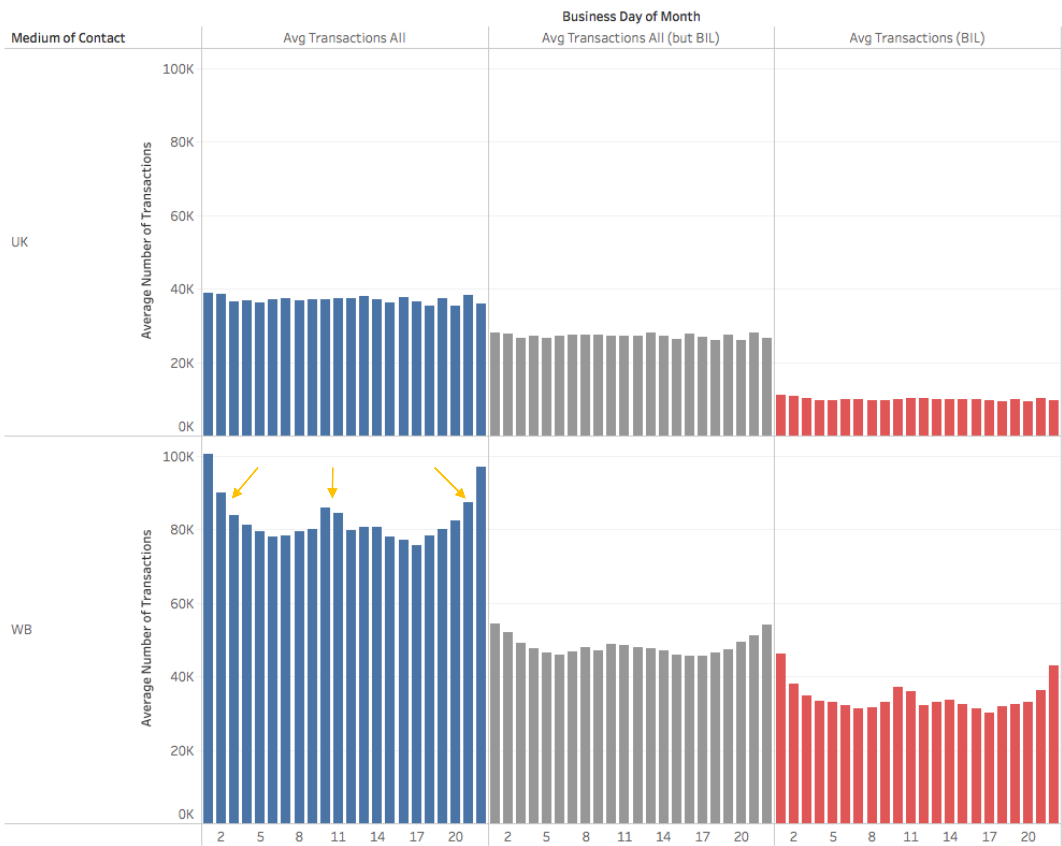


Figure 4.3: Averaged number of transactions per day for each channel

transaction-level data.

4.4 Individual Customer's Call Arrival Prediction

Our goal is to develop a feature-based predictive model to rolling forecast the occurrence of a call event by a policyholder over a set period of time ahead. We choose 30 days as the rolling forecast window, to serve as a proxy to the number of days in the next month. Hence, our prediction problem in this chapter is to forecast whether a policyholder will call the company within the next 30 days, given the characteristics of the customer's segment and his past interactions with the company through either their website or call centers. Let I denote the total number of (unique) policy numbers. For every policy number $i \in \{1, ..., I\}$, define the response variable y_i as

$$y_i|\mathbf{x}_i := \begin{cases} 1 & \text{if policyholder } i \text{ contact the company} \\ & \text{in the next 30 days via Telephone} \\ 0 & \text{otherwise} \end{cases} . \tag{4.4.1}$$

Here, $\mathbf{x}_i$ indicates the vector of features explaining the ith policyholder and his past interactions. The feature vector intents to explain the information available up to the forecasting (snapshot) date. Following the suggestion of our industry partner, we choose September 1, 2015 as the snapshot date. There are two reasons supporting this choice. First, it ensures that there remains a substantial amount of customers' historical data to be used in the present analysis. Second, it secures enough data for model assessment analyses. Given this snapshot date, our study is restricted to policy numbers who had interactions at least once between January 1, 2015 and August 31, 2015. This includes 23, 627, 130 transactions (66.0% of the entire dataset). This corresponds to 5, 645, 280 policyholders, which comprises 75.6% of all policyholders in the original dataset.

Note that the target variable (4.4.1) is defined for each policyholder while we are given contact transaction-level data, it is necessary to infer policyholder contacts from the transaction-level data. This process was completed using the policy attributes and a logic provided by our industry partner as explained in the previous section.

4.4.1 Feature Modeling

The features used in our analysis (170 features in total) are reported in Table IV. Each category of the given attributes in the data is modeled by a binary feature. Participant Type, Contact Reason Type, and Contact Reason Subtype are categorized by 5, 31, and 81 binary features, respectively.

We include five features to represent the date and time of past contacts. The first feature, Weekday, specifies the day of the week for each transaction. The second feature, Holiday, specifies whether the day, when the contact is made, is a U.S. holiday (10 federal holidays and the day after thanksgiving). Three features explain the periods of the day that the last contact occurred. Motivated by Fig. 2 and daily business hours, the day is divided into three segments: 0 am-8 am, 8 am-8 pm, and 8 pm-0 am. Analyzing the last contact of each policyholder before the snapshot date shows that 16.1% of the policyholders made their last contact during 0 am-8 am, 77.2% contacted during 8 am-8 pm, and 6.7% during 8 pm-0 am. Policyholders whose last contact occurred over 0 am-8 am primarily (75.5%) had used the telephone channel. During the business hours (8 am-8 pm), Web contacts represent 66.0% of the total transactions.

Five features are created to capture the channel used used to interact with the company. The first feature, Channel of the Last Contact, identifies whether the last contact before the snapshot day occurred via Web or Telephone. In total, 55.2% of the contacts occurred via Web, while 44.8% occurred via Telephone. The second feature, Used multiple channels, explains whether the policyholder used, at least once, both channels to contact during the past eight months. In total, 18.6% of policyholders used both contact channels at least once in the past eight months. From those policyholders who used the Web channel in the last contact, 89.3% never used Telephone before. From those policyholders whose last contact was made via Telephone, 71.6% never used Web before, while 50.8% used Telephone at least once. Finally, three features are considered to capture whether the policyholder used the same channel for the last contact as the channel in the exact one contact before the last contact, and the direction of the change, whether the policyholder first used Telephone and then used Web, or vice versa. While the policyholders generally tend to use the same contact channel, 9.5% changed the channel between the last two contacts, 75.3% of which first used Web and then used Telephone, while 25.7% first used Telephone and then used Web.

We include three features to capture the billing cycle effect. This is motivated by the observation from Fig. 4.1 that the volume of transactions increases close to billing cycle dates. The feature *Billing Cycle* specifies whether the last contact occurred on the 1, 2, 10, 11, 21, or 22 business day of the month. The feature *Billing Cycle-Begin/End* of the month indicates whether the last contact occurred on the 1, 2, 21, or 22 business day of the month. Finally, the feature *Billing Cycle-Middle* of the month describes whether the last contact occurred on the 10 or 11 business day of the month. For all three features, the last contact can be through Telephone or Web.

Four *Recency* features are generated to represent the effect of recent contacts of the policyholder with the company. Three binary features indicate whether there has been at least one (either a Web or Telephone) contact in the past 1, 7, and 30 days before the snapshot day. The fourth feature, labeled by Number of Days since the Last Contact, represents the number of days since the occurrence of the last contact, as of the snapshot day.

Five *Frequency* features are created to incorporate the effect of previous contact frequency. The first feature expresses the total number of contacts by the policyholder in the entire dataset before the snapshot day. Three features represent the number of contacts in the past 1, 7, and 30 days prior to the given snapshot day. The fifth feature is the average number of days between consecutive contacts by a policyholder in the entire data (last 8 month), since the snapshot day. In the frequency feature subclass descriptions in Table IV, an event refers to a (Web or Telephone) contact event.

We consider cross-class features to include the joint effects of frequency and recency per channel and contact reason. Three channel-recency features are considered: (i) number of days since the last Telephone contact, (ii) number of days since the last contact using the same channel of the previous contact, and (iii) number of days since the last contact using a different channel. A feature is included to indicate the total number of changes (from Web-to-Telephone or Telephone-to-Web) in the contact channel used for consecutive contacts in the dataset (past 8 months). Six channel-frequency features count the number of contact events in the past 1, 7, and 30 days before the snapshot day for each channel, Web and Telephone. Finally, 21 features are generated to incorporate the joint effect of frequency, recency, and contact reason. These features count the number of contact events in the past 1, 7, and 30 days before the snapshot day for the most frequent contact reasons: Billing (`BIL`), Change (`CHX`), E-Delivery (`EDL`), Login (`LOG`), Policy Inquiry (`PIQ`), Registration (`REG`), and

Underwriting (UNW).

4.4.2 Methodology: Logistic Regression

The generated set of 170 features, outlined in Table 4.4, significantly increases the required computational efforts for predictive data analysis. Therefore, feature selection becomes fundamental to reduce dimensionality and training time, and to improve prediction performance (Guyon and Elisseeff, 2003; Yoon et al., 2005; Seltman, 2018; Groll and Tutz, 2014; Leskovec et al., 2014).

Table 4.4: Features description.

Feature Class	Feature Subclass	# of Features
Customer Related Features	Date & Time	5
	Billing Cycle	3
	Channel of Contact	5
	Participant Type	5
Motivation Related Features	Contact Reason Type of Last Event	31
	Contact Reason Subtype of Last Event	81
Recency Related Features	Occurrence Events in the Last Days	3
	Days Since the Last Event	1*
Frequency Related Features	# of Past Events	1*
	# of Events in the Past Days	3*
	Average # of Days Between Events	1*
Cross-class Features	Days Since the Last Event by Channel	3*
	Cumulative # of Changes in Channel	1*
	# Events in the Past Days by Contact Reason	21*
	# Events in the Past Days by Channel	6*

Note. (*) indicates continuous features. All others are binary.

The statistical method consists of a generalized mixed effects logistic regression calibrated by the Lasso method (Tibshirani, 1996, 1997) and a bootstrap approach. The primary goal of the feature selection process is to identify attributes that exert relevant influence on the target variable.

For the binary response variable (4.4.1), consider the model

$$\mathbb{E}\left(y_i|b_i\right) = g^{-1}\left(\mathbf{x}_i^{\top}\boldsymbol{\beta} + b_i\right), \tag{4.4.2}$$

where g is the logit link function $g(t) = \frac{e^t}{e^t+1}$ and $\mathbf{x}_i$ represents the vector of features associated with the i policyholder. This includes all the classes of features presented in Table 4.4. The parameter b_i is a normally distributed random effect with mean 0 and the standard deviation σ. We assume that the individual observations y_i are conditionally independent given b_i. Including bi in the model captures unobservable variation in callers with the same attributes, and reduces the bias in the analysis. The maximum-likelihood estimates (MLEs) of β and σ are determined by solving

$$\max_{\boldsymbol{\beta},\sigma} \log(\mathcal{L}(\boldsymbol{\beta},\sigma)) \tag{4.4.3}$$

where, $\mathcal{L}(\boldsymbol{\beta},\sigma)$ is the joint probability of observing $\{y_i\}_i$ given β and σ.

Typical MLE optimization solution methods often perform poorly when employed for large training datasets and for high dimensional feature spaces (McLachlan and Krishnan, 2008; Karl et al., 2014). It is thus preferable to first rank the features and apply a feature selection methodology. We use the Lasso method (Tibshirani, 1997; Hastie et al., 2001; Schelldorfer et al., 2011), which relies on the ℓ_1 regularized version of problem (4.4.3) given by

$$(\boldsymbol{\beta}^*,\sigma^*) = \arg\min_{\boldsymbol{\beta},\sigma} \{-\log(\mathcal{L}(\boldsymbol{\beta},\sigma)) + \lambda \|\boldsymbol{\beta}\|_1\} \tag{4.4.4}$$

The regularization parameter λ captures the penalty incurred for any non-zero element β_k included in the final model. Therefore, the Lasso penalty function $\|\beta\|_1$ makes β_k to be zero if the kth feature is insignificant. The regularization parameter λ is determined by minimizing the Bayesian Information Criterion (Hastie et al., 2001;

Schwarz et al., 1978) written as

$$\lambda^* = \arg\min_{\lambda} \{-2\log(\mathcal{L}(\boldsymbol{\beta}(\lambda), \sigma(\lambda))) + |\mathcal{A}(\lambda)| \cdot \log N\}, \tag{4.4.5}$$

where $\boldsymbol{\beta}(\lambda)$ and $\sigma(\lambda)$ be the choice of $\boldsymbol{\beta}$ and σ that solves (4.4.4) for given λ, $\mathcal{A}(\lambda) :=\{k : \beta_k(\lambda) \neq 0\}$ be the set of relevant features identified by (4.4.4). Combing this method with the Lasso penalty has demonstrated to perform well in practice, e.g., see Wang et al. (2007); Homrighausen and McDonald (2013).

Feature selection reduces the dimension of the feature space. However, the size of the training data set remains too large to be able to solve problem (4.4.3) computationally efficiently, e.g., see Bradic et al. (2016) and the references therein. To tackle this challenge, we adopt a subsampling procedure, in which random subsamples from the data is used to estimate the model, see Meinshausen and Bühlmann (2010) for details. The bootstrap method draws random subsamples from the data, executes the Lasso method to select significant features for each subsampled dataset. The output is a ranked list of the features according to the frequency of selection across the subsampled datasets. See [21]–[28] for details on the subsampling procedures.

From our initial 170 features, summarized in Table 4.4, 135 features were selected at least once for some subsampled dataset. However, only 14 features were selected in more than 50% of the subsampled datasets. These 14 relevant features are customer related (one), recency related (one), motivation related (five), and cross-class (seven) features. Participant type is the only selected customer related feature in this set of 14 features. No feature from the class of Frequency features was directly selected. However, six of the seven effective cross-class features are defined by the contact frequency. In addition, by taking into account crossclass features, among these 14 selected features, there are eight features related to the contact recency, eight features related to the contact reason, and four features related to the contact channel.

4.5 Results

4.5.1 Predictive Model Estimation

We split the dataset into 70% for training and 30% for testing. Following the method in Hastie et al. (2001); Schelldorfer et al. (2011); Schwarz et al. (1978); Wang et al. (2007); Homrighausen and McDonald (2013); Bradic et al. (2016); Meinshausen and Bühlmann (2010); Kleiner et al. (2014), S = 435 training subsample datasets of size M = 9, 080 are considered. From the initial 170 features, 135 features are selected at least once during the subsampling process. However, only 14 were selected in more than 50% of the sampled datasets. These 14 features are presented in Figure 4.4. With the exception of the class of Frequency Features, other feature classes are identified as relevant: one customer related, one recency related, five motivation related, and seven crossclass features. However, six selected features of the crossclass are defined by the frequency of contacts. The overall accuracy is 85%, and the test error rate is 15%.

Collinearity is a common problem raised in feature selection processes using the Lasso method. The concept of collinearity refers to strong correlations between the independent variables. To assess collinearity among selected features, we compute correlations among these independent variables. Fig. 4.4 illustrates these correlations. It follows from Fig. 4.4 that there is no significant correlation among the selected features.

We then evaluate the correlation between the 14 selected variables and the 156 features that are not selected. Visualizing these correlation coefficients by a

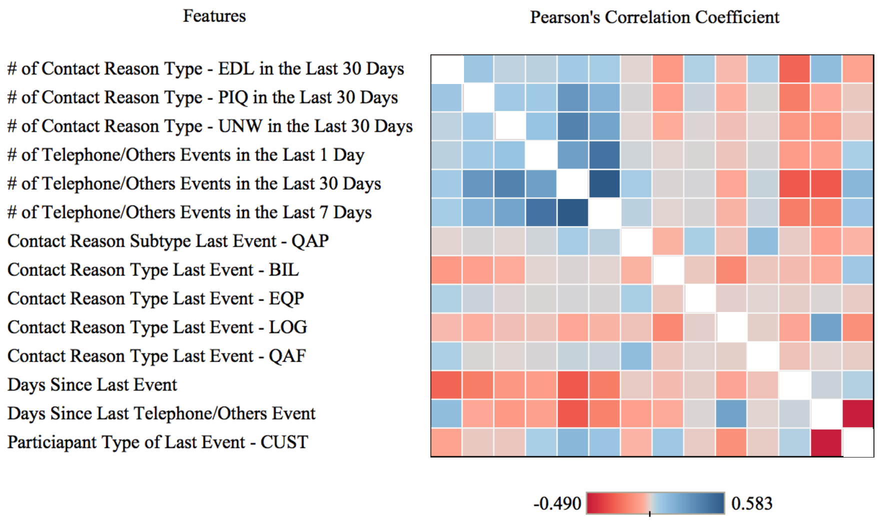


Figure 4.4: Correlation between selected features

correlation matrix similar to Figure 4.4 is challenging due to the large number of variables not selected. Figure 4.5 depicts the correlations between the selected feature, listed on the left, and the unselected features. Each segment of the horizontal bars represents a correlation corresponding to an unselected features.

Figure 4.5 shows that, with few exceptions, correlations are typically low to moderate. Four very strong correlations (> 0.8) are identified for the features: (i) Contact Reason Type of Last Event - *LOG* and Contact Reason Subtype of Last Event - *Login* with the correlation coefficient 1.00, (ii) Contact Reason Subtype of Last Event - *QAP* and Contact Reason Type of Last Event - *QFP* with the correlation coefficient 0.96, (iii) Contact Reason Type of Last Event - *EQP* and Contact Reason Subtype of Last Event - Web *eQAP* coefficient 0.87, and (iv) *Days Since Last Telephone Contact* and *Channel of Last Contact (Web)* coefficient 0.89. The first three correlations are between types and subtypes of contact reasons. The latter reveals that, although the feature *Channel of the Last Contact* was not selected, its impact is captured by the time since the last Telephone interaction, i.e., the longer the time since the last Telephone call, the greater the chances that the last contact has occurred via the Web channel. The contact reason *QFP* (Quote Form Process) refers to actions related to sending and receiving documents from the client.

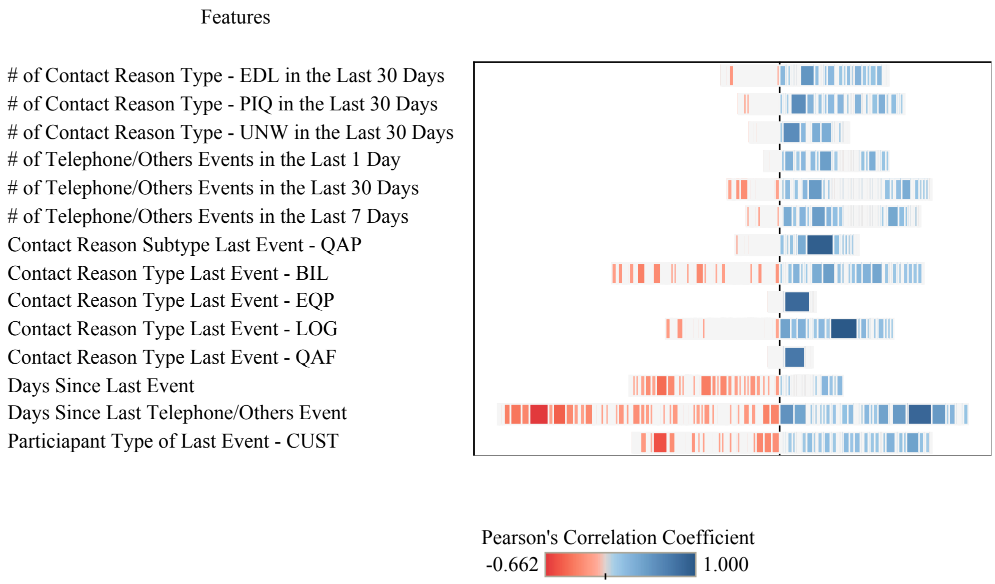


Figure 4.5: Correlation between selected and not feature

Results from the logistic regression model indicate the direction of impact of each effective feature in response variable. Positive coefficients imply the feature contributes to increasing the probability that a policyholder will call the company in the rolling forecast window (next 30 days). The intercept presents a strong negative average coefficient compared to other features. From the 14 selected features, 10 features have a positive impact on the probability of call arrival by a policyholder in the following 30 days. However, the negative impact of the other four features presents higher magnitudes. The sign of the coefficients of the features on the response variable of interest is consistent with the observations of the industry's experts.

The feature *e-delivery* in the Last 30 Days has a significant positive impact on the probability of a call arrival. This suggests that contacts of a policyholder in the past 30 days via the Web channel increases the probability that the policyholder will contact via the telephone channel in the next 30 days. This result indicates the information value of multichannel data to predict future activities in a specific channel. This evinces the effect of the policyholder's Web activities on the probability of his future calls.

The negative coefficient of the feature *Days Since the Last Event* implies that the more recent a policyholder contacted the company, the higher the probability that he will make a telephone query in the next 30 days.

Contact reason types QAF, EQP, and QAP refer to quote acceptance packages sent to customers. The positive influence of features Contact Reason of Last Event - QAF, Contact Reason of Last Event - EQP, Contact Reason of Last Event - QAP, suggests that a follow-up contact with such customers to address questions regarding the new contract will occur in the short-term.

4.5.2 Robustness Analysis

Next, we conduct a sensitivity analysis to investigate how changes to the input data impact the outcome of the predictive model. The analysis consists of changing the following factors:

- Quantity of historical data used for training the model. In addition to using all available data, we also tested only the last 90, and 30 days. Since customers do not usually maintain frequent activity over a period of 90 days, our hypothesis is that the more data is used not necessarily the better is the model performance.

- Inclusion or not of the data of clients that only had one contact in the periods evaluated for the training of the model. Three reasons lead us to believe that these data undermine the predictive power of the model, but that it should be developed a model for them. First, the data shows a significant percentage of customers who only interacted with the company once in the past. Second, this group presents lower probabilities of re-contacting in the next month, and this may increase the model's bias to predict a negative response variable. Finally, this group of samples are similar in the lack of information. As they have no history of other contacts, the amount of information collected on these clients is limited. Our hypothesis is that this compromises the ability to differentiate these people and therefore undermine the modeling performance of the model.

- Threshold used for feature selection. As discussed in Section 5.3, features are selected in the Lasso method after setting a threshold of the percentage of times the feature is included in the subsample models. The rule applied before was a majority vote, where only feature selected in over 50% of subsamples should be included. We will use different threshold (20, 40, and 50) and compare the model performances.

- Month of prediction. We trained and tested the model throughout all months of the year using three months of historical data for each experiment. The objective is to verify if the predictive power of the model is biased in the month chosen for main analysis.

Figure 4.6 shows the result for the tests using 30, 90, or all data available until August for call forecasting in September. In addition to the data quantity variation, the graph also shows the comparison with the application of different thresholds for feature selection and the inclusion or not of policyholders with only one transaction in the past. The comparison is made using the overall classification accuracy, AUC, and precision as performance metrics.

It is noteworthy that the removal of data from policyholders with only one contact in the past tends to improve the classification performance of the model in every scenario, mainly taking the AUC as a reference. Tests in which accuracy is

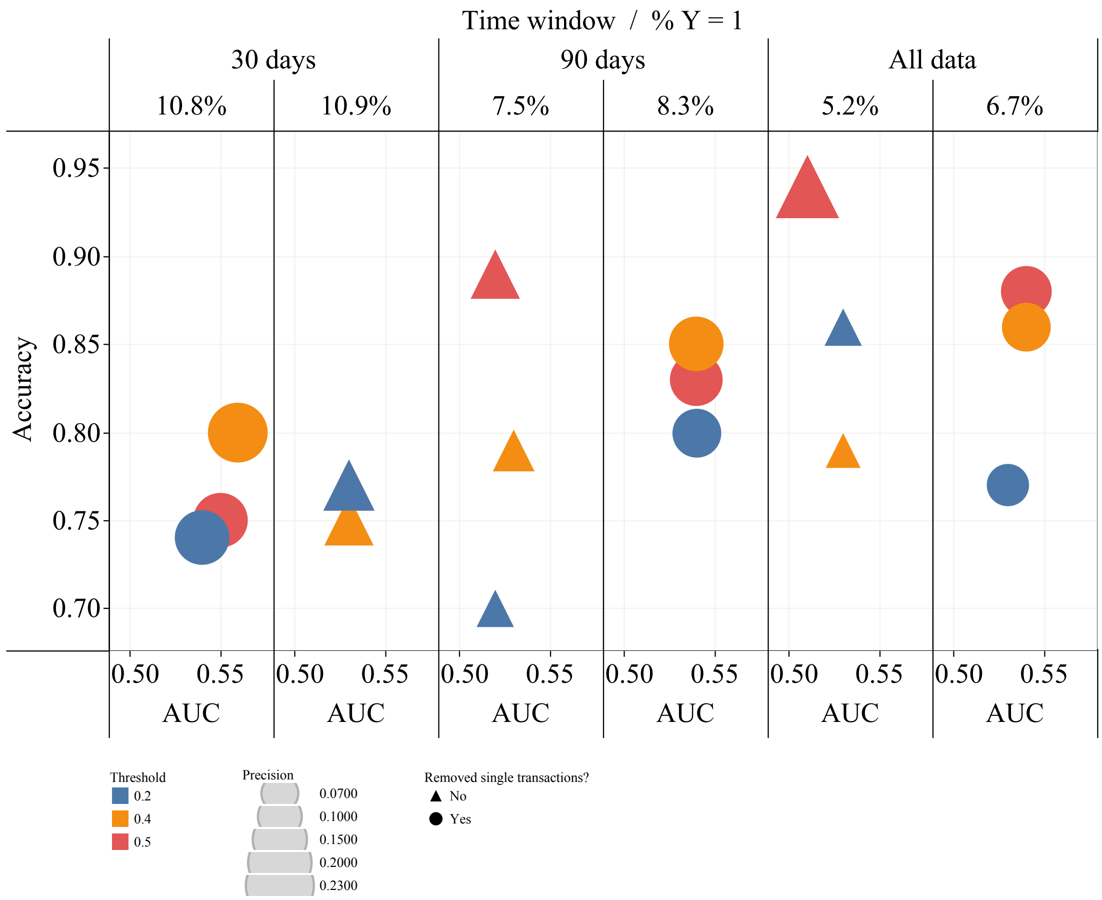


Figure 4.6: Sensitivity analysis. Change of data historical length, feature selection threshold, and inclusion or not of data of policyholders with a single transactions in the past.

above 90% have AUCs close to 50% and the imbalance ratio of the response variable is even greater. Therefore, the models are biased to predict a negative response, increasing the precision and accuracy of the model but without classification power. Furthermore, the use of 20% threshold yields consistently lower results than the other two, while 40% usually performs slightly better than 50% considering the tradeoff between AUC and accuracy.

For the following analysis we rely on a rolling window of 90 days of historical data to predict calls in the months ahead. We start with April so exactly three

Table 4.5: Data description for monthly predictions.

