Received 23 August 2022; accepted 29 March 2023. Date of publication 7 April 2023;
date of current version 28 April 2023. The review of this article was arranged by Editor James H. Lambert.

Digital Object Identifier 10.1109/OJSE.2023.3265435

Simulation Assessment of Data-Driven Channel Allocation and Contact Routing in Customer Support Systems

RODRIGO CAPORALI DE ANDRADE, PAUL T. GROGAN (Senior Member, IEEE), AND SOMAYEH MOAZENI (Senior Member, IEEE)

Stevens Institute of Technology, Hoboken, NJ 07030 USA

CORRESPONDING AUTHOR: Paul T. Grogan (e-mail: pgrogan@stevens.edu).

ABSTRACT Data-driven operations management methods can transform company operations, respond rapidly to customer demands, and enable new business models. However, companies face the challenge of measuring and evaluating how new technology will impact operational processes. This article takes a systems engineering approach to assess the tradeoffs of adopting data-driven mechanisms to improve operational processes in a multichannel customer support system. In this article, we investigate potential cost savings from two technology applications: classification methods to direct customers to efficient self-service communication channels and routing methods to match customers with agents based on the query type and available skill set. Discrete event simulation evaluates how new technology adoption affects system-level performance. What-if scenarios combine distinct configurations of customer classification mechanisms and available communication channels to evaluate the reduction in the total number of agents required to meet a target service quality level. Discussion includes practical examples of how operational managers could use experimental information to make strategic operational decisions when adopting data-driven technologies.

INDEX TERMS Data-driven methods, operations management, simulation, systems engineering.

I. INTRODUCTION

Organizations invest in data-driven technologies to improve operational processes and product quality, reduce costs, increase sales and customer engagement, sell experiences, and create new business models [1], [2], [3], [4]. For example, McDonald's invested millions of dollars in 2019 to develop artificial intelligence (AI) and machine learning tools to boost sales [5]. Technologies combine data such as weather conditions, time of day, and customer order history data obtained from reading a license plate to make order recommendations. Similarly, a U.S. bank earned \$2 billion by adopting data-driven initiatives that generated additional revenues and improved operating process efficiency [6].

The service industry provides interesting opportunities to improve business processes by leveraging new technology. Due to its labor intensity and difficulty to mechanize and automate processes [7], increases in operational productivity come largely from improvements in the use of knowledge and application of technology [8]. Machine learning algorithms identify patterns in customer data and translate them into tangible business insights to enhance customer relations, forecasting, and process automation [9], [10], [11].

With the emergence of new communication channels such as emails, online accounts, social media, and mobile apps, traditional call centers have become multichannel customer support systems [12]. Omnichannel data availability and advanced data analytics improve customer experience and enhance customer satisfaction [13], [14]. Data-driven technologies enable contact centers to become more proactive, better assist customers, and enhance their interaction experience [15]. Customer support systems can leverage transactional data to better understand customer queries and provide self-service support to improve routing and employee productivity [14], [16]. Intelligent contact routing can reduce

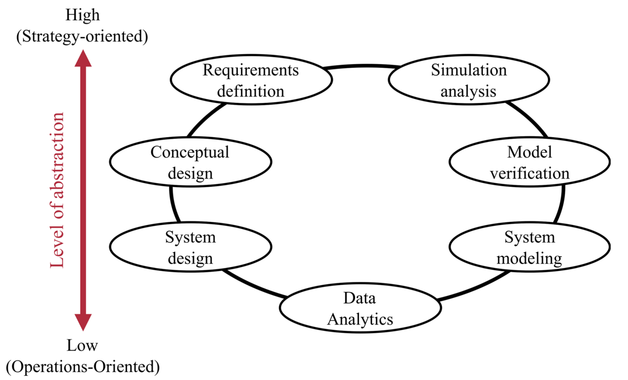

FIGURE 1. Systems engineering approach assesses the ability for data-driven methods to improve customer support system performance by interconnecting strategic goals with operational activities.

customer waiting time, improve contact resolution [17], [18], increasing customer satisfaction, and reducing operations cost [19], [20].

While successful cases motivate the pursuit of data-driven technologies, companies must carefully assess their processes, resources, customer demand, market, and competitors to identify potential benefits and guide technical implementation. The decision to pursue digital transformation is challenging due to technical, computational, operational, and cultural barriers [21], [22], [23], [24], [25], [26]. Unique scenarios require tailored analysis and, as innovation is expensive and technology advances at a rapid pace, the outcome of proposed changes is uncertain. Therefore, company transformation processes require an approach that allows for quick, inexpensive, and flexible assessment to inform decisions.

This article contributes a simulation assessment method to evaluate data-driven technology infusion in a customer support system. A systems engineering approach in Fig. 1 helps to interconnect strategic and operational perspectives. First, we define the system requirements and establish the system boundary. Next, we develop a conceptual design of the customer support system and propose new data-driven components and processes. Finally, simulation analysis assesses how adopting smart technologies and data-driven processes impact the customer support system performance measured by the required staffing levels to meet a target service quality. Alternatives vary self-service contact channels and implement data-driven mechanisms to: 1) allocate customers to channels; and 2) route customers to agents. Results identify which technology combinations can achieve four cost reduction targets of 8%, 13%, 20%, and 27%.

II. CUSTOMER SUPPORT SYSTEMS

Customer support systems strive to create a more dynamic, intuitive, and natural experience for customers. Digital channels combine with other phases of the customer support system to provide detailed data on the customer interaction processes and offer customized and lower-cost self-service alternatives [12], [13], [27].