Target Month	Data Size	Imbalance Ratio (IR)	Train Size	Test Size	Subsamples Size (M)	# of Sub-samples (S)
Apr	2,572,439	11.6	1,800,707	771,732	5,666	318
May	2,648,302	12.1	1,853,811	794,491	5,766	322
Jun	2,739,277	11.3	1,917,493	821,784	5,884	326
Jul	2,805,385	11.1	1,963,769	841,616	5,968	329
Aug	2,923,266	12.2	2,046,286	876,980	6,118	334
Sep	3,006,482	12.4	2,104,537	901,945	6,221	338
Oct	2,995,428	12.4	2,096,799	898,629	6,208	338
Nov	2,922,421	12.7	2,045,694	876,727	6,117	334
Dec	2,824,926	11.9	1,977,448	847,478	5,993	330

months of data could be used for training. The data description is presented in Table 4.5. Table 4.6 shows the comparison of the resulting performance of the nine models tested. Finally, Figure 4.7 depicts the average coefficients for each feature for the different monthly prediction models.

Table 4.6: Monthly logistics regression prediction model statistics.

Month	Accuracy	AUC	F1 Score	Precision	Recall
Apr	0.6723	0.5121	0.1347	0.0852	0.3218
May	0.7097	0.5160	0.1315	0.0853	0.2873
Jun	0.6832	0.5160	0.1393	0.0894	0.3163
Jul	0.6705	0.5136	0.1408	0.0899	0.3255
Aug	0.7363	0.5174	0.1297	0.0865	0.2594
Sep	0.7324	0.5188	0.1298	0.0857	0.2677
Oct	0.6999	0.5153	0.1291	0.0824	0.2984
Nov	0.7144	0.5191	0.1296	0.0834	0.2903
Dec	0.6980	0.5176	0.1348	0.0866	0.3041

Although the unbalance ratio varies during the months of the year, the variance is small (the difference between the months with the most and least imbalance ratios is only 1.6 units), and this has no significant impact on the performance of the regression models. The AUC, precision, and F1 Score are have discrete variances over the months. Only classification accuracy and recall present larger variations, usually in opposite directions.

With respect to the selected variables and their coefficients over the months,

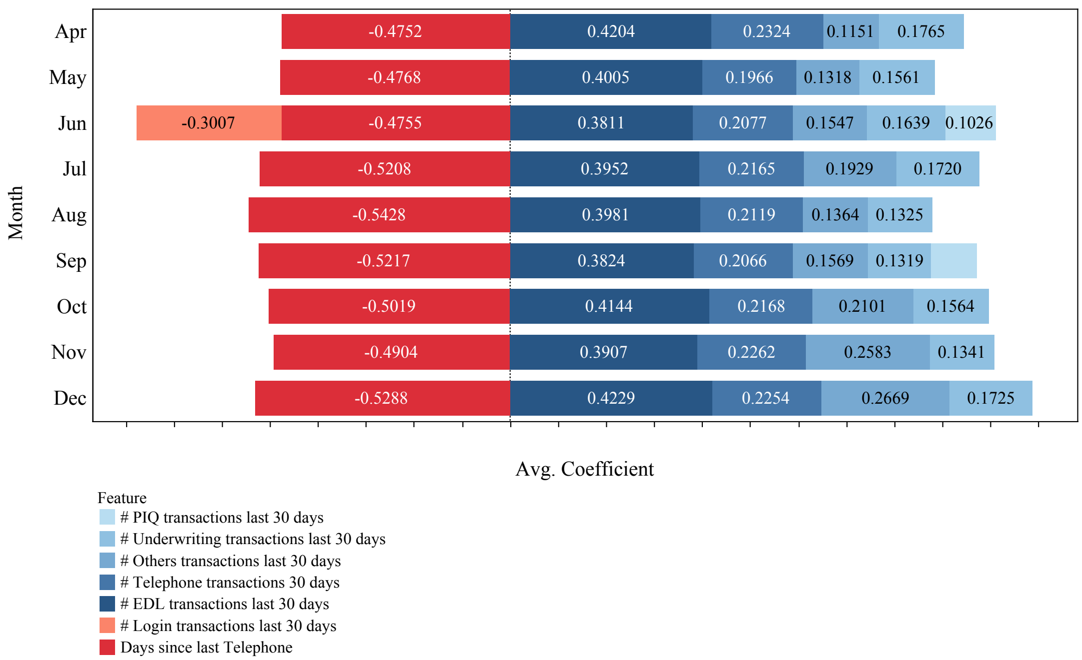


Figure 4.7: Average coefficients for robust analysis of monthly predictions.

there is consistency in both, which demonstrates the robustness of the results. Negative coefficients indicate that the more time has passed since the last of the customer's contact has been made telephone, the smaller their changes to call again in the next month. On the other hand, other variables that account for the frequency of transactions in the last 30 days, contribute to increase the likelihood of a new call in the next 30 days. These results are in aligned with the results presented in the previous section, and are further proof of the robustness of the model.

4.6 Conclusion

This chapter proposes an empirical model to analyze the effectiveness of several characteristics of a policyholder and his previous Web and telephone contacts and their reasons on the probability that he will call in the next 30 days. This analysis is based on a massive data set covering 35 million contacts of policyholders with a prominent

U.S. insurance company during 2015. The metodology relies on the statistical Lasso method and a bootstrapping procedure. A rich set of features including recency and frequency of contacts, contact reason and participant type. An important characteristic of this study is that we focus on policyholder-level call arrival prediction. We find evidence that recent web activities of a policyholder significantly increases the probability that the policyholder would make a telephone call in the next 30 days. Our results provide empirical evidence of the importance of recency and frequency aspects to predict customer behavior, which can be used for customer segmentation in marketing practices. As the time since the last contact of an individual customer increases, the probability of his call in the next 30 days decreases. Frequency of telephone calls also increases the probability of call by the policyholder.

The data-driven approach in this chapter can serve as a useful managerial tool. The modeling approach with the set of selected features enables businesses to identify opportunities to act proactively in an attempt to solve eventual problems of those customers who are more likely to call back in the short term. As an example, making advance outgoing calls to customers who have high calling probabilities can help reduce peak-time calls, leading to cost savings for the firm and providing better customer experience. In addition, the feature analysis can inform managers about the relative efficiency of their customer support channels in handling different call reasons.

Chapter 5

A Data-Driven Approach for Consumer Behavior at Voice Self-Service Platform in Insurance Call Centers

5.1 Introduction

Identifying pattern interactions of policyholders with customer service call center assists business to predict future behavior of customers, and, consequently, better design various service processes. Call centers are valuable sources of customer information (De Ruyter and Wetzels, 2000). Insights gained from analyzing call center data provide a company's business intelligence team with guidance to improve the customer experience and to enhance the quality of service (van der Aa et al., 2015).

To reduce the required workforce in a call center and consequently the associated operational costs, businesses resort to automatic service platforms, in which customers first are referred to a self-service Interactive Voice Response (IVR) system before being transferred to a call center agent. It is thus essential for businesses to ensure that IVR can efficiently handle customer calls in the sense that a call received can be successfully addressed. This not only decreases the likelihood of a customer being transferred to a customer services representative (CSR), but also reduces the risk of ineffective connection with potential customers and losing them.

This chapter studies a dataset from call centers of a major U.S. insurance company. Our research goals are: identify features that exert significant impact on the customer behavior when using the IVR system and develop a predictive model to determine the outcome of a call arriving at the voice self-service platform using a set of attributes related to the customer, policy, and service provided. To the best of our knowledge, it is a topic not yet explored in other call center studies. The literature on call centers typically focuses on forecasting call arrival rate based on time series models for different purposes including customer representatives allocation and other operational planning to meet specific performance metrics (Avramidis et al., 2004; Brown et al., 2005; Weinberg et al., 2007; Shen and Huang, 2008a,b; Taylor, 2008; Shen, 2010; Ibrahim and L'Ecuyer, 2013; Robbins, 2016; Veiga, 2016). However, with the advancement of technology, databases have developed in reliability, speed, and capacity to store detailed customer information. Although extensive research has been carried out on the call arrival rate estimation for staff scheduling purposes, no previous studies consider the outcome of a call arriving at the IVR system as the target variable, neither for a uni- or multi-variate model. Tezcan and Behzad (2012) proposed a stochastic program to determine the number of agents needed to handle customers after being transferred from IVR systems. However, they considered homogeneous customers and call outcome probabilities as a function of the arrival rate and staffing level.

The key properties of this study are: richness of real data (covering past communications between policyholders and the company with detailed information); massive dataset with nearly 10 million transactions; and a predictive model considering customers' profile to determine the likelihood of a transfer from the voice self-service platform to a customer service representative, instead of estimating the call arrival rate, which has been the typical target variable studied in the call center analytics literature.

Our analysis consists of performing data analysis, followed by feature generation, feature selection, and model coefficients estimation. For the last two phases, we used a methodology combining the Lasso (Least Absolute Shrinkage and Selection Operator) method (Tibshirani, 1996) for feature selection, and a bootstrap approach used to overcome the challenge of optimizing the objective function for large datasets. The algorithm is scalable and can be adapted for big data analytics in other business segments.

How can customer transactions data be used to explain the likelihood of a caller being transferred to a CSR? What are the most influential attributes to determine the call outcome? Are location, call motivation, type of service required, or how the customer is interacting with the IVR system relevant factors? Is the system capturing the caller information? Our hypothesis is that some features are more relevant than others to determine the transference probability of a call arriving in the IVR system.

Our findings indicate that distinct channels of contact, customer types, and specific intents are relevant features to determine how likely the call is to be transferred. The location attribute has significant impact on identifying the set of relevant features. However, there is no significant correlation between the customers' regional origin and the selected features. The results can provide managerial insights into policyholders' behavior for decision-makers seeking to make more effective use of customer data and segmentation.

Based on the predicted likelihood, it is possible to draw the profile of customers with higher probabilities to be transfered, and determine the features that contribute to increase the transfer rate. Clustering different groups with similar needs and services can be used to customize both customer service and the company's management, leading to policy and operational recommendations. The determination of the most relevant features to describe the customers can be useful to narrow which areas deserve more attention, either to diagnose potential failures in the system, or to find opportunities for new marketing strategies according to the customer's characteristics. These features may also contribute to improve the IVR system itself, indicating the subjects that demand changes. For example, improving the voice interpretation process could reduce the need to redirect the customers to a CSR, resulting in lower operational costs.

The remainder of the chapter is organized as follows. Section 5.2 explains the details of our dataset. Section 5.3 introduces the features created and describes the data preparation steps. Section 5.4 presents the statistical model and scalable analytics method used. Section 5.5 reports the results and discusses managerial insights. In Section 5.6, we conclude with a summary of the results and directions for future research.

5.2 Data Analysis

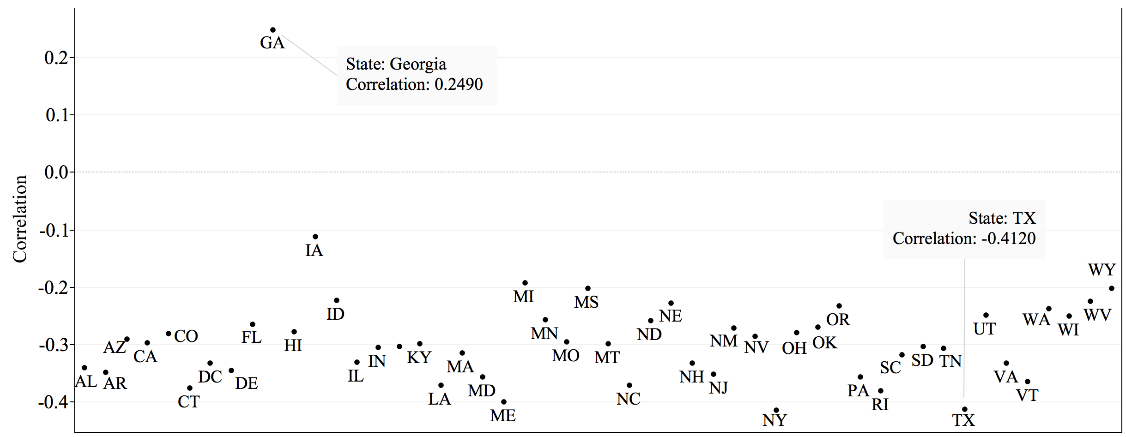
The data was gathered from call centers of a major insurance company in the United States from 2015. The call centers operate 7 days a week and 24 hours per day. The original database contains detailed information about every call arriving at the IVR system. The IVR system is a self-service platform where customers can solve their demands without speaking with a customer service representative. The system enables identification, segmentation and routing of callers to an appropriate CSR depending on their inputs.

The data consists of only inbound calls, and new records enter the database every time a customer calls any of the 1028 toll free numbers. In total, the raw dataset comprises nearly 10 million individual communications, composing a 5.19 GB file. We classify the attributes in the dataset in three categories: call-related features, service-related features, and caller-related features.

Call-Related Features. In total 15 attributes compose this class, including year, quarter, month, week of the year, day of the week, and week start and end dates. There are also call start and end times (in EST), which are used to calculate call duration, and aggregate total number of call arrivals into short time intervals such as hour, half-hour, or 15-minute. These attributes are essential for data analysis for call centers operations especially for call arrivals forecast for different time-horizon for staffing and scheduling purpose (Shen and Huang, 2008b). Ibrahim and L'Ecuyer (2013) summarizes that, to accomplish the correct "balance between quality of service and operational efficiency, call center managers need to determine the appropriate staffing levels in advance based on forecasts of incoming demand", a problem of "resource acquisition" (Aksin et al., 2007), considering problems of scheduling the available pool of agents. Using the call start and end times, it was possible to estimate the IVR call duration for each transaction. Figure 5.1 shows the correlation between the duration of calls originated from different states and the call outcome. The call outcome is defined as transferred and not transferred, taking values 1 and 0, respectively. We note that the correlation is negative for all states, except by Georgia. Georgia also presents the lower transfer rate (58%) among all other states. The negative correlation means that the longer the call length, the higher the likelihood that it will not be transferred. In light of the consistent negative correlation between the duration of the interaction with IVR and transfer outcome, we exclude this attribute from the set of features considered.

Service-Related Feature.. Service related features refer to attributes associated to the insurance product, details of the policy, and other business related elements. Calls are classified by the company in two major categories: service and sale. The company emphasizes the importance to perform parallel analysis for the two types of calls due to its operational and business characteristics differences. The two segments requires different type of support and CSRs with different skill sets to handle the calls. These capabilities can be used when routing calls (Gans et al., 2003). Additionally, there are attributes related to the policy number and policy type, two related to the

Figure 5.1: Correlation between call outcome and call duration per state.

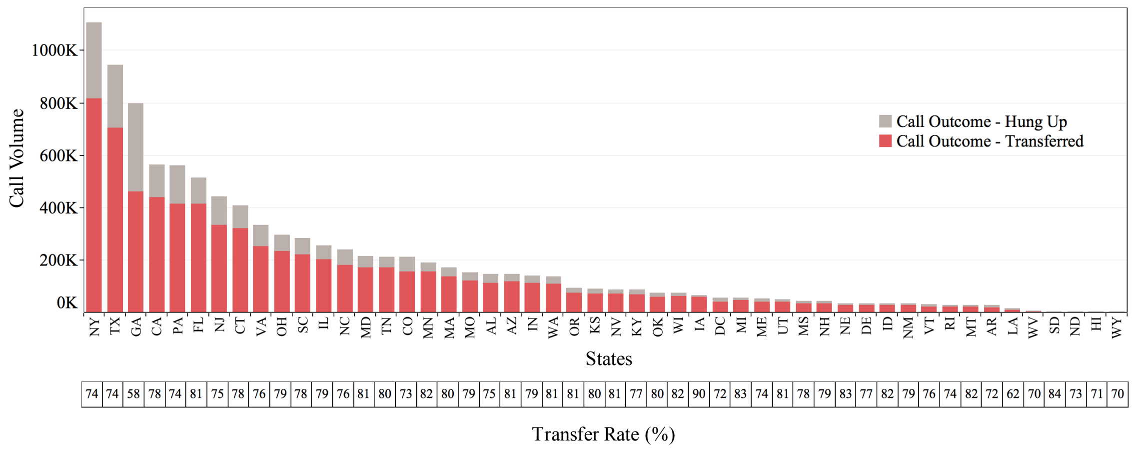


billing account number, and others presenting details of the payment method and the channel which the customer interacts with the company. Due to security and privacy restrictions, the attributes related to the policy number and billing account number were masked, making it difficult for our analysis to reach the policy level. This type of information would enable us to investigate the historical behavior of each consumer. Some factors that could be obtained from this data included are: frequency and recency of contact, frequency of transference, typical call intent, etc.

Customer-Related Features. The calls are classified according to customer information features reflecting the customer type, and customer's location. We consider the state from which the call originates. Figure 5.2 shows the number of calls by state broken down by the call outcome. The box in the bottom of the figure indicates the transfer rate, or proportion of calls with a transfer outcome, for each state. The original data set contains recorded communications from the 49 states and the District of Columbia, to which, in this chapter, we also refer as a state.

Both *Service* and *Caller-Related* classes of features are essential to understand service demand characteristics. Dividing the business into sub-groups with similar needs and some type of shared characteristics can help customize both customer

Figure 5.2: Call outcome and IVR transfer rate per state.



service and the company's management. Tsiptsis and Chorianopoulos (2011) defines segmentation "as the process of dividing the customer base into distinct and internally homogeneous groups in order to develop differentiated marketing strategies according to their characteristics". For instance, knowing the call demand by customer type, business or policy type, can be used to plan, scale and allocate representative with the right skills for the specific type of service, which is crucial for service quality and operational efficiency.

Table 5.1 presents how the set of features that will be explored are classified, grouped, described, and modeled to serve as input for the regression model.

5.3 Features Modeling

To incorporate the categorical attributes as independent variables into the regression model and yield interpretable coefficients, the described attributes are modeled as binary features. This setup enables associating a coefficient to each level of the categorical variables and direct evaluate their impact on the response variable.

In addition to the binary features provided explicitly from the original database

Table 5.1: Features descriptions for IVR call outcome predictive model.

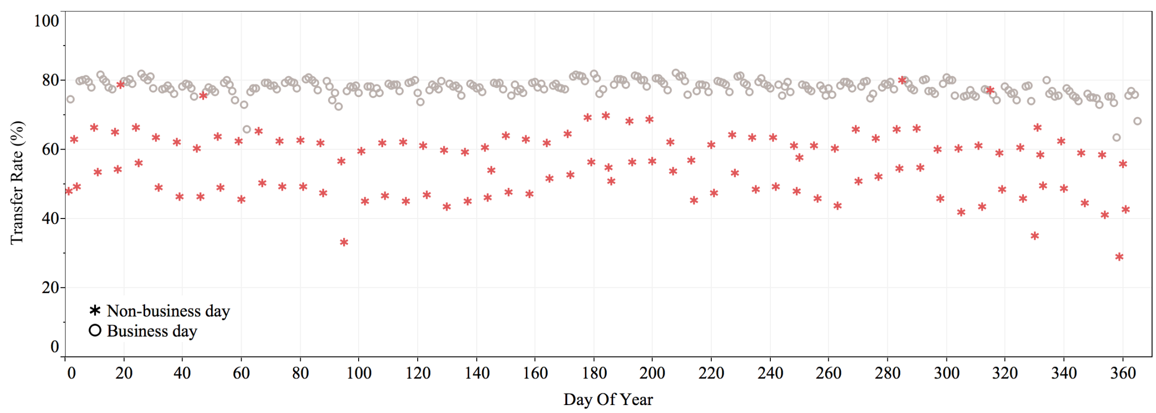
Class of Features	**Group of Features**	**Description**	**# of Features**
Call Related Features	Time	Hour of the Day {1, 2, ..., 24}	24
	Date	Day of the Week {Monday, ..., Sunday}	7
		Is Business Day {Yes, No}	1
		Is Billing Cycle {Yes, No}	1
Service Related Features	Business Unit	Company's classification {Service, Sales}	1
	Channel	{Affinity Sales, Billing, Direct Sales Internet, Direct Sales Mass Media..., Service}	8
	Policy Type	{Apartments, Automobile, Dwelling Fire, Homeowners, Not Specified, ..., Personal Liability}	13
Caller Related Features	Caller Type	{Customer, Potential Customer, Agent, Third Party}	5
	Caller Intent	{Billing, Claims, Mailing Address, Policy Inquiry} {Technical support, ..., No intent}	42
	Payment Attempted	{Yes, No}	1
	Payment Method	if Payment Attempted:{Checking, Savings, Visa, Mastercard}	4
	Interaction Modality	{Dual Tone Multi Frequency (DTMF), Speech-enabled (SPCH), DTMF & SPCH, Not Specified}	4
Exogenous Factor	*Weather*	*State Daily Average Temperature*	1
Cross-class	*Interaction Terms*	*Customer Type × Policy Type*	65
		Business Unit × Interaction Modality	4

Note. Features in *italic* were generated and do not correspond to original attributes in the raw data.

attributes, we create two features describing whether the call event occurs on a business day or a non-business day (*is business day*), and whether the call take place within the days of the billing cycle (*is billing cycle*). Non-business days include Saturdays, Sundays, and major national holidays: New Years Day, Martin Luther King Day, Presidents Day, Memorial Day, Independence Day (observed), Labor Day, Columbus Day, Veterans Day, Thanksgiving, Day after Thanksgiving, and Christmas Day. The second feature was generated to address the industry experts suggestion that call volume typically rises close to billing and payment due dates. Although the company provides flexible billing and payment due dates for the customers, the default is in the beginning and middle of the month. Therefore, we explicitly include a feature that identifies whether the call is placed in any of the following business days of the month: 1, 2, 10, 11, 21, 22.

Figure 5.3 illustrates the differences between the call transfer rates for business and non-business days along the days of the year. We observe that business days have a higher and more constant transfer rate in comparison to non-business days. Non-business days present a higher variability, ranging from around 40% up to 70%. There is also a visible discrepancy between the Saturdays and Sundays time series, with the former consistently presenting higher levels.

Figure 5.3: IVR transfer rate for business and non-business days.



Furthermore, to address an important insurance business concern, the weather condition (Karl et al., 2013; Changnon et al., 1997), we include the average temperature of the state where the call was placed. Due to the lack of zip-code information the state attribute is used to track the origin of the call. The average temperature is estimated by averaging the recorded temperatures of every available station of the states at the call date and time. The temperature data was gathered from the National Oceanic and Atmospheric Administration's National Centers for Environmental Information (NOAA-NCEI, 2016).

In contrast to the other features, the average temperature takes continuous values. The use of different units and dispersions between variables can make it difficult to interpret and compare regression coefficients and be detrimental to the variable selection process (Gelman, 2008; Tibshirani, 1997). Feature normalization seeks to minimize this problem, preventing one feature from overlapping the others and avoid the learning process from stagnating. In the present study we will make use of Z-score Standardization, which normalizes the data around the mean and standard deviation, with average 0 and variance 1.

5.4 Statistical Method

The statistical method consists of feature selection and model coefficients estimation by combining a bootstrap technique for a mixed effect logistic regression model. The primary goal of our feature selection process is to identify factors that exert relevant influence on the target variable. Moreover, selecting a subset of effective features can also facilitate the data and model understanding, reduce dimensionality, reduce training time, and improve prediction performance (Guyon and Elisseeff, 2003). Mixed models are extensively "used to model correlated and clustered responses" (Groll and Tutz, 2014), which provides a general, flexible approach in these situations, because it allows a wide variety of correlation patterns (or variance-covariance structures) to be explicitly modeled" (Seltman, 2018). The specific predictive problem under consideration is to build one model for each U.S. state to determine whether a call arriving at the IVR system will be transferred to a CSR or not, using provided historical data about previous customers' communications with the company.

Let $i \in \{1, ..., I\}$ denote the ith call with I being the total number of calls. The binary response variable y_i^s denotes the result of call i originated in state s, where s is one of the 49 U.S. states or District of Columbia. Then we define

$$y_i^s = \begin{cases} 1 & \text{if the outcome of the IVR transaction call } i \text{ is transfer} \\ 0 & \text{otherwise} \end{cases}, \tag{5.4.1}$$

We assume that

$$\mathbb{E}\left(y_i^s|b_i^s\right) = g^{-1}\left(\left(\mathbf{x}_i^s\right)^T \boldsymbol{\beta}^s + b_i^s\right), \tag{5.4.2}$$

where g denotes the logit function, the covariate vector $\mathbf{x}_i^s$ represents the fixed effects of features associated with the ith policy number (Seltman, 2018), and b_i^s is the random effect is the parameter (Laird and Ware, 1982) estimated as the model variance and interpreted as the variability of "personal" coefficients from the mean fixed effects coefficients (Seltman, 2018). We assume $b_i^s \sim \mathcal{N}\left(0, (\sigma^s)^2\right)$, where $(\sigma^s)^2$ represents random variation between y_i^s observations. Given the covariate vector and the random effect parameter, the joint probability of observing y_i^s can be written as

$$\mathcal{L}(\boldsymbol{\beta}^s, \sigma^s) = \prod_{i=1}^{I} \int_{-\infty}^{\infty} \left(\frac{e^{\left(\mathbf{x}_i^s\right)^T \boldsymbol{\beta}^s + b_i^s}}{1 + e^{\left(\mathbf{x}_i^s\right)^T \boldsymbol{\beta}^s + b_i^s}}\right)^{y_i^s} \left(\frac{1}{1 + e^{\left(\mathbf{x}_i^s\right)^T \boldsymbol{\beta}^s + b_i^s}}\right)^{1-y_i^s} \frac{e^{-\left(b_i^s\right)^2/2(\sigma^s)^2}}{\sqrt{2\pi\left(\sigma^s\right)^2}} db_i^s \tag{5.4.3}$$

for which maximum-likelihood estimates (MLEs) of $\boldsymbol{\beta}^s$ and σ^s are the solution of

$$\max_{\boldsymbol{\beta}^s, \sigma^s} \log \mathcal{L}(\boldsymbol{\beta}^s, \sigma^s) \tag{5.4.4}$$

For high-dimension data, as it is the case in the present chapter, it is typical to use the Lasso regularization version (Tibshirani, 1996; Hastie et al., 2001; Schelldorfer et al., 2011) of problem (5.4.4), which combines the MLE estimation and ℓ_1 penalty function $\|\boldsymbol{\beta}^s\|_1 = \sum_j |\beta_k^s|$ that have built-in feature selection. We then let λ be the regularization parameter, and problem 5.4.4 is then substitute by

$$(\boldsymbol{\beta}^{s*}, \sigma^{s*}) = \arg\min_{\boldsymbol{\beta}^s, \sigma^s} \left\{-\log \mathcal{L}(\boldsymbol{\beta}^s, \sigma^s) + \lambda \|\boldsymbol{\beta}^s\|_1\right\} \tag{5.4.5}$$

To determine the regularization parameter λ, we solve the following equation

5.4.6 by choosing λ that minimize the Bayesian Information Criterion (BIC) (Schwarz et al., 1978; Hastie et al., 2001):

$$\lambda^* = \arg\min_{\lambda} \left\{-2\log\mathcal{L}(\boldsymbol{\beta}^s(\lambda), \sigma^s(\lambda)) + |\mathcal{A}^s(\lambda)| \cdot \log N\right\}, \tag{5.4.6}$$

where $\boldsymbol{\beta}^s(\lambda)$ and $\sigma^s(\lambda)$ be the choice of $\boldsymbol{\beta}^s$ and σ^s that solves (5.4.5) for given λ, $\mathcal{A}^s(\lambda) = \{k : \beta_k^s(\lambda) \neq 0\}$ be the set of relevant features identified by (5.4.5).

In order to obtain a stable final subset of relevant features (Bi et al., 2003), address the difficulty to perform the regression on a high dimension data set, and reduce the bias of using a single random subsample of the data (Meinshausen and Bühlmann, 2010), we adopt a bootstrap procedure for feature selection and coefficients estimation. By drawing M^s subsamples, of size U^s, from the state data, we obtain M^s different solutions λ_s^* and distinct $\mathcal{A}^s(\lambda_s^*)$ (Meinshausen and Bühlmann, 2010). Let N^s be the total number of communications of state s data, the size U^s is set to $U^s = (N^s)^\gamma$ for $0.5 \leq \gamma < 1$ following (Kleiner et al., 2014) study recommendation. Finally, $M^s = \frac{N}{U^s}$ (Hastie et al., 2001). The union of the M^s subsets constitute the final active set. We consider that the kth feature is more likely to be significant as it is more frequently selected by the subsets. The final set of selected features is determined setting a percentage threshold of the times the features were selected in the M^s Lasso models. The kth feature is included in the final model if

$$\frac{1}{M^s}\sum_{m=1}^{M^s} 1_{k\in\mathcal{A}^s(\lambda^*)} \geq \alpha, \tag{5.4.7}$$

where k is incorporated if it is selected in the proportion of samples that exceeds a threshold α, where $0.5 \leq \alpha < 1$.

In possession of the final active set of features, the subsampling process is repeated to compute 4.4.3 using only solely features of $\mathcal{A}^s(\lambda)$ and estimate the regression coefficients for each of them. The final coefficients are calculated by averaging the coefficients obtained for each M subsample:

$$\bar{\beta}_k^s = \frac{1}{M^s} \sum_{m=1}^{M^s} \hat{\beta}_{k,m}^s \tag{5.4.8}$$

The standard error, t-statistic, and the corresponding p-value are calculated using standard procedures.

5.5 Results

5.5.1 Feature Selection

The first set of analysis examines the feature selection results for each state. 27 features among the initial 181 were selected by at least one state model. Figure 5.4 presents the number of states that included each feature broken down by the direction of impact in the likelihood of a call to be transferred, and the respective average coefficients among the models in which the feature is included. A positive coefficient suggests that the feature contributes to increase the likelihood of a customer to be transferred. Features are ranked from the most to the least frequently selected among the state models. Table 5.2 shows the number of features selected by each state. The set of features that constitute each state model varies from state to state.

The results outlined in Figure 5.4 and 5.2 indicate the caller location (State) plays a key role on determining the contribution of these variables on the call outcome. Moreover, features from different classes are identified to be relevant. However, only 7 features are selected in more than 25 states. We observe that average temperature is not selected among the top effective features. From these 27 features, 7 are call

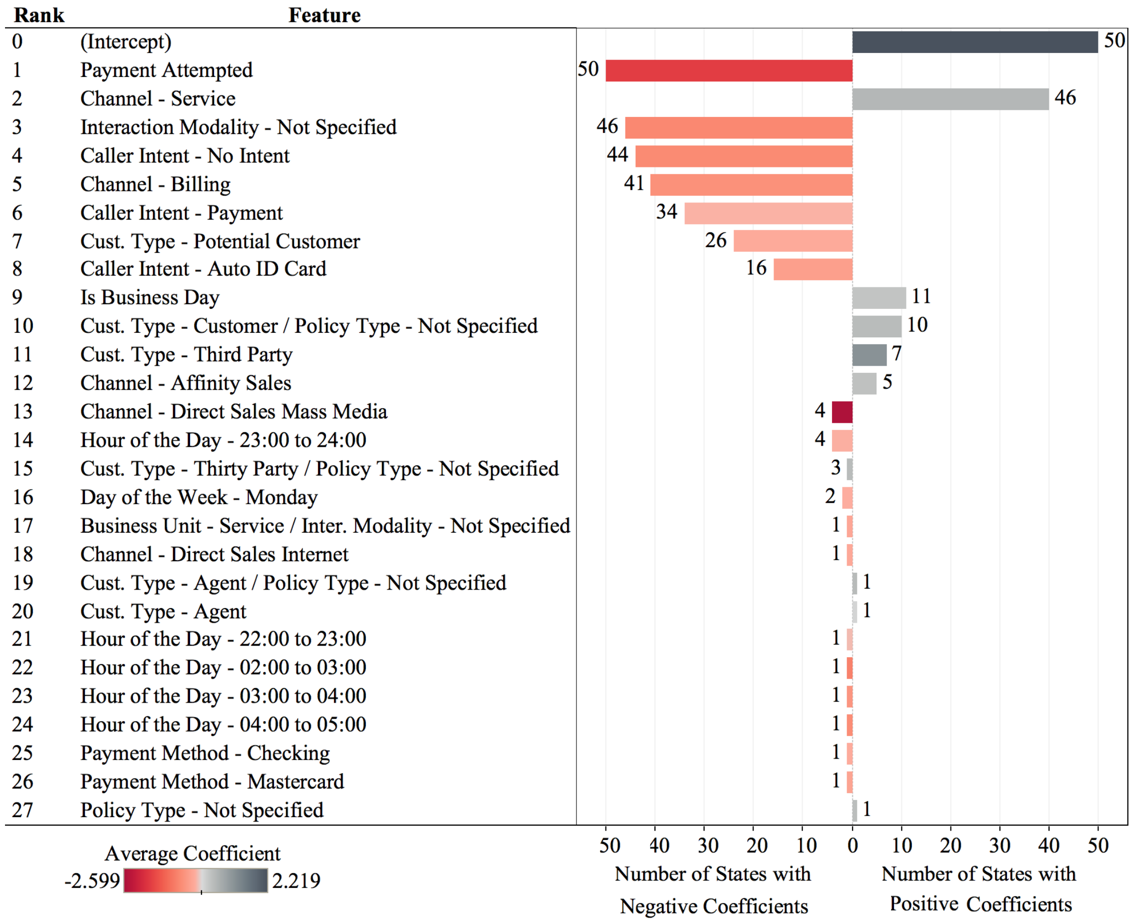


Figure 5.4: Features that are selected for at least one state, direction of impact, average of positive and negative coefficients.

related, 6 are service related, 10 are related to the caller, and 4 are cross-class features.

In addition to presenting each of the selected variables and discussing the managerial insights, we will make a complementary analysis of correlations between the selected variables and those not selected in each one of the models. We report only relevant correlations with Pearson correlation coefficients greater than 0.50. Exceptional cases are highlighted. The objective is to identify variables that, even if not directly selected, have an influence on the response variable but, due to a high correlation with the most effective variables, see its impact diminished.

Managerial Insights. We begin by discussing the results for caller related features. The results suggest that payment related features such as *Payment Attempted*, *Caller Intent - Payment*, and *Payment Methods* exerts consistent negative impacts on call likelihood to be transferred from the IVR system to a CSR. Calls with no payment attempt corresponded to 89.2% of total call volume for the 50 states and presented a high transfer rate of 84%. In contrast, in calls where there was an attempt to pay, the transfer rates drop considerably to 5.1%. Additionally, 98.9% of the non-transferred calls were contained. *Caller Intent - Payment* is selected by 34 states comprising 84.1% of total dataset. Calls with *Caller Intent - Payment* added up to 16.8%, presenting a low transfer rate of 29.1% and a also a high contained rate of 89.9%. In summary, these results indicate that the current IVR system is efficient to solve these types of demand.

The correlation analysis shows that *Payment Attempted* typically present Pearson correlation coefficients greater than 0.5 with *Payment Method* features, specially *Checking* (in 45 out of 50 states) and *Vista* (in 39 out of 50 states), that were not selected by the state models. In a similar vein, *Caller Intent - Payment* is also positive correlated with the two payment methods. When the *Caller Intent - Payment* is not a selected feature, it also presented a positive correlation with *Channel - Billing* in 8 states. Both cases are in line with the fact that to have a payment method one needs to make a payment attempt. At the same time, is logical to have *Payment* as the contact motivation, which occurs in 96.3% of *Payment Attempted* transactions.

Interestingly, *Caller Type - Potential Customer* feature is a relevant feature for 26 states. It exerts negative effects in 25 states, whereas in the state of Missouri it has a positive influence with an estimated average coefficient of 0.04. Among the 25 states for which *Potential Customer* has a negative impact, the average negative coefficient of -0.30 is significantly higher in comparison to the only positive coefficient. The 26 states comprised 81.74% of the country's total call volume, and *Potential Customer* calls corresponded to 6.64% transactions with 47.3% transfer rate. The transfer rate for 25 states (for which *Potential Customer* has a negative impact) and Missouri (positive impact) were, in fact, very distinct. The transfer rate for the first group was 91%, while for Missouri was 68.1%. However, from the not transferred calls for the 26 states, only 0.03% of the calls were contained. If we expand these analysis for the entire country, we observe that 50.0% of *Potential Customer*'s calls outcome were transferred. For those which were not transferred, 96.4% of the callers had no specific intent. For the ones transferred, 48.1% also had no specific intent related. Finally, sales intent accounted for 26.0% of the potential customers transferred calls. These suggests that the current IVR system is not effective or do not offer capabilities to handle potential customers. The immediate insights from these results is that the IVR system should be improved to better identify potential customers and facilitate the redirection to specialized CSR. The company may also consider creating another direct channel for this target audience, or offer more information and support in other mediums of contact (e.g. the company web site) to minimize number of calls with no clear or vague motivation to the call centers.

By investigating the correlation between *Caller Type - Potential Customer* and not selected features for the 26 states, we identified that *Caller Type - Potential Customer* is consistently positive correlated with the cross-class feature *Caller Type - Potential Customer/Pol. Type – Not Specified*. In fact, 99.9% of *Caller Type - Potential Customer* calls had no police type specified. Assuming that a potential customer does not hold a policy (only 0.34% of potential customer calls are placed by existing customers interested in a new policy), not having a specified policy type is the direct consequence. Since both features practically match and, therefore, carry the same information, it is reasonable to have only one selected by the state models. Correlations ranging from 0.50 to 0.68 were also estimated between *Caller Type*

- *Potential Customer* and *Channel – Direct Sales Broadmarket*, in 7 states (Mississippi, Ohio, Wisconsin, Connecticut, Colorado, Pennsylvania, and Washington), and between *Caller Type - Potential Customer* and *Channel – Direct Sales Internet* in 4 states (Georgia, Illinois, Nevada, and New York. Interesting, *Channel – Direct Sales Broadmarket* was not selected by any state and *Channel – Direct Sales Internet* is only included in the Michigan state model. *Channel – Direct Sales Broadmarket* is the most used channel by *Potential Customer* from the 7 states (66.9% of the calls) with 62.1% transfer rate. *Channel – Direct Sales Internet* is also the most used channel by *Potential Customer* from the other 4 states, consisting of 48.8% of the calls, however, presented a lower transfer rate of 31.2%. Although the two scenarios differ substantially on the call transfer rate, both did not present any contained call. Additionally, in both circumstances non-transferred and not contained calls are mostly related to no specific intent (99.9% and 88.2% calls, in the first and second scenarios respectively). The lower transfer rate associated with the non-solution of call inquiries, which are also not effectively captured by the current IVR system, corroborates with our previous evaluation and recommendations. Improve the customer service platform to effectively handle potential customer calls can potentially lead to new sales and revenues.

As shown in Figure 5.4, *Caller Type - Third Party* is an important factor selected by 7 states (Arizona, California, Indiana, Michigan, Ohio, Pennsylvania, South Carolina, and Texas) and has a strong positive impact (second highest average coefficient in magnitude among all features) on the call transference likelihood. In 2015, 6.68% ($200,062$ calls out of $2,994,569$ calls) from these seven states were placed by third party customers, and 98.16% of these calls were transferred. Furthermore, only one non-transferred call was contained. Among the transferred calls, 41.3% had no intent specified and 24.8% were policy inquiries. Similar to the *Potential*

Customers case, the IVR system might not have been able to capture most of the intents or there is no capability offered. It would be useful to future review in detail the reasons of these calls with business precision. Possible actions could be towards the creation of a specific channel for this type of customer with a personalized IVR system more prepared to handle this public or provide similar capabilities in of other mediums of contact to help contain the demand.

The correlation analysis between *Caller Type - Third Party* and non-selected features yields interesting results. Positive correlations with coefficients ranging from 0.55 and 0.57 were observed with *Caller Type - Third Party/Policy Type - Homeowners* in 4 states (Arizona, California, Ohio and South Carolina) and moderate correlation (0.47) in Michigan. The positive correlation with a interaction feature including the same caller type is expected, but the appearance of *Policy Type - Homeowners* is worth noticing, since it was not selected by any state model. Narrowing the analysis of third party calls with this type of policy, which comprise 32%, reveals an even more significant call transfer rate of 99.9% in these states. Transferred calls intents are commonly classified as *Policy Inquiry* (45.1%), *Billing* (19.6%), and *Proof of Insurance* requests (12.6%). Remarkably, only 77 calls (or 0.22%) presented *Caller Intent - No Intent*, which is a positive aspect. Despite the high transfer rate and the only non-transferred call contained (and improvements should certainly be made to enhance the self-service capabilities for this type of demand), this analysis provides evidence that the current IVR system can at least effectively identify the contact motivations for this audience.

The second class with more features selected among all states consists of service related features. The second most frequent selected feature is *Channel - Service*. This feature was selected by 46 states, which accounted for 96.9% of the country's total call volume. *Channel - Service* was the most frequent used channel in the 46 states comprising 5,304,683 calls (or 56.7%) with 89.9% overall transfer rate. From the 534,050 non-transferred calls, only 12.8% were contained within the IVR system, which justify the fact that, apart from Indiana, Oklahoma, and Texas, *Channel – Service* presented positive impact on the target variable. The average positive coefficient is 0.53, two orders of magnitude higher than the average negative coefficient of -0.0061.

Surprisingly, the *Business Unit* was never selected as an effective feature. The finding was unexpected since it was emphasized its importance in all operational analysis carried out by the company. This rather contradictory result may be due to significant correlations between the *Business Unit* and other selected features which somehow inherit and distribute its impact on the call transfer likelihood. Absolute values of Pearson correlation coefficients are rarely higher than 0.5. This level is only surpassed in the following situations: correlations with *Channel - Service* in the states of Michigan (0.51), Georgia (0.52), and South Dakota (0.64); correlations with *Caller Type - Potential Customer* in the states of Pennsylvania (-0.50), New Jersey (-0.50), and Georgia (-0.67); and finally with *Interaction Modality - Not Specified* in the state of New Hampshire (0.90).

Channel - Affinity Sales is selected by Alabama, Georgia, Missouri, Ohio, and Oregon states, and accounted for 2.81% total call volume for these states, also contributes to increase the call transfer probability in these states. The total call volume for the five states is 15.45% of all dataset. A closer investigation of the data reveals that over 88% of these calls were transferred, and the caller intent was typically (63.7%) not specified. Since this type of channel is already designed to serve a specific audience, it can suggest that the type of customer using this medium of contact chooses to skip the IVR system, being directly transferred to a CSR. The immediate insights from this situation is the need to redesign or expand the capabilities of the

IVR system to provide the necessary information this type of customer requires, or the take the opportunity to offer these services online.

In contrast to *Channel - Service*, *Channel - Billing*, *Channel - Direct Sales Internet*, *Channel - Direct Sales Mass Media*, and all exhibit negative impacts on the target variable. *Channel - Billing* is the second most used channel across the country, and was selected as an effective feature by 41 states. These states comprise 96.9% calls, with *Channel – Billing* accounting for 29.2% of the calls. *Channel – Billing* calls in the 41 states presented a transfer rate of 48.2% and, most interesting, the contained rate for the non-transferred calls was 81.5% of the calls. These results suggest that *Channel - Billing* is more effective to handle and solve calls for which it was designed than other channels, in which the contained rates are practically irrelevant.

Channel - Direct Sales Internet was only selected by the Michigan state. *Channel - Direct Sales Internet* calls consisted of 5.9% of state transactions and presented a low transfer rate of 37.2% of the calls. *Direct Sales Mass Media* has the highest (in magnitude) average coefficient (-2.60) among all features. However, while these results might suggest that these channels are effective to handle the calls, a closer investigation of *Direct Sales Mass Media* calls in the state of Florida reveals that 96% were not transferred and, in line with *Channel - Direct Sales Internet* calls in Michigan, none were contained. Given the fact that these channels were designed to provide support to calls with specific motivation linked to sales, these numbers suggest that the system is not effective and potential revenues were lost.

Call-related features class is also represented in Figure 5.4. Results of 11 states indicate that whether the call is placed during business days or not is a relevant factor, while specific days of the week are not. The only exception is *Sunday*, which is selected by only two states, Massachusetts and Ohio, and exerts a negative impact on the call likelihood to be transferred. Together the two states comprised 4.86% of all dataset and had an overall transfer rate of 79.6%. However, if we restrict to the 13,013 Sunday calls, the transfer rate drops to nearly 50%. In total, only 3.5% non-transferred calls were contained. The other five relevant call related features concern the hour of the day the call is made, and all present a negative effect. Georgia, the state with the third largest call volume 8.29% of the country's call volume, selected four hour of the day features, covering hours from 2am to 5am and from 11pm to midnight. Georgia also presented the lowest overall transfer rate among all states with 57.7%. For the effective hour of the features cited, the call volume was 46,026 (or 5.75%), and the transfer rate was even lower, 9.99% or 4,597 calls. Non-transferred and contained calls added up to 1,934 or only 0.24%. *Hour of the Day – 23:00 to 24:00* was selected by Georgia and other three states: Colorado, Oregon, and Texas. The four states comprised 2,055,076 calls, with an overall transfer rate of 68.8%. By limiting the analysis for the last hour of the day, the call volume was 30,671 or only 1.5%, presenting a lower transfer rate of 23.2%. Curiously, contained calls added up to 9,280 or 39.4%, which is high in comparison to what we observed in other situations so far. The contained calls were usually motivated by inquiries related to *Payment* (89%). In summary, these results suggest that overall the time of call is not important to predict the call transfer likelihood as it is to predict the call arrival rate as investigated in the call center literature discussed in Section 2. However, we conclude that, in the process of developing a refined predictive model at the state level, local characteristics are revealed and contribute to these variables become relevant again. The importance extends further in helping to analyze and evaluate the quality of service offered.

Finally, *Cross-Class Related* features contains four features selected. Three of them are interactions of *Caller Type* and *Policy Type - Not Specified*. Combined to features such as *Interaction Modality - Not specified*, it can another indication that the current IVR system is not efficient, limited in terms of customer usability, independently of their effect on the call outcome. *Caller Type - Customer/Policy Type - Not Specified*, *Caller Type - Agent/Policy Type - Not Specified*, and *Policy Type - Not specified* are important features contributing negatively to the call transference likelihood. However, such lack of detailed information on the caller can be harmful for all calls, transferred and not transferred. Callers that may hang up because the system is ineffective can represent, for example, the loss of potential sales. On the other hand, customer that are transferred before this information is provided or captured by the system, can increase the total time a CSR will spend with the customer because of the extra effort to get basic information.

The problem of usability of call center IVRs has been the focus of research (Ndwe and Dlodlo, 2015; Suhm and Peterson, 2002). Suhm and Peterson (2002) used the length of the CSR-caller dialog to measure the benefit of IVR automation. His study proposed a methodology to evaluate cost-effectiveness and usability of IVR systems. Their analysis and described tools can help call center managers identify failures and how to improve them. Our work can be taken as complementary to the study in Suhm and Peterson (2002) that evaluated the interaction modality and estimate the benefits of usability re-engineering of the IVR system. However, the data explored in their study is very limited in size (only 5,530 calls) and it is not as rich in customer information as our dataset.

As was pointed out in the introduction paragraph of this section, the caller location is important to determine the set of effective features for each state. Table 5.2 groups states with the same size of set of selected features. Although we have several states with the same number of features, their composition is not necessarily homogeneous. The complete report of state models results is extensive and for space consideration are not presented here. Nonetheless, the main findings are highlighted as follows.

Table 5.2: Number of features selected by states.

# of selected features	Number of States	States
15	1	GA
12	3	OH, PA, TX
11	3	**AL**, **MO**, OR
10	4	CO, NY, SC, WA
9	4	CA, FL, **IL**, **KY**
8	7	AZ, IN, **KS**, NJ, **OK**, **TN**, WI
7	8	**CT**, MA, ME, MI, MN, **MS**, **NV**, VA
6	5	MD, **NC**, NE, **NM**, VT
5	8	**AR**, **DC**, **DE**, **IA**, **NH**, ID, **RI***, **UT***
3	4	**HI**, **LA**, MT, **WV**
2	3	**ND**, **SD**, **WY**

Note. States in **bold** have identical set of selected features. * RI and UT states have the same set of features but are different from the other states listed in the same row.

We begin by comparing models with 12 features: Ohio, Pennsylvania, and Texas. The three states have selected different set of features, but they all included: *Caller Intent - No Intent* and *Payment*; *Channel - Billing* and *Service*, *Customer Type - Third Party* and *Potential Customer*; *Interaction Modality - No Specified*; and *Payment Attempted*. For the group of states with 11 features, Alabama and Missouri have identical sets of selected features, and only *Customer Type - Potential Customer* diverge in the direction of impact (positive for Alabama, negative for Missouri). While the four states which selected 10 features do not have the same set of features, from states that include 9 features in their models, Illinois and Kentucky have identical sets of selected features. All features agree on the magnitude and direction of impact on the likelihood of call being transferred. California diverges from others for selecting features related to *Customer Type - Third Type*. Florida selected *Channel - Direct Sales Mass Media*, which have a very strong negative coefficient. As previously stated, 96% of *Direct Sales Mass Media* are not transferred also not contained. Since the channel of contact is specific for a certain sales type, these calls could have meant a new businesses.

The majority of the states selected between 5 to 8 features. From the states with 8 features selected, Kansas, Oklahoma, and Tennessee have an identical set of selected features. Features agree on the sign of coefficient, except by *Channel - Service* for the Oklahoma model which presents a negative effect. However, it is the only one not statistically significant. States that selected 7 features constituted the larger group with 8 states. Of those, only three Connecticut, Mississippi, and Nevada have identical set of selected features. All sign of coefficients agree among the three models. The *intercept* for the Nevada model is the highest in comparison to the other two, which means that Nevada callers are more biased to be transferred. *Potential Customer* feature impact on the call outcome is slightly higher for the Mississippi state. Moreover, *Customer Type - Potential Customer* in Mississippi represents 10% of total calls originated in that state, which is proportionally twice of what is observed in Connecticut and Nevada. In the group of state models with 6 features, the feature selection process resulted in identical set of features for New Mexico and North Carolina. All features agree on the direction of their effect on the call outcome, however, New Mexico intercept coefficient is significantly higher than in North Caroline, suggesting that callers from the southernmost state are more inclined to be transferred, all else being equal.

The group of states with model consisting of 5 feature is the most homogeneous in terms of features selected. Although the group is composed by 8 states, 5 (Arkansas, District of Columbia, Iowa, Delaware, and New Hampshire) have identical set of features, and other 2 (Rhode Island and Utah) also performed the exact feature selection, differently from the first five, evidently. The features are consistent on the sign of influence call transference likelihood. Only *Channel - Service* (and the intercept) have positive influence on the target variable. Iowa has the highest *intercept* (4.21) among all 50 models, which is quite expected considering that 9 out of 10 calls originated there are transferred. Finally, 3 out of 4 states with models including only 3 features have identical set of features: Hawaii, Louisiana, and West Virginia. The features are the top ranked in Figure 5.4: *Payment Attempted*, *Channel - Service*, and *Interaction Modality - Not Specified*. West Virginia stands out by having the highest coefficient in magnitude for the *Payment Attempted* feature (-3.81). North Dakota, South Dakota, and Wyoming only included the top 2 features. It is worth noting that, as the set of features become more selective, the selected features are restricted to only the most important and frequently selected among other states.

5.5.2 Model Performance

The final models were also evaluated by two performance metrics widely used for classification problems: the overall accuracy and the Area Under the Receiver Operating Characteristic Curve (AUC), used to evaluate classification ability of the regression, "assessing the trade-off between sensitivity and specificity" (Hastie et al., 2001)(see also Bradley (1997). The AUC score ranges from 0 to 1, where 0.5 indicates that the method is randomly classifying the response variable, and 1 represents a perfect classifier. The outcome of the regression model is the predicted likelihood of call to be transferred y_i which is a value ranging from 0 to 1 for each observation of the test data. The predicted call outcome y_i' is classified as follows

$$y_i' = \begin{cases} 1 & \text{if } 0.5 \leq y_i \leq 1 \\ 0 & \text{if } 0 \leq y_i < 0.5 \end{cases} . \tag{5.5.1}$$

Since the overall call transfer rate is relatively high, reaching up to 90% for Iowa state, relying solely on the overall accuracy could lead to misinterpretation of the results. Figure 5.5 plots the overall accuracy versus the AUC score for all 50 models. The highlighted area shows that all models have both accuracy and AUC score above 75% indicating that in general all models performed well. Models for Georgia state presented both highest accuracy (95.44%) and AUC score (90.44%).

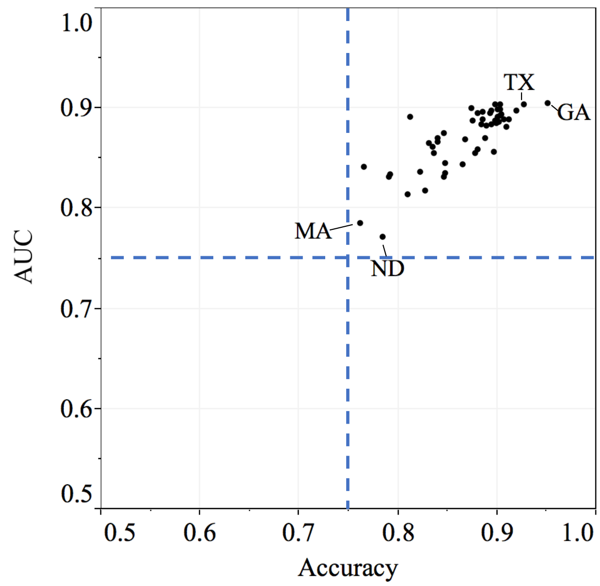


Figure 5.5: Overall accuracy vs. area under the ROC curve.

5.5.3 Customers' Segmentation Analysis

In order to draw caller's profiles and assess how they compare to the entire population based on the top 27 features of Figure 5.4, we segment the test population of each state according to their predicted probability to be transfer. First, we rank the callers by the predicted probabilities and divide the test population into 10 deciles. The first decile corresponds to the 10% of the population with the highest predicted likelihood, while the tenth decile refers to the 10% with the lowest likelihood to be transferred. We then introduce the Standard Deviation Index (SDI) as a metric to compare the features for each decile of the population. The SDI is calculated by

$$\sigma_{F,d}^{s} = \frac{\mu_{F,d}^{s} - \mu_{F}^{US}}{\sigma_{F}^{US}}, \tag{5.5.2}$$

where $\mu^s_{F,d}$ is the mean value of feature F, for the d decile of test population of state s, μ^{US}_F is the mean value of feature F in the entire country population, and σ^{US}_F is the standard deviation of of feature F in the entire country population.

The SDI allows us to evaluate how much a feature of a given group from a specific state differs from the average of the national population. Figure 5.6 makes it possible to draw a parallel between callers with the highest likelihood to be transferred in all states. It shows that different states have particularities that are only captured because of the "personalized" regression models built. Although several features such as *Channel - Service*, *Payment Methods - Mastercard* and *Checking*, *Payment Attempted*, and *Customer Type - Potential Customer* diverge similarly from the national average customer characteristics for all state caller profiles, we observe, for example, that for Michigan, Ohio, South Dakota, and Texas, *Customer Type - Third Type* present a more significant and positive divergence from the national average, comparing to callers from New York, Oregon, and Illinois, just to list a few. It can be seen from the heatmap that callers from South Dakota and Texas have very similar profiles, as well as the pair of caller profiles from Illinois and New York. While the correlation analysis showed no significant correlation between the set of selected features for each state or their coefficients and geographic regions and divisions, these similarities could be used to cluster customers with identical demands. On the other hand, the singularities could also be useful to identify opportunities for local improvements.

Considering that over half of all calls in 2015 originated from only 7 states, we will analyze their profiles in more detail. Figure 5.7 shows the SDI chart for New York, Texas, Georgia, California, Pennsylvania, Florida, and New Jersey, ordered from highest to lowest call volume. The features are sorted descending based on the SDIs for the New York state. The chart illustrates how *Customer Type - Third*

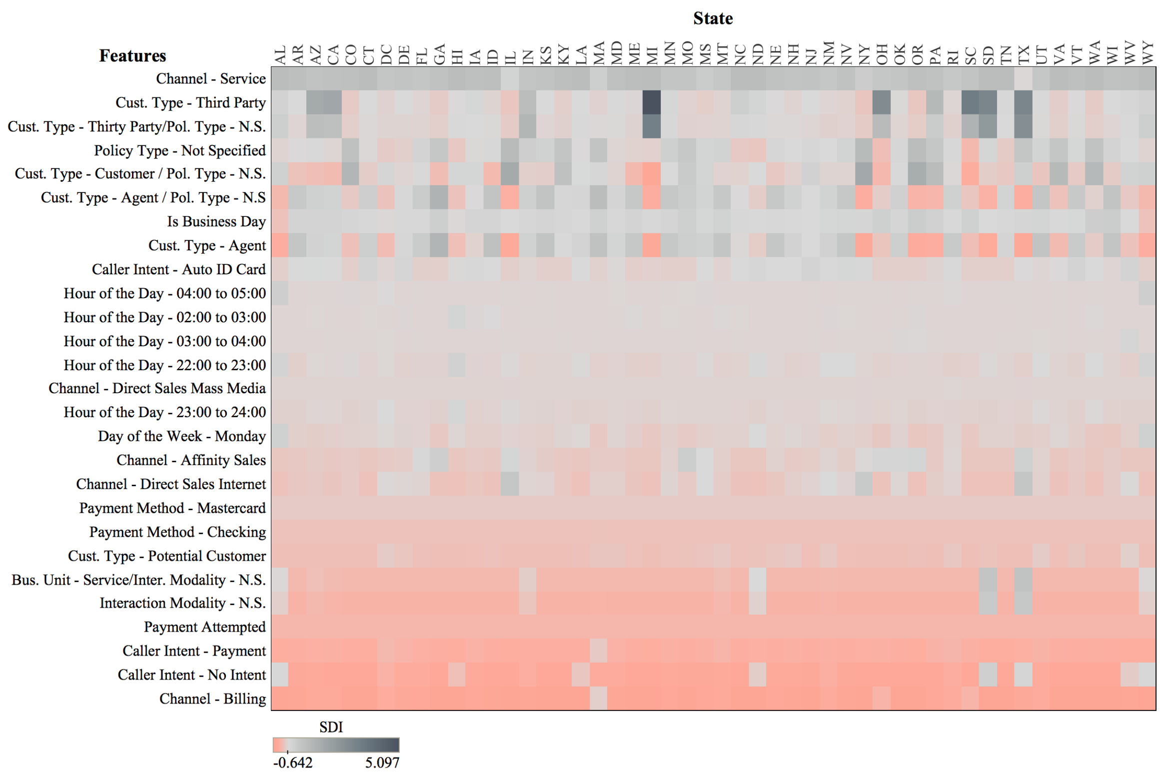


Figure 5.6: Heatmap of Standard Deviation Index (SDI) for each feature broken down by the top 10% population with highest predicted likelihood to be transferred of each state.

Party for Texas and California diverge significantly from the national average. It also shows that *Customer Type - Customer* for New York, and *Customer Type - Agent* for Georgia are essential characteristics that deviate from the national average. Finally, the Florida and New Jersey profiles are quite similar and, in contrast to other states, for Texas and New York the most significant *Channel* is *Direct Sales Internet*.

5.5.4 Correlation Analysis

As previously stated, the *average temperature* feature reveals not to be a relevant feature for any state model. We expected that the average temperature could reflect, for example, seasonality or also be a grouping factor that would exert similar effects

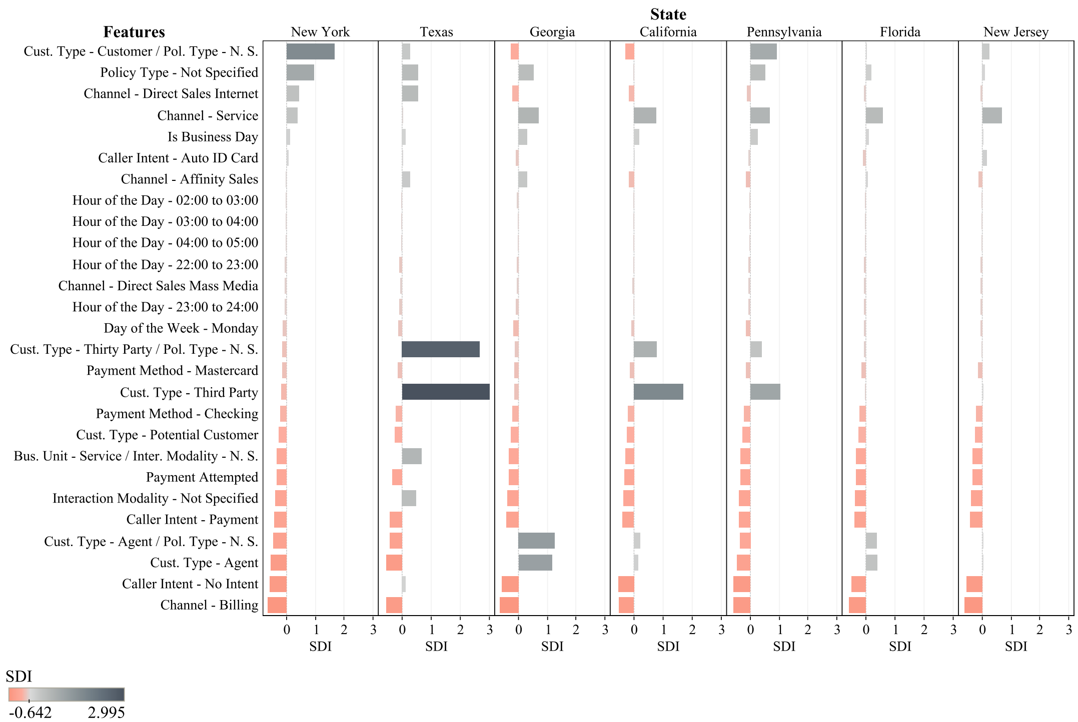


Figure 5.7: Standard Deviation Index for each feature for top 10% population with highest predicted likelihood to be transferred of each state.

on the call outcome for neighbor states or with comparable weather characteristics. While the results reported pointed out the singularities of different state models, numerous associations were also disclosed. As the temperature could not capture or represent none of this associations, we calculate the correlation between the features selected by each state and social and geographic aspects.

The considered aspects are: education index, population density, and the country's geographic region and divisions. Geographic information, state population and area were gathered from Census (2017a), and Census (2017b). The education index is represented by the percentage of the state population with 25 years and over with a bachelor's degree (Census, 2017c).

However, the results were consistent with the temperature feature had not been selected. We found no significant correlation between customer's regional origin and the selected features, besides a correlation of 0.55 between the geographic Division 2 and *Channel - Direct Sales Mass Media*, select by New Jersey, Pennsylvania, Florida, and Maryland, and correlation of 0.57 with *Payment Method - Checking*. There is also no significant correlation between state education index the selected features.

5.6 Conclusions and Future Research Recommendations

In this chapter, based on customer detail information, we study the most effective features for determining the outcome of a call arriving at the IVR system of a major insurance company call centers. The richness and volume of the real data investigated are key properties of this research that makes it stand out in relation to other works on the field.

The results indicates the caller location (State) plays a key role on determining the most relevant features and their contribution on the call outcome. From the initial 181 features considered for our regression model, only 27 were selected by at least one of the states. Our analysis underscores that, although some factors have consistent effects on the call transfer likelihood for all states, in order to have an accurate understanding and prediction of customer behavior it is crucial to have local models. While we found no correlation between the caller regional origin or demographic characteristic, the results reveal numerous local characteristics that could be closely investigated by the company, and could mean opportunities for new business, and provide support for specific reasons or products for a particular state. The case of *Channel - Direct Sales Mass Media* for Florida, where almost none of the calls are contained, illustrates the opportunity to redesign the IVR system to handle this type of calls.

The limitations of the IVR system and how it impacts the call outcome becomes evident as relevant features reveals the lack of specified information. In total, 4 features are related directly to the not specification of the policy type, and all present a positive impact on the likelihood of a call to be transferred to a customer service representative. In addition to fact that the features *Interaction Modality - Not Specified* and *Caller Intent - No intent* are also relevant, it can be an indication that the current platform is not intuitive or not effective to capture this information. This can be interpreted in two ways. First, calls that are not transferred do not implicate in the call center operational cost, since no CSR is needed. However, in the case of transferred calls, the agent loses time having to request and confirm basic customer information that could already be captured by the IVR. Secondly, calls that did not require the representative services, are often not contained, and the company could have lost a potential sales, for example.

Finally, we can summarize from the analysis that transferred calls are typically from *Service Channel*, and are made during business days by *Agent* or *Third Party*. Calls with motivations related to billing and payment tend to not be transferred to a CSR, suggesting that the system is effective to handle those demands.

The fact that we did not have access to attributes used to identify existing customers, such as the policy account number, was certainly a limitation for our analysis, as we assumed that every call was made by different people, and no intrinsic characteristics of the individual could be take into account in our model. This type of information would allow us to investigate the history of contact of the customer and extract additional information related to recency, frequency of contact, sequence of contact through different mediums of contact, past call outcomes. Future research could fruitfully explore more correlations between the relevant set of features identified in this study, possibly applying machine learning methodologies, such as neural network, that can implicit detect complex nonlinear relationships between input and output.

Chapter 6

Simulation-based Assessment of Data-Driven Prediction in Customer Support Systems

6.1 Introduction

Organizations from all sectors have invested in research to develop data-driven technologies to improve operational processes and product quality, reduce costs, increase sales and customer engagement, sell experiences rather than services, and create new business models (Ransbotham et al., 2017; Cohen, 2018; Davenport and Ronanki, 2018; Hanifa et al., 2014). In 2019, McDonald's invested millions of dollars to develop Artificial Intelligence (AI) and machine learning tools to boost sales (Yaffe-Bellany, 2019). Technologies combine diverse information such as weather conditions, time of day, and customer order history data obtained from reading a license plate to make order recommendations. According to a McKinsey article (Brocchi et al., 2018), a US bank gained $2 billion dollars by adopting data-driven initiatives that generated additional revenues and made operating processes more efficient.

In the service industry, companies use data-driven technologies to enhance customer relations, forecasting, and sales automation. While customers look for unique, different, and close-to-reality experiences, companies seek to treat each customer according to their peculiarities, needs, and expectations (Ransbotham et al., 2017). Machine learning algorithms identify patterns in customers' data and translate them into tangible business insights. The goal is to predict, for example, what type of product a customer is most likely to buy, detect fraud in real-time (Chen et al., 2015), or customize digital marketing campaigns (Andrews et al., 2015; Moazeni et al., 2019).

Moreover, chatbots (e.g., used by banks or mobile food ordering apps), service and product recommendation systems (e.g., used by Amazon, Netflix, and Spotify) contribute to increasing customer service personalization as well as encouraging interaction between customers and businesses (Chakrabarti and Luger, 2015; Chen et al., 2016). Virtual assistants, such as Apple's Siri, Amazon's Alexa, Microsoft's Cortana, and Google's Google Assistant, collect customer behavioral data, understand their input and feedback, stimulate customer engagement, and enhance the customer experience (Jiang et al., 2015; Kiseleva et al., 2016).

Data-driven technologies, referring to a solid base of data that generates valuable insights to support decision-making rather than intuition (McAfee and Brynjolfsson, 2012; Brynjolfsson and McElheran, 2016), now constitute the core business model of many start-up firms, e.g., see Hartmann et al. (2016) for a review. With the emergence of new communication channels, such as emails, online accounts, social media, and mobile apps, traditional call centers have turned into multichannel customer support systems, or contact centers (Saberi et al., 2017). Available omnichannel data and advanced data analytics integrate several aspects of their operation strategies to improve customer experience and enhance customer satisfaction (Deloitte, 2017; Andrade et al., 2020). Data-driven technologies enable contact centers to shift from being reactive to more proactive to better assist customers and enhance their interaction experience (Moazeni and Andrade, 2018; Moazeni, 2018).

Customer support systems can leverage customer transactional data collected in the contact channel to better understand customer queries and provide self-service support to improve routing and customer service representative (CSR) productivity (Rustamov et al., 2018). Intelligent contact routing can reduce customer waiting time and increase agent effectiveness for contact resolution (Mehrotra et al., 2012; Taylor, 2017). Improvements in both aspects can lead to higher customer satisfaction

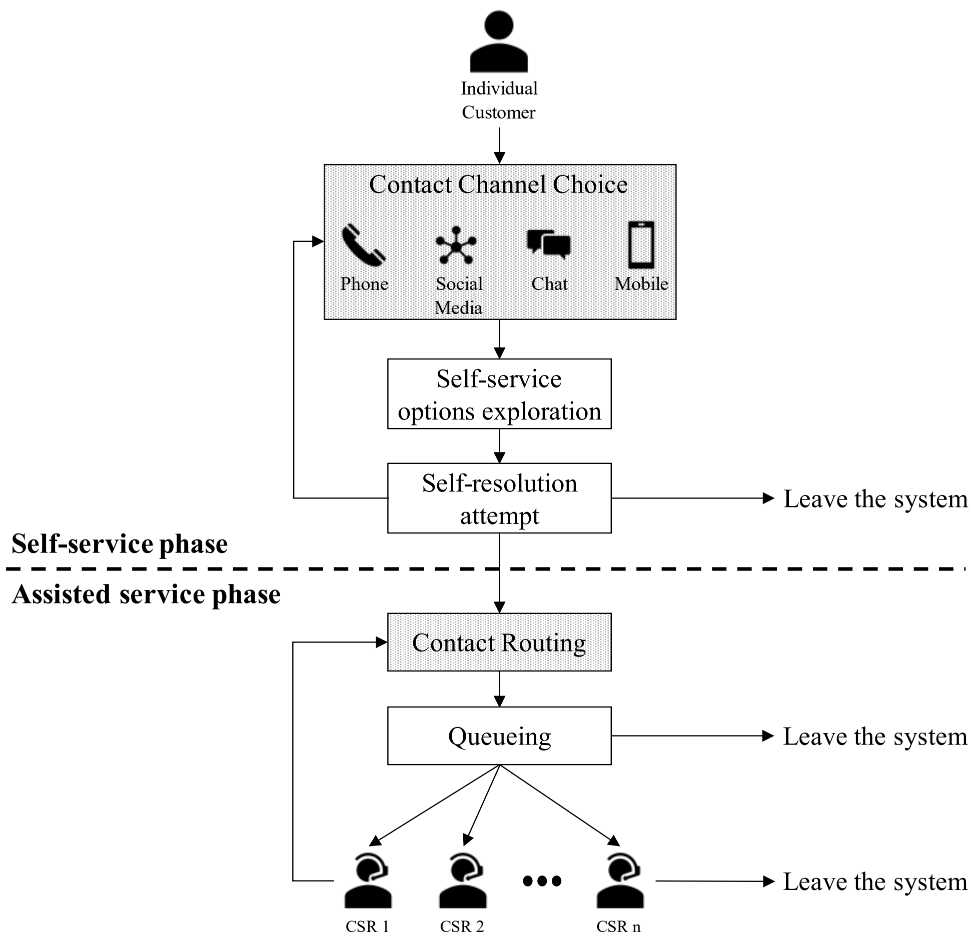


Figure 6.1: Customer service support flow diagram.

and lower operations cost (Aksin et al., 2007; Seada and Eltawil, 2015). While the data-driven technologies advance at an ever-increasing pace, company transformation processes require an approach that allows for quick, inexpensive, and flexible research to inform decisions.

This paper studies how to evaluate technology infusion for a generic multi-channel customer support system depicted in Figure 6.1 that comprises customer interaction processes and choices that can benefit from data-driven technologies. We distinguish between two main layers of the customer support process: the self-service phase and the assisted phase. The first phase includes customer support requests and a selected contact channel. Customers explore self-service options and attempt to resolve their query. If successful, the customer leaves the system, otherwise they can opt to request a CSR assistance, seek another communication channel, or abandon the system. The second phase includes contact distribution processes and CSR assistance. Contact distribution combines routing and queuing activities to connect and match customers to agents with the right skills (Chan et al., 2014). The system accounts for intra-agent transfers, and customers either abandon the system while waiting in queue or after receiving assistance.

Using modeling and simulation, this study seeks to provide operational managers with an approach to assess how adopting smart technologies and data-driven processes impact the customer support system performance. The objective, however, is not to develop an algorithm and produce quantifiable results for a specific case, but propose a process to inform strategic system design decisions. The experimental model investigates the infusion of digital technology in the contact center under various system configurations. Alternative system configurations include one or two self-service contact channels and data-driven mechanisms to: 1) allocate customers to channels and 2) route customers to agents characterized by the functional performance (accuracy). The two concepts directly impact the processes highlighted in Figure 6.1. Experimental analysis tests alternative system design decisions and the impacts in required staffing levels to meet a target service quality. Alternative design decisions combine changes to the system configuration (adding a new self-service channel) and improving the accuracy of existing customer-channel allocation and customer-agent routing processes. Resulting discussions identify which technology combinations can achieve four cost reduction targets (8%, 13%, 20%, and 27%).

This paper contributes a generalizable simulation-based assessment approach to evaluate the impacts of lower level design decisions at the system level before making a technology investment. Recognizing that different technologies present a range of capabilities to improve customer support, the simulation model is a tool that evaluates complex problems using a simplified model of system functions. Results generate insights from a system perspective based on system behavior observations without requiring the implementation of learning algorithms. We demonstrate with a notional case how to use simulation models as the benchmark tool that bridges strategy and operations to inform strategic decision.

The remaining sections of the paper are structured as follows. Section 6.2 reviews background literature on customer support systems and how customer services benefit from data-driven technologies. Section 6.3 describes the multichannel customer service support system design. Section 6.4 explains the experimental design and presents simulation results. Finally, Section 6.5 draws the main conclusions a provide directions for future work.

6.2 Customer Support Systems

Customer support services focus on creating a more dynamic, intuitive, and natural experience. Customers are more connected to companies, especially in digital channels (Saberi et al., 2017; Deloitte, 2017; Mueller et al., 2015), which provide detailed data on the customer interaction processes and permit the opportunity to offer customized and lower-cost self-service alternatives. Other phases of the customer service system also generate large volumes of data with high speed and variety, with the potential to create new design solutions for the processes involved. In this section, we review how data-driven technology impacts the processes depicted in Figure 6.1.

6.2.1 Self-Service Phase

Companies constantly adapt to new trends and customer demands in a dynamic domain. The use of smartphones created a customer interaction environment that did not exist decades ago and companies have been developing new ways to relate to them. The opportunity expanded the use of self-service environments to take advantage of lower operational costs, while also targeting higher customer satisfaction, retention, and loyalty (Kumar and Telang, 2012).

For instance, the use of Interactive Voice Response (IVR) units allows users to interact over the phone without assistance from human agents (Tezcan and Behzad, 2012). Web platforms allow customers to track orders, review items, cancel unshipped items, and request returns or refunds (Keblis and Chen, 2006). By using advanced text mining and natural language processing techniques, chatbots can quick digest customer data digestion and provide effective and personalized service (Scherer et al., 2015; Verhagen et al., 2014).

Shifting demand from higher-cost assisted channels to efficient self-service channels is a strategic decision to improve customer experience and reduce operations costs (Tezcan and Behzad, 2012). Moazeni and Andrade (2018) shows website use decreases the likelihood insurance company customers will call back in the future compared to individuals who use the telephone. Therefore, investing in online platforms to offer functionalities previously only available over the phone and encouraging customers to use these channels could lead to a reduction in operating costs.

6.2.2 Assisted Service Phase

Customers unsuccessful in solving queries in the self-service channel may request further assistance from a CSR. *Contact routing* directs customers to CSRs with the appropriate skillset and is an essential step for effective and efficient service. An accurate routing process should limit, for example, transfers between agents to reduce time spent in ineffective service, resulting in faster service and higher customer satisfaction (Tang et al., 2003).

The mechanisms that seek to match customers to agents commonly combine routing and queuing processes. The simplest method to define how customers wait in queues is the first-in, first-out (FIFO). However, complex service systems with heterogeneous servers and customers often require rules to prioritize certain customers. For example, customers allocated to an agent with a mismatched skillset can have higher priority when transferred to another queue. Different customer profiles may also require different priority levels. Calling customers are impatient, so phone inquiries should have higher priorities than e-mail messages (Ahghari and Balcioglu, 2009).

Typical routing and queuing strategies accommodate specialized agent or cross-trained agents. This latter setting is commonly referred to as "skill-based routing," where customers are directed to different CSRs based on the type of query they are trained to handle (Ali III, 2010; Mehrotra et al., 2012).

Research in call centers solves the problem of optimizing routing mechanisms with dynamic programming techniques, incorporating agent and customer data to improve performances (Mehrotra et al., 2012). Chan et al. (2014) proposed a dynamic model for contact routing that combines the heterogeneity of agents and customers to match a type of call and a group of agents. Other studies that consider heterogeneous servers and customers include Mehrotra et al. (2012) and Ibrahim et al. (2016). Mehrotra et al. (2012) investigates dynamic routing strategies to match call types with agent skills, subject to their effectiveness for each call type. Finally, Ibrahim et al. (2016) incorporates agent and call type heterogeneity, and time-dependent agent performance to model inter-dependent service times.

6.2.3 Channel Allocation

Companies can use information from multiple sources and data-driven prediction methods to anticipate the reason, time, and channel by which customer queries arrive at a customer support service system, and to further allocate individuals to a specific channel that is considered most appropriate (Moazeni and Andrade, 2018; Jerath et al., 2015). For instance, a web channel may be more suitable to solve standardized services (Boyer et al., 2002). In contrast, human agents tend to have a greater understanding and resolution capacity for more complex inquiries. Therefore, directing customers to various service channels based on predicted demand and channel performance can bring benefits in terms of service quality and operational efficiency.

Data-driven technologies can leverage customer attributes from recorded interactions to more accurately predict future interactions. In Jerath et al. (2015), a stochastic function of customer information needs models customer query frequency and contact channel choice in multichannel customer support services. The approach uses individual-customer-level data on claims and channel usage to predict demand for different contact channels at aggregate and individual levels. Studies show that several factors influence the adoption and use of various channels, such as customer demographics, the type of information structure available in each communication venue, and the type of query (Xue et al., 2007; Campbell and Frei, 2010; Kumar and Telang, 2012). Furthermore, Jerath et al. (2015) emphasizes the potential use of more detailed data from transcripts of telephone conversations and usage patterns of other channels to fully develop and estimate the model.

6.2.4 Contact Routing

Enhanced contact routing can benefit from accurate classification of customers and query types using customer attributes in the self-service phase and historical customer multichannel interaction data. The high volume, variety, and speed at which data is generated and processed requires varied and advanced analytics techniques to improve real-time customer support service quality. Customer routing and queuing can also be designed to use agent profiles to determine the attributes that best match each customer's demographics and improve contact outcomes through machine learning methods that aggregate different types and levels of data. The result is a more flexible system that extends traditional skills-based routing, allowing for more personalized customer experience (Schoeller and Heffner, 2014).

Furthermore, studies on natural language processing (NLP) improve contact routing processes based on unstructured data collected in real-time during service delivery (Chu-Carroll and Carpenter, 1999; Zitouni et al., 2003; Tyson and Matula, 2004; Sarikaya et al., 2014). NLP techniques seek to make the call routing process more convenient for customers using speech-enabled systems. Customers often struggle to define which option best fits the customer query when using closed-menu systems, leading to frustration and decreased service satisfaction. The challenge, therefore, is to understand the query from the customer's own description, rather than using touch-tone or speech-enabled menus (Rustamov et al., 2018).

Big data from contact centers opened up opportunities to study matching customers with agents based on behavioral profiles. The motivation of the approach is based on a *homophily effect* between customers and agents. Social homophily is the process by which individuals form bonds based on shared characteristics (Domingue et al., 2018). The goal is to consider customer/CSR chemistry, in addition to the type of problem and agent training, to determine the best matching and enhance customer satisfaction. Ali (2011) proposes a contact routing model to maximize customer satisfaction based on underlying customer and CSR demographics, psychographics, and historical performance data. The proposed method uses several machine learning techniques to score available CSRs against the customer waiting in the queue. The highest score determines which agent should be paired with the customer.

In summary, companies can achieve a competitive advantage from implementing data-driven technologies at different stages of the customer service process. Transforming data from consumer activities, agents, and operational processes into insights enable the development of systems with dynamic, adaptive capabilities that create value in customer satisfaction or operational efficiency (Erevelles et al., 2016).

However, when evaluating significant changes to the contact center system design, historical data is no longer available for new configurations, limiting a direct comparison between the current and proposed models. In design problems, decisions require analysis of what-if scenarios varying key assumptions and constraints to find a balanced solution. Implementing changes in the system, even on a small scale, to gather data and compute the potential benefits can be costly and infeasible to replicate many times (Rosen et al., 2015). Relying on synthetic data from simulation models is ideally suited to describe the real-life tradeoffs. Thus, this study develops a simulation-based assessment of a multidimensional trade-off analysis between lower-level design and system-level strategies and costs savings from reduced workforce levels.

6.3 Model Description

Simulation models are tools that enable the investigation of how lower-level design decisions affect system-level performance. Simulation analysis is a bottom-up approach that assesses, for example, how a technology decision in the components of the customer support system (such as implementing a new contact classification algorithm) influence the overall system performance metrics such as service level, average speed to answer, and customer satisfaction.

This study seeks to understand the system-level consequences of implementing data-driven technologies for contact centers. We investigate the infusion of digital technology and data-driven approaches in the customer support system under three technology immersion scenarios:

1. Offer an additional contact channel that provides self-service query resolution options to customers.

2. Implement an allocation platform that guides customers to the communication channel with highest contact resolution capacity.

3. Improve the routing mechanism to match customers with the most appropriate CSR based on the agent skill set and customer query type.

This study models four customer support system configurations and simulates the system-level effects from changes to lower level system designs by varying the performance levels of the channel allocation platform (item 2) and contact routing mechanism (item 3).

6.3.1 Discrete-Event Simulation

The activity diagram in Figure 6.2 describes the proposed customer support service system as a discrete event simulation with Run, Handle Customer, CSR assistance activities. The gray boxes indicate the processes influenced by the data-driven technologies of interest: the channel allocation platform, and the contact routing mechanism. Each activity is described as follows:

1. *Run Activity:* Generates customer demands by creating a new customer every inter-arrival time step with randomly assigned query type, channel preference, and mean patience.

2. *Handle Customer Activity:* Processes both self-service and assisted service phases of the customer support system. During the self-service stage, a platform allocates the customer to a channel. The customer checks the channel availability (may be unavailable due to scheduled or unscheduled maintenance) and attempts to resolve the query using available capabilities. The customer either leaves the system if the query is solved or abandons without requesting a CSR assistance.

3. *CSR Assistance Activity:* Processes a customer request for an additional human-agent support. Checks if the contact routing mechanism successfully directs the customer to an agent with the proper skills to assist with the query, otherwise, assigns the customer to an agent lacking the required skill set. The customer requests CSR support and waits in the corresponding queue if all agents are busy. If the customer does not abandon the system while waiting, the CSR

provides the requested assistance and is released after the service time elapses. If the customer was initially routed to the appropriate CSR, the customer leaves the system, otherwise they are re-classified and transferred to new agent with the appropriate skill set.

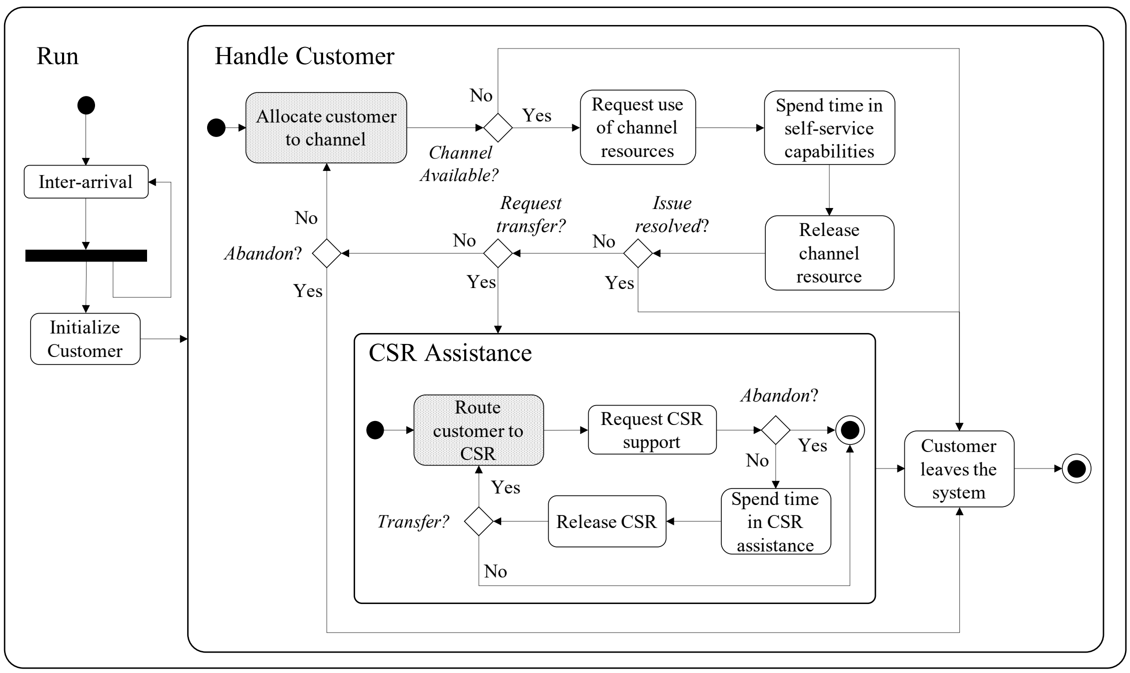


Figure 6.2: Simulation Activity Diagram.

6.3.2 Model Details and Assumptions

This subsection describes modeling details including key assumptions grouped into four major categories.

Individual Customer: Customers arrive following a Poisson process with rate λ. Customers have one of two query types $q \in \{A, B\}$, sampled from a discrete uniform distribution. In addition, customer patience is exponentially distributed with mean time θ, independently generated each time they are put "on-hold".

Contact Channel: Customers have the option to use two contact channels, Telephone (T) and Web (W), directed by a channel allocation platform. The customer channel allocation S follows a Bernoulli distribution with parameter γ_1. This allocation is "accurate" ($S = 1$) if the customer chooses, with probability γ_1, to use the channel $c \in \{T, W\}$ with highest contact resolution probability $\rho_c \in [0, 1]$ and the capability to solve query q. After allocation, the random variable $Y \sim Bernoulli(\rho_c)$ denotes whether query was successfully resolved ($Y = 1$) with probability ρ_c after a customer uses channel c. In addition, $X \sim Bernoulli(\varrho_c)$ defines whether an unresolved query in channel c leads to a CSR assistance request ($X = 1$) or abandonment ($X = 0$) with probability $\varrho_c \in [0, 1]$. The time customers spend in contact channel c is considered to be an independent identically distributed (i.i.d.) exponential random variable with rate μ_c. The random variable $V \sim Bernoulli(\psi)$ determines whether the communication channel is available at the time a customer seeks support with probability $\psi \in [0, 1]$. Finally, the capacity of communication channels is assumed to be 10,000 customers.

Contact Routing: The random variable $Z \sim Bernoulli(\gamma_2)$ defines whether the contact routing mechanism successfully matches the customer to an appropriate agent ($Z = 1$) with probability γ_2, referred to as the *contact routing accuracy probability*. We assume the automated contact routing platform has an adaptable capability to allocate customers to the CSR queue capable of handling their query. Queues follow a first come, first-served rule.

CSR Assistance: We assume heterogeneous CSRs with no cross-training and service time dependent on the agent skill set and accurate contact routing. The service time for customers matched to an agent lacking the skill set to resolve their query is an i.i.d. random variable following a exponential distribution with rate ω. This service duration models the time required to transfer to another agent and is assumed to be independent of query type. The service times for agent with the appropriate skill set to solve query q are i.i.d. exponential random variables with rate ν_q. Finally, all CSRs solve customers queries with a fixed contact resolution probability $\varphi \in [0, 1]$.

Table 6.1 shows the complete list of model input parameters and associated values further discussed in Section 6.3.6.

Table 6.1: Simulation Model Input Parameters.

Classes	Parameter	Value	Unit
Contextual Parameters	λ - Customer arrival rate	20	cust/min
	θ - Mean customer patience time	4	min
	ρ_T - Contact resol. prob. - Tel.	50.00	%
	ρ_W - Contact resol. prob. - Web	75.00	%
	ϱ_T - Transfer prob. - Tel.	80.00	%
	ϱ_W - Transfer prob. - Web	80.00	%
	μ_T - Service rate - Tel.	1	cust/min
	μ_W - Service rate - Web	0.50	cust/min
	ν_A - CSR Service rate - Query A	0.50	cust/min
	ν_B - CSR Service rate - Query B	0.25	cust/min
	ω - Service rate - transfer CSR	2	cust/min
	φ - Contact resol. prob. CSR	95.00	%
	τ - Maximum waiting time	0.5	min
	ψ - Contact channels availability	99.00	%
	Contact channels capacity	10,000	cust
System Configuration	Resolution capability - Tel	A&B	query
	Resolution capability - Web	A, B, A&B	query
	γ_1 - Channel allocation accuracy	50-100	%
	γ_2 - Contact routing accuracy	60-100	%

6.3.3 System Value Performance Metrics

Contact centers use several metrics to assess operational performance (see Gans et al. (2003) and references therein). Besides controlling the customer service system performance, these metrics are also important for operational planning, such as workforce scheduling based on a target performance. These metrics are constraints imposed on optimization problems to define the minimum number of agents required to meet the desired levels of service quality.

Our experiments seek to assess how the changes to the system design impact the number of CSRs needed to met a determined service quality. The service quality definition and measurement is widely discussed topic in service industry (Andrade et al., 2020) and it is out of the scope of this paper. For simplicity, we use in this paper the service level (SL) to measure service quality. SL is the percentage of customers that waited to be served by a CSR for a time below a target threshold. As the workforce can represent up to 70% of the budget in call centers (Gans et al., 2003), reducing the number of CSRs can result in a significant reduction in the contact center's operating cost.

Let t_i denote the time that a customer i waits in queue to be assigned to a CSR, where $i = \{1, 2, ..., I\}$ and I is the total number of customers that requested assistance over a given simulation time. Then, the percentage of customers with waiting time below the given threshold level τ is denoted by $SL(\tau)$ and is equal to

$$SL(\tau) = \frac{\sum_{i=1}^{I} \Gamma(t_i)}{I} \tag{6.3.1}$$

where $\Gamma(t)$ is an indicator function which returns 1 if $t \leq \tau$, i.e., the waiting time is within the response time threshold τ, and returns 0 is $t > \tau$.

Let G_q be the subset of customers with query type q. The percentage of customers with query type q assigned to a CSR within the given response time threshold τ is given by

$$SL_q(\tau) = \frac{\sum_{i \in \mathrm{G_q}} \Gamma(t_i)}{|G_q|}, \forall q \in \{A, B\} \tag{6.3.2}$$

where $|G_q|$ is the number of customers in the subset G_q. For given system performance requirements ϕ, ϕ_A, ϕ_B, we can compute the minimum number of CSRs by solving the following integer linear optimization problem:

$$
\begin{aligned}
\min \quad & \sum_{q\in\{A,B\}} N_q \\
\text{subject to} \quad & SL(\tau) \geq \phi \\
& SL_q(\tau) \geq \phi_q \quad ,\forall q \in \{A,B\} \\
& N_q \geq 0 \quad ,\forall q \in \{A,B\}
\end{aligned}
\tag{6.3.3}
$$

Here, the decision variables N_A and N_B, respectively, indicate the number of CSRs with skill sets to handle query types A and B. In our simulation study, we set $\phi = 80\%$ and $\phi_A = \phi_B = 75\%$ for the overall and individual queue service level targets.

6.3.4 Model Implementation

The simulation model is implemented in Python 3.7 using the SimPy (v. 3.0.11) library, an open-source process-based discrete-event simulation framework (Meurer et al., 2017), and the source code is available on a public repository. [1]. Four object-oriented programming classes are created based on the activity diagram: Customer, Channel, Agent, and Contact Center. While the latter defines the simulation structure, the other three define the attributes of its elements. The Customer class defines channel preference, type of query, and patience for each individual. The Channel class assigns the medium of contact, capacity, average service time, the list of capabilities, the contact resolution probability, and the venue availability. The Agent class defines the CSR training type, average service time, contact resolution probability, and the number of agents. Both Channel and Agent classes are defined as shared resources. Finally, the Contact Center class performs the simulation by generating demand and executing the Handle Customer function that implements the processes illustrated in

[1] Available at: github.com/rodrigocaporali/contact_center_simulation

the activity diagram.

6.3.5 Model Verification

To ensure the proposed model provides valid insights from simulation experiments, we first verify the assumptions with the steady-state Erlang A model. Erlang A extends the Erlang C, the simplest and most widely-used model for call centers, by allowing abandonment and is represented as M/M/N+M. The model assumes calls arrive according to a Poisson process and are served in a first-come first-served queue discipline. Furthermore, customers have finite patience, service time follows an exponential distribution, and agents are homogeneous and independent of each other. Finally, processes of arrivals, patience, and service are mutually independent (Garnett et al., 2002; Mandelbaum, 2004; Gans et al., 2003; Brown et al., 2005).

Verification experiments vary four parameters: inter-arrival rate λ, average service time μ, number of servers N, and patience θ as follows: $\lambda = \{1, 10, 25\}$, $\mu = \{1, 2\}$, $N = \{\lambda\mu + 1,\ \lambda\mu + 3,\ \lambda\mu + 5\}$, and $\theta = \{2\mu\}$. For each scenario, we simulate a total of 100 runs for a period of 3000 minutes, selected based on graphical and statistical analysis of steady-state behavior. We use the first 1000 minutes as a warm-up period, and compute different performance metrics using the remaining 2000 minutes. Analysis compares the abandon rate (percentage of customers leaving the system before CSR service), average time customers spend in the system, and service level. Results attest the validity of the model as all theoretical estimates fall within in the 95% confidence intervals of simulated results.

6.3.6 Model Validation

Validation evaluates whether the simulation model resembles the real-world system of interest. Our simulation model structure incorporates the main components of real customer support systems such as multiple communication channels with distinct capabilities, capacity, and efficiency, heterogeneous customers and agents, and multiple contact types. The importance of accounting for heterogeneous customers and agents, multiple query types, and multichannel environment, is widely recognized in the literature to enrich experimental analysis and extend conclusions to managerial recommendations (Mandelbaum and Zeltyn, 2009; Ahghari and Balcioglu, 2009; Jerath et al., 2015). Our model relies on requirements often taken by contact centers, assumptions applied in related literature, and real data gathered from contact centers of a major U.S. insurance company (see Moazeni and Andrade (2018) for a detailed data description).

Contact center customer arrival rates vary depending on the month of the year, day of the week, time of the day, and also contact channel type (Aksin et al., 2007; Taylor, 2012; Andrade and Moazeni, 2017). Although the arrival rate can be as high as a 100 contacts per minute, we assume a constant $\lambda = 20$ customers per minute to simulate a medium-to-large size contact center operating in steady state while balancing computational limitations.

In terms of distribution selection, customer arrivals are typically assumed to follow a Poisson process (Cezik and L'Ecuyer, 2008; Avramidis et al., 2010; Khudyakov et al., 2010; Ibrahim and L'Ecuyer, 2012). To determine the average patience time we assume a rule-of-thumb to be approximate to two times the average service time (Mandelbaum and Zeltyn, 2009).

We account for heterogeneity in agent service rates and the 80% self-service contact resolution probabilities assumed are in harmony with data from a U.S. bank explored in (Gans et al., 2010). Finally, based on insurance company data, we assume customers have an 80% probability to request assistance in non-resolved interactions in self-service systems.

The target acceptable waiting time in a queue that defines the service level is usually in the range of 20 to 30 seconds. For example, Pot et al. (2008), Cezik and L'Ecuyer (2008) and Avramidis et al. (2010) assume a 20 seconds threshold, while the insurance company data (explored in Moazeni and Andrade (2018) and Moazeni (2018)) contain attributes that flag transactions answered withing a 30 seconds limit. Contact centers often target service level of 80% (Cezik and L'Ecuyer, 2008; Pot et al., 2008; Avramidis et al., 2010; Chan et al., 2014). Therefore, the desired service quality level in our simulations is is set to be greater than 80% defined by a 30 seconds threshold.

6.4 Simulation Study

6.4.1 Experimental Design

We perform simulation executions for four different customer service support system configurations, based on the available communication channels and contact resolution capabilities. Table 6.2 describes the four considered system configurations.

Table 6.2: Experimental System Configurations

Contact Channel	**Telephone**		**Web**	
Query Resolution Capability	**A**	**B**	**A**	**B**
System Configuration 1	X	X		
System Configuration 2	X	X	X	
System Configuration 3	X	X		X
System Configuration 4	X	X	X	X

System Configuration 1 includes the single contact channel "Telephone" and can handle two query types. The other three system configurations involve a new channel "Web", which refers to digital channels. The Web channel has a higher contact resolution capacity and is, therefore, more efficient. System Configurations 2 and 3 consider a two-channel system design with the Web channel designed to solve only one of the two queries. System Configuration 4 assumes both Web and Telephone channels can solve both query types. We choose two channels and two query types to represent the simplest model that allows such system design changes to occur.

Experiments vary the channel allocation accuracy probability $\gamma_1 \in \{0.5, 0.55, 0.6, \ldots, 1.0\}$ and the contact routing accuracy probability $\gamma_2 \in \{0.6, 0.65, 0.7, \ldots, 1.0\}$ independently to assess their individual impact on the system. Varying γ_1 and γ_2 creates a total of 66 different scenarios. For each configuration and for each scenario (γ_1, γ_2), we perform 100 simulation runs for a period of 3000 minutes, similar approach described in Section 6.3.5. For each run, the first 1000 minutes are used as a warm-up period followed by 2000 minutes of data recording.

6.4.2 Simulation Results and Discussion

Figure 6.3 summarizes the main results of our simulation experiments. The graphs on the left and the right correspond, respectively, to scenarios varying the channel allocation accuracy and the contact routing accuracy. The lines show the optimal number of CSRs to solve the staffing problem (6.3.3) for each configuration. Each curve in Figure 3 corresponds to a system configuration in Table 6.2. The labeled marks show the relative percentage change in the number of agents necessary to maintain the target SL with respect to each scenario.

Let Ω_k be a set of all contextual parameters and input performance metrics defined in Section 6.3.2 and system configuration $k = \{1, 2, 3, 4\}$ as described in Table 6.2. Assume $N(\gamma_1, \gamma_2, \Omega_k)$ is the total number of CSRs to solve problem (6.3.3). Then, let

$$r_1 = \frac{N(\gamma_1, \gamma_2 = 0.6, \Omega_k)}{N(\gamma_1 = 0.5, \gamma_2 = 0.6, \Omega_k)} - 1, \qquad (6.4.1)$$

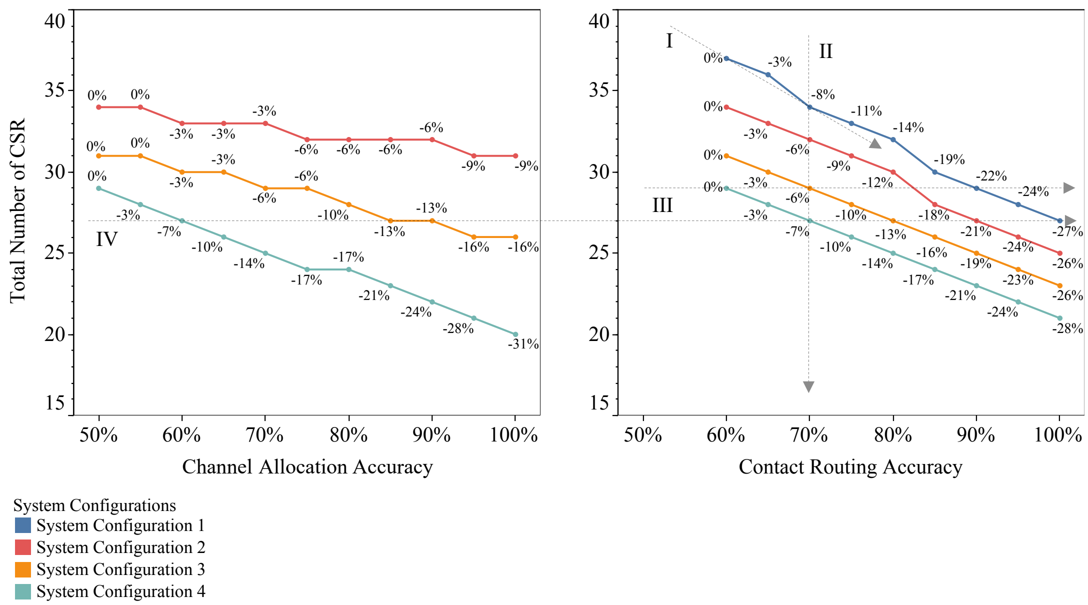


Figure 6.3: Simulation results: total number of customer service representatives necessary to maintain a 80% service level for distinct combinations of channel matching accuracy, contact routing accuracy, and system configurations.

and

$$r_2 = \frac{N(\gamma_1 = 0.5, \gamma_2, \Omega_k)}{N(\gamma_1 = 0.5, \gamma_2 = 0.6, \Omega_k)} - 1 \tag{6.4.2}$$

be the percent changes in the total number of CSRs shown, respectively, on the left and on the right charts of Figure 6.3.

Dashed lines I, II, III, and IV provide examples of how simulation results can guide operations managers to make decisions when designing operations strategies such as adopting data-driven technologies in the customer support system process. Four hypothetical situations discuss potential recommendations based on the results.

Example I

The company wants to reduce operational costs by investing in more efficient contact routing software. Assume the company only offers the option of Telephone contacts, similar to System Configuration 1. In this context, line I shows the impact of improving the contact routing accuracy on the number of CSR necessary to maintain service quality above 80%. If routing accuracy of a new software increases in performance from 60% to 70%, the number of agents reduce from 37 to 34, equivalent to 8% reduction in costs. In a simplistic estimation of the potential cost reduction, assuming a $30/hour per agent (including salary, training, infrastructure), 40 working-hours/week, and 52 weeks/year would expect $187,200 savings per year.

Example II

After deciding to invest in adopting technologies that improve contact routing accuracy to 70%, the company seeks to reduce the number of CSRs to 27. Line II illustrates how the company can meet this target by investing in new communication channels while maintaining the same contact routing performance. Enhancing contact classification and routing optimization algorithms to match customers to agents with greater than 70% accuracy becomes more complex and financially prohibitive. Therefore, investing in additional channel options becomes an attractive alternative.

Example III

In a third hypothetical scenario, the company wants to cut workforce costs by at least 20% to retain only 29 of the original 37 CSRs. Line III shows four alternatives to meet the new expenditure requirement:

1. Invest only in contact routing technologies to achieve accuracy of 90%, which is unlikely due to uncertainties in the service support process (e.g. customer inputs using closed-menu and speech-enabled systems, chatbots, and query classification mechanisms).

2. Invest in contact routing technologies to achieve accuracy of about 82.5% and

add a new web communication channel with the ability to solve type A queries.

3. Invest in contact routing technologies to achieve an accuracy of 70% and add a new web communication channel capable of solving type B queries.

4. Invest only to add a new web communication channel with the ability to solve both types of queries.

Example IV

Finally, we analyze how the company can eliminate 10 CSRs as a 27% reduction from the base case scenario with possible actions to enhance the contact routing accuracy, offer new communication channel with multiple configurations, or adopt data-driven technologies to match customers to the most efficient channel. Line IV guides the company's decision-maker with six alternative decisions:

1. Invest only in contact routing technologies to achieve accuracy of 100% which is unlikely due to uncertainties in the service support process.

2. Invest in contact routing technologies to achieve accuracy of 90% and add a new Web communication channel with the ability to solve type A queries.

3. Invest in contact routing technologies to achieve an accuracy of 80% and add a new Web communication channel capable of solving type B queries.

4. Invest in contact routing technologies to achieve an accuracy of 70% and add a new Web communication channel with the ability to solve both types of queries.

5. Invest in channel allocation technologies to achieve an accuracy of 85% and add a new Web communication channel capable of solving type B queries.

6. Invest in channel allocation technologies to achieve an accuracy of 60% and add a new Web communication channel that can solve both types of queries.

Results reveal an immediate and considerable reduction in staffing costs with the addition of the Web channel, even without modifying the channel allocation and contact routing accuracy. Apart from the scenario considering both Telephone and Web channels with capabilities to solve query types A and B, changes in the contact routing accuracy lead to larger decreases in costs compared to the channel allocation accuracy. Yet, the combination of offering a new channel and improvements to customer-channel allocation and the contact routing efficiencies boosts its benefits.

The reduction in the number of CSRs present distinct patterns depending on the system design and Web channel configuration. Although there is a linear relationship between accuracy and number of agents, the rate at which the costs decrease is different in each scenario. The combination of both contact channels with full contact resolution capabilities is about twice as effective than having the Web channel with features to solve only the query type B, and more than three times compared to only having the ability to solve query type A. The mean CSR service time for query B is twice as long for query A to explain the results. As a result of the workforce optimization performed manually which only allows integer values, we notice flat levels in the graph on the left. Since the impact of higher channel allocation accuracy on the number of agents is relatively low, despite an increase in idleness, the number of agents only reduces when all constraints are satisfied.

The following section describes the main aspects and implications of the simulation results.

6.4.3 Implications for Practice in Operations Management

In short, Figure 6.3 enables a multidimensional trade-off analysis between lower-level and system-level design strategies, and costs that can support design operations decisions. We emphasize this study is a notional case, not intended to represent any real case or to provide evidence of results of cost reduction. Our purpose is to present a methodology to assess the trade-offs of data-driven methods in customer service support systems and demonstrate the value of the simulation approach for operations management. Programming libraries such as SimPy allow the seamless development of complex systems models and simulation experiments. The result is a powerful tool to support decision-making in operations systems due to its convenience to accommodate and test changes to operational processes.

Changes to contextual parameters such as arrival rate or service duration have a limited impact on the insights gained from our simulation experiments. A higher volume of customers joining the modeled system requires more computational resources to execute similar analyses or results in a larger number of agents to handle the extra workload, as shown in the model verification in Section 6.3.5. However, it does not change the dynamic of cost reduction observed from changing characteristics of the system design. In contrast, factors that would bring our model closer to reality, such as a greater variety of queries, agents with more complex training types, and how contact channels accommodate solutions to different query types, have the potential to transform the system behavior.

Although the recommendations discussed in Section 6.4.2 suggest significant reductions in operating costs from adopting new technology options, it is important to highlight the limitations of these conclusions based on model assumptions. First, our analysis was performed under steady-state conditions and assuming a single mean arrival rate. Workforce planning in contact centers is typically optimized for periods of one hour, half an hour, or even fifteen minutes. There is high volatility in contact volume throughout the day, which requires a dynamic workforce scheduling to achieve optimal operating costs. Therefore, models should incorporate mechanisms to account for the intra-day seasonality and adapt the ideal number of agents to allow a more realistic assessment of the advantages of adopting the aforementioned system design strategies.

Three other important factors to consider when extending our model to real-world applications include: accounting for greater diversity of customers types with different priority levels; allowing customers to have more than one query in the same contact; and introducing outbound interactions. In our experiments, we restricted customer types to only two, given their communication channel preferences, and no priority was imposed, although the model is structured and prepared for it. In reality, there may be multiple instances of customer classification and service priorities. Second, our simulation assumes consumers have only one query per interaction. It would be interesting to incorporate the possibility of customers having more than one problem which could interfere with the intra-agent transfer dynamics and, consequently, the contact distribution mechanisms. Finally, our model replicates an inbound customer support system but many contact centers also make outbound communications. Agents often follow up on previous inbound contacts, act proactively, or perform telemarketing activities. This could change the system processes, impact CSR utilization rates and skill sets, and workforce planning.

6.5 Conclusions

Data-driven technology is changing businesses models, processes, operations, product development, and the dynamics between customers and organizations. Companies spend millions to develop technologies that generate significant financial returns, either as savings or additional revenues. The search to adopt digital technologies that leverage data to support better business decisions has become a necessity for companies to remain competitive in their fields.

In customer services, technology has made it possible to expand customer communication channels beyond the telephone. Call centers became contact centers by offering multiple service channels such as mobile, social networks, e-mail, SMS, chat. Contact centers are the direct reflection of the digital transformation and the new type of consumer it has created. Customer support services incorporate AI, machine learning, and big data analytics techniques to provide a cohesive and fluid user experience throughout the customer journey across all points of contact with the company.

This paper contributes a simulation-based approach to analyze how the infusion of data-driven technologies into customer service operational processes impact system-level performance to better inform targeted investments in digital technology. The main contribution draws on the synthesized approach to evaluate impacts, at the system level, of lower-level system design strategies to adopt digital technologies. From a systems perspective, we compare the benefits of implementing different technologies on an equal footing. Simulation experiments assess the tradeoffs between investing in a more efficient communication channel and enhanced customer classification mechanisms and potential cost savings from reduced workforce levels in a multichannel customer support system. We implemented the simulation model using

SimPy Python 3.7. The source code is available in an online repository for public use and future research.

Our model allows different types of customers with multiple query types, distinct individual priorities and preferences. The model can be adapted to incorporate additional contact channels, each with specific characteristics. Examples of configuration attributes include the channel capacity and availability, service time based on customer or request type, contact resolution efficiency, and the set of self-service features offered. The model also allows changing factors that interfere with the contact routing process by altering the contact routing accuracy. Finally, it is possible to define the characteristics of the CSRs such as different skill sets, average service time (e.g., with the problem or customer type distinction), as well as problem-solving probability.

Chapter 7

Conclusions and Directions for Future Research

This chapter unfolds the research questions and summarizes the conclusions and key contributions. Finally, it discusses the implications of this work and directions for future research.

7.1 Research Summary

Organizations around the world benefit from integrating digital technologies into their operations strategies to improve operational efficiency, support robust decision-making and increase agility to respond to market trends, demands, and innovations. These technologies enable creating business value from that data and redesigning their business models. The data generated and collected from digital technologies allow for in-depth and real-time control of business processes from product design to customer support services. However, there are scarce literature bridging operations, more technical, and strategic management levels to understand the tangible benefits of leveraging big data analytics in their business processes. Therefore, the central research question driving this dissertation is

How can data analytics improve decision-making processes in customer support services operations management?

To address the main question, we pose the following three secondary research questions. The key contributions of this dissertation derive from studies we conducted to explore and answer each question.

1. *What are the forces that influence contact center behavior and how can system reliability measurement achieve consistent desired behavior?*

Chapter 3 uses systems thinking tools to address this question and develop a contact center operations management system model. The system model incorporates concepts of enterprise transformation, and describes its components and functions. As reliability is one of the most important dimensions of service quality, we propose a mathematical model to measure the reliability of the system based on the new system design. The developed system architecture and reliability metric are general enough and can be implemented in any industry. An empirical analysis using data from a major insurance company is provided to illustrate the applicability of the system design, reliability, and performance control using a control chart.

2. *How can a feature-based model leverage multichannel transaction level customer support data to predict customer behavior?*

This research addressed this question in two parts exploring two different problems in operations. Chapters 4 and 5 transform the data gathered from contact centers of a major insurance company into managerial insights, respectively, for operations planning and real-time process automation.

Chapter 4 develops a model to predict future Telephone queries by an individual customer within the next thirty days, based on the multichannel data from insurance contact centers. The model incorporates information related to the past Web activities of an individual customer to predict his future telephone queries. Various characteristics related to the customer segment, recency and frequency of customer interactions, and cross-class features are considered. We find evidence that some of the recent web activities of a policyholder significantly increases the probability that the policyholder would make a telephone call in the next 30 days. In addition, recency and frequency of contacts impact the probability of the policyholder's call, for a specific set of reasons for past contacts. In addition, recency and frequency of contacts impact the probability of the policyholder's call, for a specific set of reasons for past contacts

Chapter 5 studies the effective features for determining the outcome of a call arriving at the self-service Interactive Voice Response system. Our findings indicate that distinct channels of contact, average temperature, customer types, and specific intents are relevant features to determine how likely the call will be transferred. The location attribute has a significant impact on identifying the model active set of features but there is no significant correlation between the customer country's regional origin and the selected features. While the developed algorithms are scalable and can be adopted for other big data analytics problems, the results can provide various managerial insights into customers' behavior seeking to make more effective use of customer data and segmentation.

3. *How can operations managers assess the trade-offs between investing in data-driven technologies that improve customer matching and classification and their operational cost savings?*

To address this last question, Chapter 6 investigates potential cost savings from introducing more accurate classification methods to direct customers to more efficient self-service communication channels, and to improve customer-agent matching based on detailed data from customers, processes, and agents. We conduct simulation experiments that allow anticipate the effects of the changes in the system and infer how lower-level system designs affect system-level performance. The result discussions include practical examples of how operational managers could use the experimental information to make strategic operational decisions when adopting data-driven technologies.

7.2 Key Contributions

This research contributions to broader literature in customer support operations management result from the studies thoroughly developed in Chapters 3 to 6:

1. Developed a contact center operations management system architecture that supports a novel mathematical model of the system reliability for real-time measurement of service quality. The system architecture is general and provides a holistic understanding of customer support services. The proposed probabilistic formulation for the system reliability model capture the uncertainties of the service processes and is adaptable to different scenarios and configurations of contact centers.

2. Developed feature-based models that leverage multichannel transaction-level customer support data to predict customers behavior at the individual level for resource planning and real-time automation.

3. Demonstrated how to assess the impacts, at the system level, of changing the system design when adopting data-driven technologies for contact center operational processes. The simulation-based approach allows to anticipate the effects of technology infusion and is used as a tool to inform strategic decision, bridging strategy and operations.

7.3 Limitations

This section critically reviews the results of this research and discusses its implications for the broader literature on operations management, and data analytics. Our analysis is extensive to managers and decision-makers in companies that can benefit from our insights to motivate systematic evaluation of changes to their business processes.

Chapters 4 and 5 develop feature-based models with a similar approach, albeit with different goals. the first study, the main objective is to predict a future customer call, information that can support the planning of resource allocation or creation of proactive measures. The second study seeks to predict the request for human assistance after the customer is not successful in the self-service stage of the support service. A direct application of this model is to improve the automatic transfer mechanisms in real-time, based on past information and data collected during the service process. However, a common limitation to both studies is the result of the technique used to create predictive models.

The method used in both cases combines logistic regression with boostrap and presents positive and negative points. Among the advantages, we include the ability to interpret the results of the model that allows drawing conclusions beyond the estimated likelihood. The model's coefficients allows to infer the individual impacts of input variables. The variables associated with the coefficients seek to translate customer behavior and relevant characteristics of the service processes. Thus, analysing the coefficients helps to identify problems and opportunities for improvements in operational processes. Our model results uncover customer behavior insights that are challenging to extract from simpler data analyzes and descriptive statistics. Furthermore, the models identify effective features influencing operations processes that can be used as input for optimization models seeking to enhance services performance.

Nevertheless, the predictive power of regression models is limited. The imbalanced data and the non-linearity of the underlying customer behavior function contribute to lower predictive performance. Section A in the appendix discusses the challenges and how to work with imbalanced datasets. Furthermore, the models do not consider non-transaction customer data. As a result, prediction performance for customers that will use the contact center for the first time gets particularly compromised. Richer data from external sources of the customer support service processes can allow adding more information about a customer with no transaction history with the contact center and potentially enhance the models' prediction power. Finally, the linear characteristics of logistic regressions cannot implicitly detect complex nonlinear relationships between input and output. Artificial Neural Network or Random Forest are alternative machine learning techniques to solve the non-linearity problem and, consequently, improve predictive performance. However, relying on these methods compromise the interpretative power of the impacts of the inputs.

Chapter 6 seeks to connect the lessons from the previous chapters by creating a framework for evaluating the implementation of data analytics from a view with a higher level of abstraction. The experimental approach models customer service operations and simulates various scenarios of infusing data-driven technologies into operational processes. Moreover, the study develops a customer support service system that simplifies a real system and builds upon several assumptions that remain constant throughout the simulations. In contrast, real-life customer service operations are dynamic. For instance, customer arrival rates and the number of customer service representatives vary during hours of the day, and customers may have multiple queries at a time. Our models assumes that customers only have one of two query types configured, and there is no cross-trained agents capable to resolve both queries. Additionally, our model replicates an inbound contact center and, therefore, does not account for outbound communications. Incorporating interactions initiated from inside the contact center can change the system process, impact agents' utilization rates, and workforce planning. Despite the uncertainties involved in the simulation model, our approach allows for a quick and versatile analysis that provides extract valuable insights for decision making at the strategic level.

7.4 Future Research

Future research can explore how to use customer data analytics to design smart mechanisms that improve service quality, enhance customer experience, and reduce operations costs in multichannel customer service support systems. We propose the following projects:

Project I - Proactive Customer Management

After accurately identify customers with high probabilities to interact with the company and predict demand (Chapters 4 and 5), a smart mechanism to optimize the use of workforce with the available information can be developed. The company can allocate CSR's to contact customers at non-peak hours by proactively contacting customers to resolve potential future queries. The expectation is a demand reduction in peak hours, while also allowing for more personalized customer service. The CSR should contact customers with the background information prepared beforehand to enable more effective and faster services. Additional benefits include reduced waiting times during peak hours, higher CSR utilization, and lower operational costs.

Project II - Incentive Programs

A common practice in the service industry is to use reward programs as a promotional tool to attract and develop consumer loyalty, and also as a method of capacity and demand management (Byung-Do Kim, Mengze Shi, 2001; Kim et al., 2004). Examples of sectors that make use of these programs include hotels, car rentals, telecommunications companies, and airlines. The congestion of contact centers is expected to reduce through the design of data-driven incentive programs that encourage customers to use alternative communication channels. The goal is to create programs that dynamically optimize rewards in real-time based on customer and operations data, such as the type of service required, date and time of contact, and customer behavior related features.

Project III - Smart Customer Management Platforms

This study seeks to improve customer service operations by developing a data-driven mechanism to manage customer service duration while using a self-service system, based on the accurate prediction of the transference likelihood, which was the model proposed in Chapter 5. The customer management platform must be flexible to identify the need to quickly transfer the customer to a CSR or have him/she navigating for a longer period through the support system. Eventually, if the system identifies that a potential customer with a high hanging up probability, it would be ideal to quickly transfer this customer to a skilled CSR and increase the chances of making a sale. In contrast, in high call demand hours, keeping the customer for a longer time in the IVR system may increase the likelihood to solve the original problem and eliminate transference chances, or simply reduce the queue duration, which is one of the key performance metrics for call center services.

Appendix A

Challenges in Imbalanced Data Classification

This Appendix discusses challenges in classification problems in presence of highly imbalanced classes. Section A.1 describes the problem, and Section A.2 discuss the main strategies in the machine learning literature to deal with highly imbalanced classes in classification problems.

A.1 Background

Imbalanced learning problem occurs when there is a large disproportion between the number of examples in each class in the data under investigation. As a result, the predictive model incorrectly classifies the minority class examples due to bias towards majority class (Prati et al., 2009). The problem is present in real-world situations such as anomaly detection (Khreich et al., 2010), medical diagnosis (Chen, 2016), and microarray research (Yu et al., 2013). The problem of learning from imbalance data gained attention over the years (see Chawla et al. (2002); He and Garcia (2009); Sun et al. (2009); López et al. (2013)). More recently, challenges for imbalanced classification in Big Data problems were also discussed in Fernández et al. (2017).

Algorithm performance becomes strongly compromised when the learning problem depends on imbalanced data sets since standard machine learning algorithms assume balanced classes or equal misclassification costs (He and Garcia, 2009). The performance of machine learning algorithms on classification problems is typically evaluated by accuracy, or the percentage of correctly classified samples (Hastie et al., 2001). Given the confusion matrix in Table A.1, the overall accuracy (ACC) for a two classes problem is calculated by

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}. \tag{A.1.1}$$

Table A.1: Confusion matrix.

Classes		**Actual**	
		0	1
Predicted	0	True Negative (TN)	False Positive (TP)
	1	False Negative (FN)	True Positive (TP)

Although the classification error/accuracy is typically used to evaluate the classifier algorithm performance, the accuracy rate does not take into account the prevalence of the majority class (Chawla et al., 2002). For instance, if the majority class corresponds to 90% of the data, it is trivial to build a classifier that is accurate 90% of the time, simply by predicting the majority class. This becomes an issue especially when the minority class is the one of most interest (e.g. population with a particularly rare disease), and a classifier that fails to identify the minority class is useless.

A.2 Approaches for Class Imbalance Problems

A.2.1 Main Categories

According to Fernández et al. (2017), there are three categories of techniques that seek to overcome the challenge of using imbalanced data sets in learning algorithms: (1) methods that modify existing data seeking to balance class distributions (Chawla et al., 2002; Batista et al., 2004), (2) methods that adapt algorithm learning processes towards minority classes (Ramentol et al., 2015), and (3) cost-sensitivity solutions (Domingos, 1999; Veropoulos et al., 1999; Hall et al., 2009). An additional category comprises the ensemble solutions that combine distinct categories.

The first category consists of sampling methods that either delete some of the data related to the majority class, or replicate examples of the minority class while seeking to balance the data set. These strategy types are referred to as under-sampling and over-sampling, and can be executed using distinct mechanisms (Japkowicz, 2000; Haibo He et al., 2008; He and Garcia, 2009). However, these strategies can eliminate important examples for learning or cause overfitting (Batista et al., 2004). To address these limitations, different authors proposed methods using heuristics to insert or remove examples have been proposed. The Synthetic Minority Oversampling Technique (SMOTE) is a type of over-sampling method that interpolates minority class examples and their k-nearest neighbors to create new data points (Chawla et al., 2002). The SMOTE technique can be combined with other techniques such as the Edited Nearest Neighbors (ENN) and Tomek Links creating hybrid methods. For further details refer to Chawla et al. (2002); Batista et al. (2003); Chawla et al. (2003); Batista et al. (2004); Lusa et al. (2013)

The second category comprises solutions that adapt the existing classification algorithms to improve minority class performance (Wu and Chang, 2003; Lin et al., 2002; Bernadó-Mansilla and Garrell-Guiu, 2003; Huang et al., 2006). For instance, Wu and Chang (2003) introduces a class-boundary-alignment algorithm for Support Vector Machines (SVMs) adjusts the class boundary either by transforming the kernel function when the training data can be represented in a vector space, or by modifying the kernel matrix when the data do not have a vector-space representation. In Cieslak et al. (2012), a decision tree technique uses the Hellinger distance, a measure of distributional convergence (Kailath, 1967), as the splitting criterion instead of entropy. The authors demonstrate the superiority over common used sampling and ensemble methods such as C4.4 (Provost and Domingos, 2003) and CART(Gini)

(Breiman et al., 1984).

The third category consists of the cost-sensitivity solutions. Cost-sensitive methods combine both algorithm and data level approaches to incorporate different misclassification costs for each class in the learning phase. It biases the classifier towards the minority class assuming higher misclassification costs for the class and seeking to minimize the total cost errors of both classes (Galar et al., 2012). Misclassification costs are typically described by a cost matrix. The cost of a classification error of a positive instance, which belongs to the minority class, is greater than the cost of a classification error of a negative instance. As a result, the weights of both classes in the training stage are equalized. Domingos (1999) proposes a procedure for making error-based cost-sensitivity classifiers. It combines the predictive benefits of bagging with a comprehensive model for cost-sensitivity prediction (Witten et al., 2016). The method relabels the training examples with their estimated minimal-cost class and apply the error-based learner to the new training set. The expected cost of predicting that x belongs to class i is estimated by minimizing the Bayes conditional risk. Several studies also combine cost-sensitive solutions to algorithm level techniques to achieve better classification performances (Masnadi-Shirazi and Vasconcelos, 2010; Zong et al., 2013; Iranmehr et al., 2019).

Table A.2 compiles the main observations from the literature and provides a comparison of the characteristics of the three groups of techniques commonly use to deal with imbalance data problems.

A.2.2 Ensemble Techniques

A common challenge in machine learning problems is to increase algorithm classification and prediction performance. There are techniques that make use of ensembles of classifiers to increase the accuracy of a single classifier (Polikar, 2006; Rokach, 2010).

These strategies train different classifiers and combining their decisions to output a single class label (Galar et al., 2012). Data level and cost-sensitivity approaches are typically combined with ensemble learning algorithms (Wang and Yao, 2009; Seiffert et al., 2010; Wang and Pineau, 2016; Rayhan et al., 2017).

Ensemble methods are defined as "a class of highly successful machine learning algorithms which combine several different models to obtain an ensemble which is, hopefully, more accurate than its individual members" (Sołtys et al., 2015). In addition to search for the maximization of the generalization ability, machine committees are motivated by great availability of computational resources to enable the synthesis of multiple proposals for solutions for all or part of the problem being investigated, and demonstration, through the free lunch theorem, that there are no generic machine learning model that, on average, present better performance than any other model for any class of problems type (Wolpert and Macready, 1997). Therefore, part of the strategy to build an ensemble system is the need to combine the outputs of individual classifiers so the correct decisions are amplified, and incorrect ones are canceled out (Polikar, 2006).

Popular ensemble methods used in machine learning are based on *bagging*, *boosting*, and *stacking* approaches (Breiman, 1996; Freund and Schapire, 1997). In both bagging and boosting procedures, the training data is manipulated and multiple classifiers are generated through the training set sampling. Furthermore, the two methods use voting for classification or averaging for numeric prediction to combine output of individual models. However, the boosting algorithm attributes weights to the samples, while bagging works by re-sampling subsets of the training data. "The effect of combining multiple hypotheses can be viewed through a theoretical device known as the *bias–variance decomposition*" (Witten et al., 2016). The improvement in performance is usually a result of variance reduction.

Bagging, an acronym for "bootstrap aggregating", are intuitive procedures that randomly draw subsamples with replacement from the learning set and generates multiples versions of the predictors for each subset. The predictors are then aggregated by plurality voting when predicting a class, or averages over the versions when predicting a numerical outcome (Breiman, 1996).

A straightforward advantage of bagging is the intuitiveness which makes it easy to implement. Moreover, it can be applied to wide variety of fitting methods such as convention regression, logistic regression, and discriminant function analysis (Berk, 2005). Disadvantages include the failure to learning algorithms whose output is insensitive to small changes in the input, and the dependency between fitted values across samples created by the bootstrap samples with replacement. The nonadditivity of bias and variance can also influence on the result of the bagging procedure.

Differently than Bagging, Boosting algorithms repeatedly run the learning algorithm but focus on misclassification examples. In addition, boosting procedures are iterative and each new model is influenced by the performance of those previously built, while in bagging models are developed separately (Witten et al., 2016). The main idea is to train a strong classifier by combining weak classifiers, those whose error rate is just slightly better than 50% (Hastie et al., 2001). AdaBoost (Freund and Schapire, 1997) is the more general version of the original boosting algorithm and have received most attention . For each example should be provided a weight, where misclassified examples get higher weight, and base classifiers get a weighted vote depending on its accuracy. The final classification is obtained by plurality voting.

Unlike *bagging* and *boosting* that draw m multiple samples from the training data set and apply m models of the same type, *Stacking, or stacked generalization*, combines the output of different types of base classifiers (Wolpert, 1992). The output of first-level classifiers serve as input for a second-level learning algorithm. Stacking methods attempt to learn which of the base classifiers are the most reliable and to find out the best way to combine their outputs using a meta classifier. In summary, the learning property in stacking algorithms takes place in the meta level, at which base level classifiers are combined Sikora et al. (2015).

Zhang et al. (2018) proposes a novel method that combines stacking and inverse random undersampling (SIRUS) techniques to improve classification performance for imbalance problems. The study compares the new method performance against three other methods: inverse random under sampling (IRUS) Tahir et al. (2012), Chan's method Philip and Chan (1998), and decision tree (C4.5). The SIRUS procedure is two-fold. The first phase consists of repeating two steps of under sampling the majority class and apply a group of classification algorithms to learn the boundaries between the two classes. Finally, use the output of each algorithm as input data for a meta classifier. Experimental results show that SIRUS outperforms the benchmark methods in 11 out of 18 tests using distinct data sets based on the AUC. The proposed method also presents the exact same performance for one data set, and is the second best in other 6 tests.

Table A.3 compiles the main observations from the literature and provides a summary and a comparison between the characteristics of Stacking, Ensemble, and Boosting techniques.

Table A.2: Comparison of techniques for imbalance data problems.

	Data Level		**Algorithm Level**	**Cost-sensitivity**
Objective	Replicate existing minority class examples or eliminate examples of the majority class.	Create new examples of the minority class to achieve balance with majority class.	Adapt algorithm learning process towards minority class.	Bias the classifier by assuming higher misclassification costs for minority class.
Data Partition/ Generation	Random	Heuristic	-	-
Advantages	Flexible, can generate any possible distribution between classes; independent of the underlying classifier, can be combined with other strategies	Avoid overfitting by generating new minority class examples by inter-polation; independent of the under lying classifier, can be combined with other strategies	Do not eliminate data; appropriate for nonlinear problems.	Do not eliminate data; can be effective for Neural Networks training process
Disadvantages	May introduce bias; can eliminate potentially useful data, increase likelihood of over-fitting, oversampling can increase processing time	Over-generalization with occurrence of overlapping between classes; can increase variance	More sophisticated, requires a good insight into the modified learning algorithm and a precise identification of reasons for its failure in mining skewed distributions; Not applicable to any method	Challenging to determine a cost representation of a given domain, may lead to over-fitting, since misclassifications are often unknown; Poor scalability
References	Japkowicz (2000) Batista et al. (2004) He and Garcia (2009) Galar et al. (2012) Tahir et al. (2012)	Chawla et al. (2002) Wang and Japkowicz (2004) Haibo He et al. (2008) Zio (2009) Hu et al. (2009) Seiffert et al. (2010)	Lin et al. (2002) Batista et al. (2003) Wu and Chang (2003) Ertekin et al. (2007) Zong et al. (2013) Iranmehr et al. (2019)	Domingos (1999) Veropoulos et al. (1999) Hall et al. (2009) Galar et al. (2012) Wang and Pineau (2016) Zhao et al. (2019)

Table A.3: Comparison between families of ensemble methods.

Family	**Bagging**	**Boosting**	**Stacking**
Data Partitioning	Random drawing with replacement	Misclassified samples are given higher preference	Various
Objective	Minimize variance	Increase predictive power	Both
Example of Ensambling Function	(Weighted) average	Weighted majority vote	Logistic Regression, Random Forest, Multilayer Perceptron, Support Vector Machine
Advantages	Intuitiveness; can applied to wide variety of fitting methods; easy implementation	Ability to build a powerful classifier from very simple ones; Flexibility, since different loss functions can be used; Easy implementation with no parameters to be tune; Good generalization as the error converges to zero.	Potential to ensemble diverse group of strong base learners; Well-suited for big data since models can be trained independently.
Disadvantages	Fails to learn from algorithms whose output is insensitive to small changes in the input Boostrap samples with replacement can create dependency between fitted values across samples	Dependence on data; Dependence on weak learner; Susceptible to uniform noise; Instances with low weight may not be included into the re-sampled data set.	Difficult to analyze results and asses the exact individual contribution of covariates, and require large amount of computation.
References	Breiman et al. (1996) He and Garcia (2009) Khoshgoftaar et al. (2011) Galar et al. (2012) Altman and Krzywinski (2017)	Schapire (1990) Sun et al. (2009) Seiffert et al. (2010) Galar et al. (2013) Witten et al. (2016) Rayhan et al. (2017)	Wolpert and Macready (1997) Sesmero et al. (2015) Sikora et al. (2015) Zhang et al. (2018)

Bibliography

Addo-Tenkorang, R. and P. T. Helo (2016). Big data applications in operations/supply-chain management: A literature review. *Computers and Industrial Engineering 101*, 528–543.

Agarwal, R., G. Gao, C. DesRoches, and A. K. Jha (2010). Research commentary—the digital transformation of healthcare: Current status and the road ahead. *Information Systems Research 21*(4), 796–809.

Ahghari, M. and B. Balcioglu (2009). Benefits of cross-training in a skill-based routing contact center with priority queues and impatient customers. *IIE Transactions 41*(6), 524–536.

Aier, S. (2014). The role of organizational culture for grounding, management, guidance and effectiveness of enterprise architecture principles. *Information Systems and e-Business Management 12*(1), 43–70.

Aier, S. and B. Gleichauf (2010). Application of Enterprise Models for Engineering Enterprise Transformation. *Journal of Enterprise Architecture 1*(5), 56–72.

Aksin, Z., M. Armony, and V. Mehrotra (2007). The modern call-center: A multi-disciplinary perspective on operations management research. *Production and Operations Management 16*(6), 665–688.

Aldor-Noiman, S., P. D. Feigin, and A. Mandelbaum (2009). Workload forecasting for a call center: Methodology and a case study. *The Annals of Applied Statistics 3*(4), 1403–1447.

Ali, A. R. (2011). Intelligent call routing: optimizing contact center throughput. In *Proceedings of the Eleventh International Workshop on Multimedia Data Mining*, pp. 1–9.

Ali III, L. F. (2010). *A Call Center Simulation Study: Comparing the Reliability of Cross-Trained Agents to Specialized Agents*. Master's thesis, University of Tennessee.

Altman, N. and M. Krzywinski (2017). Points of Significance: Ensemble methods: Bagging and random forests. *Nature Methods 14*(10), 933–934.

Anderson, E. W., C. Fornell, and D. R. Lehmann (1994). Customer Satisfaction, Market Share, and Profitability: Findings from Sweden. *Journal of Marketing 58*(3), 53.

Anderson, E. W., C. Fornell, and S. K. S. Mazvancheryl (2004). Customer Satisfaction and Shareholder Value. *Journal of Marketing 68*(4), 172–185.

Andrade, R. and S. Moazeni (2017). Feature selection for voice self-service transfer rate in insurance call centers. In *Presented at INFORMS Annual Meeting Conference*, Houston, TX.

Andrade, R., S. Moazeni, and J. E. Ramirez-Marquez (2020). A systems perspective on contact centers and customer service reliability modeling. *Systems Engineering 23*(2), 221–236.

Andrews, M., X. Luo, Z. Fang, and A. Ghose (2015). Mobile ad effectiveness: Hyper-contextual targeting with crowdedness. *Marketing Science 35*(2), 218–233.

Antipov, A. and N. Meade (2002). Forecasting call frequency at a financial services call centre. *Journal of the Operational Research Society 53*(9), 953–960.

Armony, M. and A. R. Ward (2010). Fair dynamic routing in large-scale heterogeneous-server systems. *Operations Research 58*(3), 624–637.

Atlason, J., M. A. Epelman, and S. G. Henderson (2008). Optimizing call center staffing using simulation and analytic center cutting-plane methods. *Management Science 54*(2), 295–309.

Aven, T. (2012). *Reliability and risk analysis*. Springer Science & Business Media.

Avramidis, A. N., W. Chan, M. Gendreau, P. L'Ecuyer, and O. Pisacane (2010). Optimizing daily agent scheduling in a multiskill call center. *European Journal of Operational Research 200*(3), 822–832.

Avramidis, A. N., A. Deslauriers, and P. L'Ecuyer (2004). Modeling daily arrivals to a telephone call center. *Management Science 50*(7), 896–908.

Avramidis, A. N. and P. L'Ecuyer (2005). Modeling and simulation of call centers. In *Proceedings of the 37th Conference on Winter Simulation*, WSC '05, pp. 144–152. Winter Simulation Conference.

Baines, T. S., H. W. Lightfoot, O. Benedettini, and J. M. Kay (2009). The servitization of manufacturing: A review of literature and reflection on future challenges. *Journal of Manufacturing Technology Management 20*(5), 547–567.

Barrow, D. and N. Kourentzes (2018). The impact of special days in call arrivals forecasting: A neural network approach to modelling special days. *European Journal of Operational Research 264*(3), 967–977.

Batista, G. E., A. L. Bazzan, and M. C. Monard (2003). Balancing training data for automated annotation of keywords: a case study. In *WOB*, pp. 10–18.

Batista, G. E., R. C. Prati, and M. C. Monard (2004). A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter 6*(1), 20–29.

Bayuk, J. L. and B. M. Horowitz (2011). An architectural systems engineering methodology for addressing cyber security. *Systems Engineering 14*(3), 294–304.

Berk, R. A. (2005). Data mining within a regression framework. In *Data Mining and Knowledge Discovery Handbook*, pp. 231–255. Springer.

Berman, S. J. (2012). Digital transformation: opportunities to create new business models. *Strategy & Leadership 40*(2), 16–24.

Bernadó-Mansilla, E. and J. M. Garrell-Guiu (2003). Accuracy-based learning classifier systems: models, analysis and applications to classification tasks. *Evolutionary Computation 11*(3), 209–238.

Bertsimas, D., N. Kallus, and A. Hussain (2016). Inventory Management in the Era of Big Data. *Production and Operations Management 25*(12), 2006–2009.

Bi, J., K. Bennett, M. Embrechts, C. Breneman, and M. Song (2003). Dimensionality reduction via sparse support vector machines. *Journal of Machine Learning*

Research 3(Mar), 1229–1243.

Blois, K. J. (1984). Productivity and effectiveness in service firms. *The Service Industries Journal 4*(3), 49–60.

Boardman, J. and B. Sauser (2008). *Systems thinking: Coping with 21st century problems*. CRC Press.

Boardman, J. and B. Sauser (2013). *Systemic thinking: building maps for worlds of systems*. John Wiley & Sons.

Bodur, M. and J. R. Luedtke (2017). Mixed-integer rounding enhanced benders decomposition for multiclass service-system staffing and scheduling with arrival rate uncertainty. *Management Science 63*(7), 2073–2091.

Borst, S. C., A. Mandelbaum, and M. Reiman (2000). Dimensioning large call centers. *Operations Research 52*, 17–34.

Boulding, W., R. Staelin, M. Ehret, and W. J. Johnston (2005). A Customer Relationship Management Roadmap: What Is Known, Potential Pitfalls, and Where to Go. *Journal of Marketing 69*(4), 155–166.

Boyer, K. K., R. Hallowell, and A. V. Roth (2002). E-services- operating strategy—a case study and a method for analyzing operational benefits. *Journal of Operations Management 20*(2), 172–188.

Bradic, J. et al. (2016). Randomized maximum-contrast selection: Subagging for large-scale regression. *Electronic Journal of Statistics 10*(1), 121–170.

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition 30*(7), 1145–1159.

Breiman, L. (1996). Bagging predictors. *Machine learning 24*(2), 123–140.

Breiman, L. et al. (1996). Heuristics of instability and stabilization in model selection. *The Annals of Statistics 24*(6), 2350–2383.

Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen (1984). *Classification and Regression Trees*. CRC press.

Brocchi, C., D. Grande, K. Rowshankish, T. Saleh, and A. Weinberg (2018). Designing a data transformation that delivers value right from the start. *McKinsey on Payments - Special Edition on Advanced Analytics in Banking 11*(28), 1–8.

Brown, L., N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, and L. Zhao (2005). Statistical analysis of a telephone call center: A queueing-science perspective. *Journal of the American Statistical Association 100*(469), 36–50.

Brown, T. J., G. A. Churchill, and J. P. Peter (1993). Improving the measurement of service quality.

Brynjolfsson, E. and K. McElheran (2016). The rapid adoption of data-driven decision-making. *American Economic Review 106*(5), 133–139.

Buzzel, R. D. (1990). Commentary on "Unobservable Effects and Business Performance". *Marketing Science 9*(1), 86–87.

Byung-Do Kim, Mengze Shi, K. S. (2001). Reward Programs and Tacit Collusion. *Marketing Science 20*(2), 99–120.

Campbell, D. and F. Frei (2010). Cost structure, customer profitability, and retention implications of self-service distribution channels: Evidence from customer behavior in an online banking channel. *Management Science 56*(1), 4–24.

Census (2017a). U.S. Census Bureau. Available at the Census website `https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf`.

Census (2017b). U.S. Census Bureau. Available at the Census website `https://www.census.gov/geo/reference/state-area.html`.

Census (2017c). U.S. Census Bureau, 2011-2015 American Community Survey 5-Year Estimates.

Cezik, M. T. and P. L'Ecuyer (2008). Staffing multiskill call centers via linear programming and simulation. *Management Science 54*(2), 310–323.

Chakrabarti, C. and G. F. Luger (2015). Artificial conversations for customer service chatter bots: Architecture, algorithms, and evaluation metrics. *Expert Systems with Applications 42*(20), 6878–6897.

Chan, W., G. Koole, and P. L'Ecuyer (2014). Dynamic call center routing policies using call waiting and agent idle times. *Manufacturing & Service Operations Management 16*(4), 544–560.

Changnon, S. A., D. Changnon, E. R. Fosse, D. C. Hoganson, R. J. Roth Sr, and J. M. Totsch (1997). Effects of recent weather extremes on the insurance industry: major implications for the atmospheric sciences. *Bulletin of the American Meteorological Society 78*(3), 425–435.

Channouf, N., P. L'Ecuyer, A. Ingolfsson, and A. N. Avramidis (2007). The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta. *Health Care Management Science 10*(1), 25–45.

Chawla, N. V., K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer (2002). Smote: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research 16*, 321–357.

Chawla, N. V., A. Lazarevic, L. O. Hall, and K. W. Bowyer (2003). Smoteboost: Improving prediction of the minority class in boosting. In *European Conference on Principles of Data Mining and Knowledge Discovery*, pp. 107–119. Springer.

Chen, J., Y. Tao, H. Wang, and T. Chen (2015). Big data based fraud risk management at Alibaba. *The Journal of Finance and Data Science 1*(1), 1–10.

Chen, Q. G., S. Jasin, and I. Duenyas (2016). Real-time dynamic pricing with minimal and flexible price adjustment. *Management Science 62*(8), 2437–2455.

Chen, Y., H. Chen, A. Gorkhali, Y. Lu, Y. Ma, and L. Li (2016). Big data analytics and big data science: a survey. *Journal of Management Analytics 3*(1), 1–42.

Chen, Y.-S. (2016). An empirical study of a hybrid imbalanced-class dt-rst classification procedure to elucidate therapeutic effects in uremia patients. *Medical & Biological Engineering & Computing 54*(6), 983–1001.

Choi, T.-M., S. W. Wallace, and Y. Wang (2018). Big data analytics in operations management. *Production and Operations Management 27*(10), 1868–1883.

Chong, A. Y. L., E. Ch'ng, M. J. Liu, and B. Li (2017). Predicting consumer product demands via Big Data: the roles of online promotional marketing and online reviews. *International Journal of Production Research 55*(17), 5142–5156.

Chu-Carroll, J. and B. Carpenter (1999). Vector-based natural language call routing. *Computational linguistics 25*(3), 361–388.

Chumpitaz Caceres, R. and N. G. Paparoidamis (2007). Service quality, relationship satisfaction, trust, commitment and business-to-business loyalty. *European Journal of Marketing 41*(7/8), 836–867.

Churchill Jr, G. A. and C. Surprenant (1982). An investigation into the determinants of customer satisfaction. *Journal of Marketing Research 19*(4), 491–504.

Cieslak, D. A., T. R. Hoens, N. V. Chawla, and W. P. Kegelmeyer (2012). Hellinger distance decision trees are robust and skew-insensitive. *Data Mining and Knowledge Discovery 24*(1), 136–158.

Cilli, M., G. S. Parnell, R. Cloutier, and T. Zigh (2015). A Systems Engineering Perspective on the Revised Defense Acquisition System. *Systems Engineering 18*(6), 584–603.

Cohen, M. C. (2018). Big data and service operations. *Production and Operations Management 27*(9), 1709–1723.

Cronin, J. J., M. K. Brady, and G. T. M. Hult (2000). Assessing the Effects of Quality, Value, and Customer Satisfaction on Consumer Behavioral Intentions in Service Environments. *Journal of Retailing 76*(2), 193–218.

Davenport, T. H. and R. Ronanki (2018). Artificial intelligence for the real world. *Harvard Business Review 96*(1), 108–116.

Davies, R. (2015). Industry 4.0. digitalisation for productivity and growth. Technical Report September, European Parliamentary Research Service.

de la Boutetière, H., A. Montagner, and A. Reich (2018). Unlocking success in digital transformations. Technical Report October, McKinsey & Company.

De Ruyter, K. and M. G. Wetzels (2000). The impact of perceived listening behavior in voice-to-voice service encounters. *Journal of Service Research 2*(3), 276–284.

Delana, K., N. Savva, and T. Tezcan (2020). Proactive Customer Service: Operational Benefits and Economic Frictions. *Manufacturing & Service Operations Management* (May), 1–18.

Delen, D., M. Erraguntla, and R. J. M. C.-n. Wu (2011). Better management of blood supply-chain with GIS-based analytics. *Annals of Operations Research 185*(1), 181–193.

Deloitte (2017). Global contact center survey. Technical report, Deloitte.

Deslauriers, A., P. L'Ecuyer, J. Pichitlamken, A. Ingolfsson, and A. N. Avramidis (2007). Markov chain models of a telephone call center with call blending. *Computers & Operations Research 34*(6), 1616–1645.

Despeisse, M., P. D. Ball, S. Evans, and A. Levers (2012). Industrial ecology at factory level - A conceptual model. *Journal of Cleaner Production 31*, 30–39.

DoD (2018). Digital Engineering Strategy. Technical Report June, U.S. Department of Defense, Office of the Deputy Assistant Secretary of Defense for Systems

Engineering, Washington, D.C.

Domingos, P. (1999). Metacost: A general method for making classifiers cost-sensitive. In *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 155–164. ACM.

Domingue, B. W., D. W. Belsky, J. M. Fletcher, D. Conley, J. D. Boardman, and K. M. Harris (2018). The social genome of friends and schoolmates in the national longitudinal study of adolescent to adult health. *Proceedings of the National Academy of Sciences of the United States of America 115*(4), 702–707.

Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and Perspectives 38*(1), 105.

Dyche, J. (2001). *The CRM Handbook: A Business Guide to Customer Relationship Management* (1 ed.). Upper Saddle River, NJ: Addison-Wesley.

Elsayed, E. A. (2012). *Reliability Engineering* (2nd ed.). New Brunswick: Wiley Publishing.

Erevelles, S., N. Fukawa, and L. Swayne (2016). Big data consumer analytics and the transformation of marketing. *Journal of Business Research 69*(2), 897–904.

Ertekin, S., J. Huang, L. Bottou, and C. Lee Giles (2007). Learning on the border: Active learning in imbalanced data classification. In *International Conference on Information and Knowledge Management, Proceedings*, pp. 127–136.

Feinberg, R. A., I.-s. Kim, L. Hokama, K. de Ruyter, and C. Keen (2000). Operational determinants of caller satisfaction in the call center. *International Journal of Service Industry Management 11*(2), 131–141.

Fernández, A., S. del Río, N. V. Chawla, and F. Herrera (2017). An insight into imbalanced big data classification: outcomes and challenges. *Complex & Intelligent Systems 3*(2), 105–120.

Fornell, C. (1992). A National Customer Satisfaction Barometer: The Swedish Experience. *Journal of Marketing 56*(1), 6.

Fornell, C., S. Mithas, F. V. Morgeson, and M. Krishnan (2006). Customer Satisfaction and Stock Prices: High Returns, Low Risk. *Journal of Marketing 70*(1), 3–14.

Forsberg, K. and H. Mooz (1991). The relationship of system engineering to the project cycle. In *INCOSE International Symposium*, Volume 1, pp. 57–65. Wiley Online Library.

Freund, Y. and R. E. Schapire (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences 55*(1), 119–139.

Friginal, E. (2009). Threats to the sustainability of the outsourced call center industry in the philippines: implications for language policy. *Language Policy 8*(1), 51.

Fukunaga, A., E. Hamilton, J. Fama, D. Andre, O. Matan, and I. Nourbakhsh (2002). Staff Scheduling for Inbound Call Centers and Customer Contact Centers. *AI Magazine 23*(4), 822–829.

Gabor, D. and S. Brooks (2017). The digital revolution in financial inclusion: international development in the fintech era. *New Political Economy 22*(4), 423–436.

Galar, M., A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera (2012). A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 42*(4), 463–484.

Galar, M., A. Fernández, E. Barrenechea, and F. Herrera (2013). EUSBoost: Enhancing ensembles for highly imbalanced data-sets by evolutionary undersampling. *Pattern Recognition 46*(12), 3460–3471.

Gallino, S. and A. Moreno (2014). Integration of online and offline channels in retail: The impact of sharing reliable inventory availability information. *Management Science 60*(6), 1434–1451.

Gama Dessavre, D., J. E. Ramirez-Marquez, and K. Barker (2016). Multidimensional approach to complex system resilience analysis. *Reliability Engineering and System Safety 149*, 34–43.

Gans, N., G. Koole, and A. Mandelbaum (2003). Telephone call centers: Tutorial, review, and research prospects. *Manufacturing & Service Operations Management Publication 5*(2), 79–141.

Gans, N., N. Liu, A. Mandelbaum, H. Shen, and H. Ye (2010). Service times in call centers: Agent heterogeneity and learning with some operational consequences. In *Borrowing Strength: Theory Powering Applications*, Volume 6, pp. 99–123. Institute of Mathematical Statistics.

Garnett, O., A. Mandelbaum, and M. Reiman (2002). Designing a call center with impatient customers. *Manufacturing & Service Operations Management 4*(3), 208–227.

Geissbauer, R., J. Vedso, and S. Schrauf (2016). Industry 4.0: Building the digital enterprise. Technical report, PwC.

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine 27*(15), 2865–2873.

Giachetti, R. (2010). *Design of Enterprise Systems*. Boca Raton: CRC Press.

Gölzer, P. and A. Fritzsche (2017). Data-driven operations management: organisational implications of the digital transformation in industrial practice. *Production Planning and Control 28*(16), 1332–1343.

Groll, A. and G. Tutz (2014). Variable selection for generalized linear mixed models by l1-penalized estimation. *Statistics and Computing 24*(2), 137–154.

Gronholdt, L., A. Martensen, and K. Kristensen (2010). The relationship between customer satisfaction and loyalty: Cross-industry differences. *Total Quality Management 11*(April 2012), 37–41.

Gronroos, C. (1990). *Service Management and Marketing: Managing the Moments of Truth in Service Competition*. Lexington books.

Guyon, I. and A. Elisseeff (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research 3*(Mar), 1157–1182.

Haibo He, Yang Bai, E. A. Garcia, and Shutao Li (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, Volume 3, pp. 1322–1328. IEEE.

Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten (2009). The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter 11*(1), 10–18.

Hanifa, N., M. Zaidin, M. Nurazri, M. Diah, H. Yee, and S. Sorooshian (2014). Quality management in industry 4.0 era. *Journal of Management and Science 4*(3), 2250–1819.

Hartmann, P. M., M. Zaki, N. Feldmann, and A. Neely (2016). Capturing value from big data – a taxonomy of data-driven business models used by start-up firms. *International Journal of Operations and Production Management 36*(10), 1382–1406.

Hastie, T., R. Tibshirani, and J. Friedman (2001). *The Elements of Statistical Learning*. Springer, New York.

He, H. and E. A. Garcia (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering 21*(9), 1263–1284.

Heizer, J. and B. Render (2014). *Operations Management: Sustainability and Supply Chain Management* (11 ed.). Upper Saddle River, New Jersey: Pearson Education, Inc.

Holland, P. (2003). Achieving a step change in contact centre performance: Exploiting data to make better. *Journal of Targeting, Measurement and Analysis for Marketing 12*(2), 103–113.

Holtewert, P., R. Wutzke, J. Seidelmann, and T. Bauernhansl (2013). Virtual Fort Knox Federative, Secure and Cloud-based Platform for Manufacturing. *Procedia CIRP 7*, 527–532.

Homrighausen, D. and D. McDonald (2013). The lasso, persistence, and cross-validation. In *International Conference on Machine Learning*, pp. 1031–1039.

Hu, S., Y. Liang, L. Ma, and Y. He (2009). MSMOTE: Improving classification performance when training data is imbalanced. In *2nd International Workshop on Computer Science and Engineering, WCSE 2009*, Volume 2, pp. 13–17. IEEE.

Huang, T. and J. A. Van Mieghem (2014). Clickstream data and inventory management: Model and empirical analysis. *Production and Operations Management 23*(3), 333–347.

Huang, Y.-M., C.-M. Hung, and H. C. Jiau (2006). Evaluation of neural networks and data mining methods on a credit assessment task for class imbalance problem. *Nonlinear Analysis: Real World Applications 7*(4), 720–747.

Hubbard, E. M., C. E. Siemieniuch, M. A. Sinclair, and A. Hodgson (2010). Working towards a holistic organisational systems model. In *2010 5th International Conference on System of Systems Engineering, SoSE 2010*, pp. 1–6. IEEE.

Hunt, H. K. (1977). Conceptualization and measurement of consumer satisfaction and dissatisfaction. *Marketing Science Institute (Report No. 77-103)*, 490.

Ibrahim, R. and P. L'Ecuyer (2012). Forecasting call center arrivals: Fixed-effects, mixed-effects, and bivariate models. *Manufacturing & Service Operations Management 15*(1), 72–85.

Ibrahim, R. and P. L'Ecuyer (2013). Forecasting call center arrivals: Fixed-effects, mixed-effects, and bivariate models. *Manufacturing & Service Operations Management 15*(1), 72–85.

Ibrahim, R., P. L'Ecuyer, H. Shen, and M. Thiongane (2016). Inter-dependent, heterogeneous, and time-varying service-time distributions in call centers. *European Journal of Operational Research 250*(2), 480–492.

Iranmehr, A., H. Masnadi-Shirazi, and N. Vasconcelos (2019). Cost-sensitive support vector machines. *Neurocomputing 343*, 50–64.

Ishikawa, D. K. (1985). *What Is Total Quality Control? The Japanese Way (Business Management).* Englewood Cliffs, N.J.: Prentice Hall Trade.

Jaaron, A. A. and C. J. Backhouse (2011). Systems thinking for call centre service design: Affective commitment implications in manufacturing enterprises. *Service Industries Journal 31*(4), 613–628.

Japkowicz, N. (2000). The class imbalance problem: Significance and strategies. In *Proceedings of the International Conference on Artificial Intelligence*.

Jerath, K., A. Kumar, and S. Netessine (2015). An information stock model of customer behavior in multichannel customer support services. *Manufacturing & Service Operations Management 17*(3), 368–383.

Jiang, J., A. Hassan Awadallah, R. Jones, U. Ozertem, I. Zitouni, R. Gurunath Kulkarni, and O. Z. Khan (2015). Automatic online evaluation of intelligent assistants. In *Proceedings of the 24th International Conference on World Wide Web*, pp. 506–516. International World Wide Web Conferences Steering Committee.

Jin, J., Y. Liu, P. Ji, and H. Liu (2016). Understanding big consumer opinion data for market-driven product design. *International Journal of Production Research 54*(10), 3019–3041.

Johnston, R. (1995). The determinants of service quality: satisfiers and dissatisfiers. *International Journal of Service Industry Management 6*(5), 53–71.

Kailath, T. (1967). The divergence and bhattacharyya distance measures in signal selection. *IEEE Transactions on Communication Technology 15*(1), 52–60.

Karl, A. T., Y. Yang, and S. L. Lohr (2014). Computation of maximum likelihood estimates for multiresponse generalized linear mixed models with non-nested, correlated random effects. *Computational Statistics & Data Analysis 73*, 146–162.

Karl, T. R., N. Nicholls, and A. Ghazi (2013). *Weather and climate extremes: changes, variations and a perspective from the insurance industry.* Springer Science & Business Media.

Katchasuwanmanee, K., R. Bateman, and K. Cheng (2016). Development of the Energy-smart Production Management system (e-ProMan): A Big Data driven approach, analysis and optimisation. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 230*(5), 972–978.

Keblis, M. F. and M. Chen (2006). Improving customer service operations at Amazon.com. *Interfaces 36*(5), 433–445.

Kelleher, J. D., B. Mac Namee, and A. D'arcy (2015). *Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies.* MIT Press.

Khoshgoftaar, T. M., J. Van Hulse, and A. Napolitano (2011). Comparing boosting and bagging techniques with noisy and imbalanced data. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans 41*(3), 552–568.

Khreich, W., E. Granger, A. Miri, and R. Sabourin (2010). Iterative boolean combination of classifiers in the roc space: An application to anomaly detection with hmms. *Pattern Recognition 43*(8), 2732–2752.

Khudyakov, P., P. D. Feigin, and A. Mandelbaum (2010). Designing a call center with an IVR (interactive voice response). *Queueing Systems 66*(3), 215–237.

Kiel, D., J. M. Muller, C. Arnold, and K.-i. Voigt (2017). Sustainable Industrial Value Creation: Benefits and Challenges of Industry 4.0. *International Journal of Innovation Management 21*(08), 34.

Kim, B. D., M. Shi, and K. Srinivasan (2004). Managing capacity through reward programs. *Management Science 50*(4), 503–520.

Kim, K. and S. Mehrotra (2015). A two-stage stochastic integer programming approach to integrated staffing and scheduling with application to nurse management. *Operations Research 63*(6), 1431–1451.

Kiseleva, J., K. Williams, A. Hassan Awadallah, A. C. Crook, I. Zitouni, and T. Anastasakos (2016). Predicting user satisfaction with intelligent assistants. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 45–54. ACM.

Kleiner, A., A. Talwalkar, P. Sarkar, and M. I. Jordan (2014). A scalable bootstrap for massive data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 76*(4), 795–816.

ko de Ruyter and M. G. M. Wetzels (2000). The Impact of Perceived Listening Behavior in Voice-to-Voice Service Encounters. *Journal of Service Research 2*(3), 276–284.

Koole, G. (2013). *Call Center Optimization.* Amsterdam: MG books.

Koole, G. and A. Mandelbaum (2002). Queueing models of call centers: An introduction. *Annals of Operations Research 113*(1-4), 41–59.

Kumar, A. and R. Telang (2012). Does the web reduce customer service cost? empirical evidence from a call center. *Information Systems Research 23*(3-part-1), 721–737.

Kumar, V. (2010). Customer Relationship Management. In J. Sheth and N. Malhotra (Eds.), *Wiley International Encyclopedia of Marketing.* Chichester, UK: John Wiley & Sons, Ltd.

Kurti, E. and D. Haftor (2015). Barriers and enablers of digital business model transformation. In *The European Conference on Information Systems Management*, pp. 262. Academic Conferences International Limited.

Laird, N. M. and J. H. Ware (1982). Random-effects models for longitudinal data. *Biometrics 38*(4), 963–974.

Lasi, H., P. Fettke, T. Feld, and M. Hoffmann (2014). Industry 4.0. *Business & Information Systems Engineering 6*, 239–242.

L'Ecuyer, P. (2006). Modeling and optimization problems in contact centers. In *Third International Conference on the Quantitative Evaluation of Systems, QEST 2006*, pp. 145–154.

Lee, J., H. A. Kao, and S. Yang (2014). Service innovation and smart analytics for Industry 4.0 and big data environment. *Procedia CIRP 16*, 3–8.

Legros, B., O. Jouini, and G. Koole (2015). Adaptive threshold policies for multi-channel call centers. *IIE Transactions 47*(4), 414–430.

Lehrer, C., A. Wieneke, J. vom Brocke, R. Jung, and S. Seidel (2018). How Big Data Analytics Enables Service Innovation: Materiality, Affordance, and the Individualization of Service. *Journal of Management Information Systems 35*(2), 424–460.

Leong, L. Y., T. S. Hew, V. H. Lee, and K. B. Ooi (2015). An SEM-artificial-neural-network analysis of the relationships between SERVPERF, customer satisfaction and loyalty among low-cost and full-service airline. *Expert Systems with Applications 42*(19), 6620–6634.

Leskovec, J., A. Rajaraman, and J. D. Ullman (2014). *Mining of Massive Datasets*. Cambridge University Press.

Liao, S., C. van Delft, and J.-P. Vial (2013). Distributionally robust workforce scheduling in call centres with uncertain arrival rates. *Optimization Methods and Software 28*(3), 501–522.

Lieu, D. K. and S. Sorby (2009). *Visualization, Modeling, and Graphics for Engineering Design*. Clifton Park, NY: Delmar, Cengage Learning.

Lin, Y., Y. Lee, and G. Wahba (2002). Support vector machines for classification in nonstandard situations. *Machine learning 46*(1-3), 191–202.

Liu, X., P. V. Singh, and K. Srinivasan (2016). A structured analysis of unstructured big data by leveraging cloud computing. *Marketing Science 35*(3), 363–388.

López, V., A. Fernández, S. García, V. Palade, and F. Herrera (2013). An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. *Information Sciences 250*, 113–141.

Luo, X. and C. B. Bhattacharya (2006). Corporate Social Responsibility, Customer Satisfaction, and Market Value. *Journal of Marketing 70*(October), 1–18.

Lusa, L. et al. (2013). Smote for high-dimensional class-imbalanced data. *BMC Bioinformatics 14*(1), 106.

Ma, J., M. Kwak, and H. M. Kim (2014). Demand trend mining for predictive life cycle design. *Journal of Cleaner Production 68*, 189–199.

Mandelbaum, A. (2004). Call centers (centres): Research bibliography with abstracts. Technical report, Israel Institute of Technology, Haifa.

Mandelbaum, A. and S. Zeltyn (2009). Staffing many-server queues with impatient customers: Constraint satisfaction in call centers. *Operations Research 57*(5), 1189–1205.

Marr, B. and S. Parry (2004). Performance management in call centers: Lessons, pitfalls and achievements in Fujitsu Services. *Measuring Business Excellence 8*(4), 55–62.

Masnadi-Shirazi, H. and N. Vasconcelos (2010). Risk minimization, probability elicitation, and cost-sensitive SVMs. In *ICML 2010 - Proceedings, 27th International Conference on Machine Learning*, pp. 759–766.

Matt, C., T. Hess, and A. Benlian (2015). Digital Transformation Strategies. *Business and Information Systems Engineering 57*(5), 339–343.

Mattia, S., F. Rossi, M. Servilio, and S. Smriglio (2017). Staffing and scheduling flexible call centers by two-stage robust optimization. *Omega (United Kingdom) 72*, 25–37.

McAfee, A. and E. Brynjolfsson (2012). Big data: The management review. *Harvard Business Review 90*(10), 60–68.

Mcgee, S. and R. Edson (2010). Extending the Conceptagon as an Analytic Framework : A Case Study of Public Preparedness in Israel. In *8th Conference on Systems Engineering Research*, Hoboken, NJ, USA, pp. 17–19. CSER.

McLachlan, G. J. and T. Krishnan (2008). *Extensions of the EM Algorithm, in The EM Algorithm and Extensions* (2nd ed.). John Wiley & Sons, Inc., Hoboken, NJ, USA.

McNair Andrew & Holmes Richard (2015). 2015 Global Contact Centre Benchmarking Report. Technical report, Dimension Data.

Mehrbod, N., A. Grilo, and A. Zutshi (2018). Caller-Agent Pairing in Call Centers Using Machine Learning Techniques with Imbalanced Data. In *2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, pp. 1–6.

Mehrotra, V., K. Ross, G. Ryder, and Y. P. Zhou (2012). Routing to manage resolution and waiting time in call centers with heterogeneous servers. *Manufacturing and Service Operations Management 14*(1), 66–81.

Meinshausen, N. and P. Bühlmann (2010). Stability selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 72*(4), 417–473.

Meurer, A., C. P. Smith, M. Paprocki, O. Čertík, S. B. Kirpichev, M. Rocklin, A. Kumar, S. Ivanov, J. K. Moore, S. Singh, et al. (2017). SymPy: symbolic computing in python. *PeerJ Computer Science 3*, e103.

Moazeni, S. (2018). Predicting individual-level call arrival from online account customer activity. In *2018 IEEE International Conference on Big Data (Big Data)*, pp. 1624–1631. IEEE.

Moazeni, S. and R. Andrade (2018, Jul). A data-driven approach to predict an individual customer's call arrival in multichannel customer support centers. In *2018 IEEE International Congress on Big Data (BigData Congress)*, pp. 66–73.

IEEE.

Moazeni, S., B. Defourny, and M. J. Wilczak (2019). Sequential learning in designing marketing campaigns for market entry. *Management Science*, 1–20.

Moldavska, A. and T. Welo (2015). Development of Manufacturing Sustainability Assessment Using Systems Thinking. *Sustainability 8*(1), 1–26.

Montgomery, D. C. (2009). *Statistical Quality Control*, Volume 7. New York: Wiley.

Moon, S. and G. J. Russell (2008). Predicting product purchase from inferred customer similarity: An autologistic model approach. *Management Science 54*(1), 71–82.

Moro, S., P. Cortez, and P. Rita (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems 62*, 22–31.

Mueller, F., H. Naujoks, H. Singh, G. Schwarz, A. Schwedel, and K. Thomson (2015). Global digital insurance benchmarking report 2015 - pathways to success in a digital world. Technical report, Bain & Company.

Mukherjee, U. K. and K. K. Sinha (2018). Product Recall Decisions in Medical Device Supply Chains: A Big Data Analytic Approach to Evaluating Judgment Bias. *Production and Operations Management 27*(10), 1816–1833.

Mularoni, M. and D. Barnett (2017). Technology, MA and geopolitical uncertainty poised to disrupt contact center industry. JLL Research, Everest Group, Bureau of Labor Statistics.

Nachlas, J. A. (2017). *Reliability Engineering: Probabilistic Models and Maintenance Methods*. Boca Raton: CRC Press.

Nagar, V. and M. V. Rajan (2005). Measuring customer relationships: The case of the retail banking industry. *Management Science 51*(6), 904–919.

Ndwe, T. J. and N. Dlodlo (2015). Relationship between affordance and cultural conventions in the design of IVR systems for oral users. In *New Trends in Networking, Computing, E-learning, Systems Sciences, and Engineering*, pp. 445–450. Springer.

Neslin, S. A., D. Grewal, R. Leghorn, V. Shankar, M. L. Teerling, J. S. Thomas, and P. C. Verhoef (2006). Challenges and opportunities in multichannel customer management. *Journal of Service Research 9*(2), 95–112.

Nguyen, T., L. ZHOU, V. Spiegler, P. Ieromonachou, and Y. Lin (2018). Big data analytics in supply chain management: A state-of-the-art literature review. *Computers and Operations Research 98*, 254–264.

NOAA-NCEI (2016). National Oceanic and Atmospheric Administration's (NOAA) National Centers for Environmental Information (NCEI). Available at the NOAA - NCEI website `http://www.ncdc.noaa.gov`.

Nunnally, J. C., I. H. Bernstein, et al. (1967). *Psychometric Theory*, Volume 226. McGraw-Hill New York.

Oh, H. (1999). Service quality, customer satisfaction, and customer value: A holistic perspective. *International Journal of Hospitality Management 18*(1), 67–82.

Oks, S. J., A. Fritzsche, and C. Lehmann (2016). The digitalisation of industry from a strategic perspective. In *R&D Management Conference 2016 "From Science to*

Society: Innovation and Value Creation", pp. 1–9.

Oliver, R. L. (1981). Measurement and Evaluation of Satisfaction Process in Retail Settings. *Journal of Retailing 57*(3), 25–48.

Opresnik, D. and M. Taisch (2015). The value of big data in servitization. *International Journal of Production Economics 165*, 174–184.

Parasuraman, A., V. A. Zeithaml, and L. L. Berry (1988). SERVQUAL: A Multiple-Item Scale For Measuring Consumer Perceptions of Service Quality. *Journal of Retailing 64*(1), 12.

Parasuraman, Arun and Berry, Leonard L and Zeithaml, V. A. (1991). Understanding customer expectations of service. *Sloan Management Review 32*(3), 39–48.

Parvatiyar, A. and J. N. Sheth (2001). Customer relationship management: Emerging practice, process, and discipline. *Journal of Economic & Social Research 3*(2), 1–34.

Patrick Eigbe, A., B. J. Sauser, and J. Boardman (2009). Soft systems analysis of the unification of test and evaluation and program management: A study of a federal aviation administration's strategy. *Systems Engineering 14*, 298–309.

Payne, A. and P. Frow (2004). The role of multichannel integration in customer relationship management. *Industrial Marketing Management 33*(6), 527–538.

Philip, K. and S. Chan (1998). Toward scalable learning with non-uniform class and cost distributions: A case study in credit card fraud detection. In *Proceeding of the Fourth International Conference on Knowledge Discovery and Data Mining*, pp. 164–168.

Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine 6*(3), 21–45.

Posada, J., C. Toro, I. Barandiaran, D. Oyarzun, D. Stricker, R. De Amicis, E. B. Pinto, P. Eisert, J. Döllner, and I. Vallarino (2015). Visual computing as a key enabling technology for industrie 4.0 and industrial internet. *IEEE Computer Graphics and Applications 35*(2), 26–40.

Pot, A., S. Bhulai, and G. Koole (2008). A simple staffing method for multiskill call centers. *Manufacturing & Service Operations Management 10*(3), 421–428.

Prati, R. C., G. E. Batista, and M. C. Monard (2009). Data mining with imbalanced class distributions: concepts and methods. In *4th Indian International Conference on Artificial Intelligence (IICAI-09)*, pp. 359–376.

Provost, F. and P. Domingos (2003). Tree induction for probability-based ranking. *Machine Learning 52*(3), 199–215.

Provost, F. and T. Fawcett (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data 1*(1), 51–59.

Ramentol, E., S. Vluymans, N. Verbiest, Y. Caballero, R. Bello, C. Cornelis, and F. Herrera (2015). Ifrowann: imbalanced fuzzy-rough ordered weighted average nearest neighbor classification. *IEEE Transactions on Fuzzy Systems 23*(5), 1622–1637.

Ransbotham, S., D. Kiron, P. Gerbert, and M. Reeves (2017). Reshaping business with artificial intelligence. *MIT Sloan Management Review and The Boston Consulting Group 59*(1), 1.

Ravichandran, K. (2010). Influence of Service Quality on Customer Satisfaction Application of SERVQUAL Model. *Journal of Business and Management 5*(4), 117–125.

Rayhan, F., S. Ahmed, A. Mahbub, M. R. Jani, S. Shatabda, D. M. Farid, and C. M. Rahman (2017). MEBoost: Mixing estimators with boosting for imbalanced data classification. In *2017 11th International Conference on Software, Knowledge Information, Industrial Management and Applications (SKIMA)*, pp. 1–6.

Reichheld, F. F. (2003). The one number you need to grow. *Harvard Business Review 81*(12), 46–55.

Rhodes, D. H., A. M. Ross, and D. J. Nightingale (2009). Architecting the system of systems Enterprise: Enabling constructs and methods from the field of engineering systems. In *2009 IEEE International Systems Conference Proceedings*, pp. 190–195. IEEE.

Rijo, R., J. Varajão, and R. Gonçalves (2012). Contact center: information systems design. *Journal of Intelligent Manufacturing 23*(3), 497–515.

Ringel, M., H. Zablit, F. Grassl, J. Manly, and C. Möller (2018). The Most Innovative Companies in 2018. Technical report, Boston Consulting Group.

Robbins, T. R. (2016). Evaluating the fit of the erlang a model in high traffic call centers. In *Winter Simulation Conference (WSC), 2016*, pp. 1790–1801. IEEE.

Roblek, V., M. Meško, and A. Krapež (2016). A Complex View of Industry 4.0. *SAGE Open 6*(2), 1–11.

Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review 33*(1-2), 1–39.

Rosen, S. L., C. P. Saunders, and S. K. Guharay (2015). A structured approach for rapidly mapping multilevel system measures via simulation metamodeling. *Systems Engineering 18*(1), 87–101.

Rouse, W. B. (2005a). Enterprises as systems: Essential challenges and approaches to transformation. *Systems Engineering 8*(2), 138–150.

Rouse, W. B. (2005b). A theory of enterprise transformation. *Systems Engineering 8*(4), 279–295.

Rouse, W. B. (2014). Human interaction with policy flight simulators. *Applied Ergonomics 45*(1), 72–77.

Rustamov, S., E. Mustafayev, and M. A. Clements (2018). Context analysis of customer requests using a hybrid adaptive neuro fuzzy inference system and hidden Markov models in the natural language call routing problem. *Open Engineering 8*(1), 61–68.

Ryzhov, I. O., B. Han, and J. Bradić (2015). Cultivating disaster donors using data analytics. *Management Science 62*(3), 849–866.

Saberi, M., O. Khadeer Hussain, and E. Chang (2017). Past, present and future of contact centers: a literature review. *Business Process Management Journal 23*(3), 574–597.

Sagaert, Y. R., E. H. Aghezzaf, N. Kourentzes, and B. Desmet (2018). Temporal Big Data for Tactical Sales Forecasting in the Tire Industry. *INFORMS Journal on Applied Analytics 48*(2), 121–129.

Sage, A. P. and C. L. Lynch (1998). Systems integration and architecting: An overview of principles, practices, and perspectives. *Systems Engineering 1*(3), 176–227.

Sahay, B. S. (2005). Multi-factor productivity measurement model for service organisation. *International Journal of Productivity and Performance Management 54*(1), 7–22.

Santouridis, I. and P. Trivellas (2010). Investigating the impact of service quality and customer satisfaction on customer loyalty in mobile telephony in Greece. *The TQM Journal 22*(3), 330–343.

Sarikaya, R., G. E. Hinton, and A. Deoras (2014). Application of deep belief networks for natural language understanding. *IEEE Transactions on Audio, Speech and Language Processing 22*(4), 778–784.

Sauser, B., M. Mansouri, and M. Omer (2011). Using Systemigrams in Problem Definition: A Case Study in Maritime Resilience for Homeland Security. *Journal of Homeland Security and Emergency Management 8*(1), 1–10.

Sawhney, R., K. Subburaman, C. Sonntag, P. R. Venkateswara Rao, and C. Capizzi (2010). A modified FMEA approach to enhance reliability of lean systems. *International Journal of Quality and Reliability Management 27*(7), 832–855.

Schapire, R. E. (1990). The Strength of Weak Learnability (Extended Abstract). *Machine Learning 5*, 197–227.

Schelldorfer, J., P. Bühlmann, G. De, and S. Van (2011). Estimation for High-Dimensional Linear Mixed-Effects Models Using L1-Penalization. *Scandinavian Journal of Statistics 38*(2), 197–214.

Scherer, A., N. V. Wünderlich, and F. Von Wangenheim (2015). The value of self-service: Long-term effects of technology-based self-service usage on customer retention. *MIS Quarterly 39*(1), 177–200.

Schoeller, A. and R. Heffner (2014). Design your contact center from the outside in. Technical report, Forrester Research, Inc., Cambridge, MA.

Schoenherr, T. and C. Speier-pero (2015). Data science, predictive analytics, and big data in supply chain management: current state and future potential. *Journal of Business Logistics 36*(1), 120–132.

Schwarz, G. et al. (1978). Estimating the dimension of a model. *The Annals of Statistics 6*(2), 461–464.

Seada, A. A. and A. B. Eltawil (2015, March). Modeling and analysis of workforce management decisions in modern call centers. In *2015 International Conference on Industrial Engineering and Operations Management (IEOM)*.

Seiffert, C., T. M. Khoshgoftaar, J. Van Hulse, and A. Napolitano (2010). RUSBoost: A hybrid approach to alleviating class imbalance. *IEEE Transactions on Systems,* *Man, and Cybernetics Part A: Systems and Humans 40*(1), 185–197.

Seltman, H. (2018). *Experimental Design and Analysis*. Online Book.

Sesmero, M. P., A. I. Ledezma, and A. Sanchis (2015). Generating ensembles of heterogeneous classifiers using stacked generalization. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 5*(1), 21–34.

Shankar, V., A. K. Smith, and A. Rangaswamy (2003). Customer satisfaction and loyalty in online and offline environments. *International Journal of Research in Marketing 20*(2), 153–175.

Sharp, D. E. (2003). Organizing and managing the call center. In D. Sharp (Ed.), *Call Center Operation*, Chapter 3, pp. 61–135. Elsevier.

Shen, H. (2010). Statistical analysis of call-center operational data: Forecasting call arrivals, and analyzing customer patience and agent service. *Wiley Encyclopedia of Operations Research and Management Science*, 1–10.

Shen, H. and J. Z. Huang (2005). Analysis of call centre arrival data using singular value decomposition. *Applied Stochastic Models in Business and Industry 21*(3), 251–263.

Shen, H. and J. Z. Huang (2008a). Forecasting time series of inhomogeneous poisson processes with application to call center workforce management. *The Annals of Applied Statistics 2*(2), 601–623.

Shen, H. and J. Z. Huang (2008b). Interday forecasting and intraday updating of call center arrivals. *Manufacturing & Service Operations Management 10*(3), 391–410.

Sikora, R. et al. (2015). A modified stacking ensemble machine learning algorithm using genetic algorithms. In *Handbook of Research on Organizational Transformations through Big Data Analytics*, pp. 43–53. IGi Global.

Sitton, M. and Y. Reich (2015). Enterprise Systems Engineering for Better Operational Interoperability. *Systems Engineering 18*(6), 625–638.

Sitton, M. and Y. Reich (2018). EPIC framework for enterprise processes integrative collaboration. *Systems Engineering 21*(1), 30–46.

Sołtys, M., S. Jaroszewicz, and P. Rzepakowski (2015). Ensemble methods for uplift modeling. *Data Mining and Knowledge Discovery 29*(6), 1531–1559.

Soujanya, M. and S. Kumar (2010). Personalized ivr system in contact center. In *Electronics and Information Engineering (ICEIE), 2010 International Conference On*, Volume 1, pp. V1–453. IEEE.

Soyer, R. and M. M. Tarimcilar (2008). Modeling and analysis of call center arrival data: A bayesian approach. *Management Science 54*(2), 266–278.

Stefanovic, N. (2015). Collaborative Predictive Business Intelligence Model for Spare Parts Inventory Replenishment. *Computer Science and Information Systems 12*(3), 911–930.

Stolletz, R. (2003). Characterization of Inbound Call Centers. *Performance Analysis and Optimization of Inbound Call Centers. Lecture Notes in Economics and Mathematical Systems 528*, 3–20.

Suhm, B. and P. Peterson (2002). A data-driven methodology for evaluating and optimizing call center IVRs. *International Journal of Speech Technology 5*(1), 23–37.

Sun, Y., A. K. Wong, and M. S. Kamel (2009). Classification of imbalanced data: A review. *International Journal of Pattern Recognition and Artificial Intelligence 23*(4), 687–719.

Tahir, M. A., J. Kittler, and F. Yan (2012). Inverse random under sampling for class imbalance problem and its application to multi-label classification. *Pattern Recognition 45*(10), 3738–3750.

Tang, M., B. Pellom, and K. Hacioglu (2003). Call-type classification and unsupervised training for the call center domain. In *2003 IEEE Workshop on Automatic Speech Recognition and Understanding*, pp. 204–208. IEEE.

Tao, F., J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui (2018). Digital twin-driven product design, manufacturing and service with big data. *International Journal of Advanced Manufacturing Technology 94*(9-12), 3563–3576.

Tao, F., L. Zhang, and A. Y. Nee (2011). A review of the application of grid technology in manufacturing. *International Journal of Production Research 49*(13), 4119–4155.

Tate, W. L. and W. van der Valk (2008). Managing the performance of outsourced customer contact centers. *Journal of Purchasing and Supply Management 14*(3), 160–169.

Taylor, C. (2017). Combine the power of ivr and wfo analytics to evolve agent and customer experience. Technical Report April, Forrester Research, Inc., Cambridge, MA.

Taylor, J. W. (2008). A comparison of univariate time series methods for forecasting intraday arrivals at a call center. *Management Science 54*(2), 253–265.

Taylor, J. W. (2012). Density forecasting of intraday call center arrivals using models based on exponential smoothing. *Management Science 58*(3), 534–549.

Teas, R. K. (1993). Consumer Expectations and the Measurement of Perceived Service Quality. *Journal of Professional Services Marketing 8*(2), 33–54.

Tezcan, T. and B. Behzad (2012). Robust design and control of call centers with flexible interactive voice response systems. *Manufacturing & Service Operations Management 14*(3), 355–484.

The Open Group Architecture Forum (2018). *The Open Group Standard - The TOGAF Standard, Version 9.2.* Reading, UK: The Open Group.

Thirumalai, S. and K. K. Sinha (2013). To personalize or not to personalize online purchase interactions: Implications of self-selection by retailers. *Information Systems Research 24*(3), 683–708.

Thomas, J. S. and U. Y. Sullivan (2005). Managing Marketing Communications with Multichannel Customers. *Journal of Marketing 69*(4), 239–251.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological) 58*(1), 267–288.

Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. *Statistics in Medicine 16*(4), 385–395.

Tiwari, S., H. Wee, and Y. Daryanto (2018). Big data analytics in supply chain management between 2010 and 2016: Insights to industries. *Computers & Industrial Engineering 115*(May 2017), 319–330.

Tsiptsis, K. K. and A. Chorianopoulos (2011). *Data Mining Techniques in CRM: Inside Customer Segmentation*. John Wiley & Sons.

Tyson, N. and V. Matula (2004). Improved lsi-based natural language call routing using speech recognition confidence scores. In *Second IEEE International Conference on Computational Cybernetics, 2004. ICCC 2004.*, pp. 409–413. IEEE.

U.S. Bureau of Labor Statistics (2019). Employment by major industry sector.

van der Aa, Z., J. Bloemer, and J. Henseler (2015). Using customer contact centres as relationship marketing instruments. *Service Business 9*(2), 185–208.

van Dun, Z., J. Bloemer, and J. Henseler (2011). Perceived customer contact centre quality: Conceptual foundation and scale development. *The Service Industries Journal 31*(8), 1347–1363.

Veiga, R. A. (2016). Alternative model for estimating the abandon rate in call centers. In *Biennial Congress of Argentina (ARGENCON), 2016 IEEE*, pp. 1–5. IEEE.

Velu, C. (2017). A Systems Perspective on Business Model Evolution: The Case of an Agricultural Information Service Provider in India. *Long Range Planning 50*(5), 603–620.

Venkatesh, V. and R. Agarwal (2006). Turning visitors into customers: A usability-centric perspective on purchase behavior in electronic channels. *Management Science 52*(3), 367–382.

Verhagen, T., J. Van Nes, F. Feldberg, and W. Van Dolen (2014). Virtual customer service agents: Using social presence and personalization to shape online service encounters. *Journal of Computer-Mediated Communication 19*(3), 529–545.

Verhoef, P. C., T. Broekhuizen, Y. Bart, A. Bhattacharya, J. Qi Dong, N. Fabian, and M. Haenlein (2019). Digital transformation: A multidisciplinary reflection and research agenda. *Journal of Business Research*.

Veropoulos, K., C. Campbell, N. Cristianini, et al. (1999). Controlling the sensitivity of support vector machines. In *Proceedings of the International Joint Conference on AI*, Volume 55, pp. 60.

von Leipzig, T., M. Gamp, D. Manz, K. Schöttle, P. Ohlhausen, G. Oosthuizen, D. Palm, and K. von Leipzig (2017). Initialising Customer-orientated Digital Transformation in Enterprises. *Procedia Manufacturing 8*(October 2016), 517–524.

Wade, J. (2016). Conceptagon. Lecture Notes, ES684 Systems Thinking, Stevens Institute of Technology.

Wahlster, W., J. Helbig, A. Hellinger, M. A. V. Stumpf, J. Blasco, H. Galloway, and H. Gestaltung (2013). Recommendations for implementing the strategic initiative INDUSTRIE 4.0. Technical Report April, Forschungsunion, Frankfurt.

Wallace, D. W., J. L. Giese, and J. L. Johnson (2004). Customer retailer loyalty in the context of multiple channel strategies. *Journal of Retailing 80*(4), 249–263.

Wang, B. and J. Pineau (2016). Online Bagging and Boosting for Imbalanced Data

Streams. *IEEE Transactions on Knowledge and Data Engineering 28*(12), 3353–3366.

Wang, B. X. and N. Japkowicz (2004). Imbalanced data set learning with synthetic samples. In *Proc. IRIS Machine Learning Workshop*, Volume 19.

Wang, G., A. Gunasekaran, and E. W. Ngai (2018). Distribution network design with big data: model and analysis. *Annals of Operations Research 270*(1-2), 539–551.

Wang, H., G. Li, and C.-L. Tsai (2007). Regression coefficient and autoregressive order shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 69*(1), 63–78.

Wang, J. and J. Zhang (2016). Big data analytics for cycle time related feature selection in the semiconductor wafer fabrication system. *International Journal of Production Research 54*(23), 7231–4244.

Wang, S. and X. Yao (2009). Diversity analysis on imbalanced data sets by using ensemble models. In *2009 IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2009 - Proceedings*. IEEE.

Warrington, P. T., E. Gangstad, R. Feinberg, and K. de Ruyter (2007). Multi-Channel Retailing and Customer Satisfaction. *International Journal of E-Business Research 3*(2), 57–69.

Weinberg, J., L. D. Brown, and J. R. Stroud (2007). Bayesian forecasting of an inhomogeneous poisson process with applications to call center data. *Journal of the American Statistical Association 102*(480), 1185–1198.

Wild, C. J. (2000). *Chance encounters: A first course in data analysis and inference.* New York: Citeseer.

Witten, I. H., E. Frank, M. A. Hall, and C. J. Pal (2016). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann.

Wolpert, D. H. (1992). Stacked generalization. *Neural Networks 5*(2), 241–259.

Wolpert, D. H. and W. G. Macready (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation 1*(1), 67–82.

Wu, G. and E. Y. Chang (2003). Class-boundary alignment for imbalanced dataset learning. In *ICML 2003 Workshop on Learning from Imbalanced Data Sets II, Washington, DC*, pp. 49–56.

Xue, M., L. M. Hitt, and P. T. Harker (2007). Customer efficiency, channel usage, and firm performance in retail banking. *Manufacturing & Service Operations Management 9*(4), 535–558.

Yaffe-Bellany, D. (2019, Oct). Would you like fries with that? McDonald's already knows the answer. *The New York Times*.

Ye, H., J. Luedtke, and H. Shen (2019). Call Center Arrivals: When to Jointly Forecast Multiple Streams? *Production and Operations Management 28*(1), 27–42.

Yoo, Y., R. J. Boland, K. Lyytinen, and A. Majchrzak (2012). Organizing for innovation in the digitized world. *Organization Science 23*(5), 1398–1408.

Yoon, H., K. Yang, and C. Shahabi (2005). Feature subset selection and feature ranking for multivariate time series. *IEEE Transactions on Knowledge and Data Engineering 17*(9), 1186–1198.

Yu, H., S. Hong, X. Yang, J. Ni, Y. Dan, and B. Qin (2013). Recognition of multiple imbalanced cancer types based on DNA microarray data using ensemble classifiers. *BioMed Research International 2013*, 1–13.

Yu, Q., G. Allon, and A. Bassamboo (2016). How do delay announcements shape customer behavior? an empirical study. *Management Science 63*(1), 1–20.

Yu, Q., G. Allon, A. Bassamboo, and S. Iravani (2017). Managing customer expectations and priorities in service systems. *Management Science 64*(8), 1–29.

Yu, Z., W. B. Rouse, and N. Serban (2011). A Computational Theory of Enterprise Transformation. *Systems Engineering 14*(4), 441–454.

Zhang, W., D. Yang, and H. Wang (2019). Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey. *IEEE Systems Journal 13*(3), 2213–2227.

Zhang, Y., G. Liu, W. Luan, C. Yan, and C. Jiang (2018). An approach to class imbalance problem based on stacking and inverse random under sampling methods. In *2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*, pp. 1–6. IEEE.

Zhang, Y., S. Ren, Y. Liu, T. Sakao, and D. Huisingh (2017). A framework for Big Data driven product lifecycle management. *Journal of Cleaner Production 159*, 229–240.

Zhao, P., Y. Zhang, M. Wu, S. C. Hoi, M. Tan, and J. Huang (2019). Adaptive Cost-Sensitive Online Classification. *IEEE Transactions on Knowledge and Data Engineering 31*(2), 214–228.

Zheng, K., Z. Zhang, and B. Song (2020). E-commerce logistics distribution mode in big-data context: A case analysis of JD.COM. *Industrial Marketing Management 86*, 154–162.

Zhong, R. Y., G. Q. Huang, S. Lan, Q. Y. Dai, X. Chen, and T. Zhang (2015). A big data approach for logistics trajectory discovery from RFID-enabled production data. *International Journal of Production Economics 165*, 260–272.

Zio, E. (2009). Reliability engineering: Old problems and new challenges. *Reliability Engineering & System Safety 94*(2), 125–141.

Zitouni, I., H.-K. J. Kuo, and C.-H. Lee (2003). Boosting and combination of classifiers for natural language call routing systems. *Speech Communication 21*, 647–661.

Zong, W., G.-B. Huang, and Y. Chen (2013). Weighted extreme learning machine for imbalance learning. *Neurocomputing 101*, 229–242.

Vita

Rodrigo Caporali de Andrade

Place of birth	Belo Horizonte, MG, Brazil
Date of birth	November 30, 1990
Education	Stevens Institute of Technology, Hoboken, NJ Doctoral Candidate in Engineering Management Anticipated graduation date: August 2020
	Stevens Institute of Technology, Hoboken, NJ Master of Science in Financial Engineering Graduation date:: May 2016
	Universidade Federal de Minas Gerais, Belo Horizonte, Brazil Bachelor in Mining Engineering Graduation date:: December 2013
Work Experience	Stevens Institute of Technology, Hoboken, NJ Teaching Assistant (January 2018 - May 2020) Lead TA for Eng. Economics class of 300+ students (undergrad and grad levels) to make sure the course run smoothly over the years. Planned and controlled lecture materials, assignments, and exams.
	Chubb, Jersey City, NJ Data Scientist Intern (June 2019 - August 2019) Developed innovative approaches using NLP methodologies to better understand claim processes and generate business. knowledge on based on large claim text data (+400MM notes).
	Stevens Institute of Technology, Hoboken, NJ Research Assistant (June 2016 - December 2017) Worked on a research project for a major U.S. insurance company and developed advanced analytics methods to investigate large datasets from customer support systems. Generated managerial insights for business, analytics, and operational teams
	Vale S.A., Nova Lima, MG, Brazil Mining Engineer Intern (May 2012 - December 2013) Developed projects of mine equipment's supply efficiency focused on operational control and continuous improvement with projected cost reduction impacts of US$ 875,000/year.

Publications	Andrade, R., Grogan, P.T., Moazeni, S. (2020). imulation-based Assessment of Data-Driven Processes in Customer Support Systems. Submitted.
	Moazeni, S., Andrade, R. (2020). A Data-Driven Approach for Consumer Behavior at Voice Self-Service Platform in Insurance Call Centers. To be submitted.
	Andrade, R., Moazeni, S., Ramirez-Marquez, J. E. (2020). A systems perspective on contact centers and customer service reliability modeling. *Systems Engineering, Vol 23,* No. 2, pp. 221-236.
	Moazeni, S., Andrade, R. (2018). A data-driven approach to predict an individual customer's call arrival in multichannel customer support centers. *Proceedings of the 2018 IEEE International Congress on Big Data (BigData Congress), pp.66-73.*
	Andrade, R., Braganca, M., Moazeni, S., (2017). Economic and Political Crisis in Brazil: An Empirical Analysis of the Private Consumption *Available at SSRN 3182066.*
Conferences	Moazeni, S., Andrade, R. (2017). Feature selection for voice self-service transfer rate in insurance call centers. *Presented at INFORMS Annual Meeting Conference*, Houston, TX.
Honors & Awards	Distinguished Leadership by a Ph.D. Student in the School of Systems and Enterprises Stevens Institute of Technology, 2020
	ASEM Outstanding Teaching Assistant Stevens Institute of Technology, 2020
	Graduate Assistantship Stevens Institute of Technology, 2016-2020
	Best Financial Engineering Master's Project Stevens Institute of Technology, 2016
	Graduate Assistantship Ministry of Education of Brazil, 2014-2016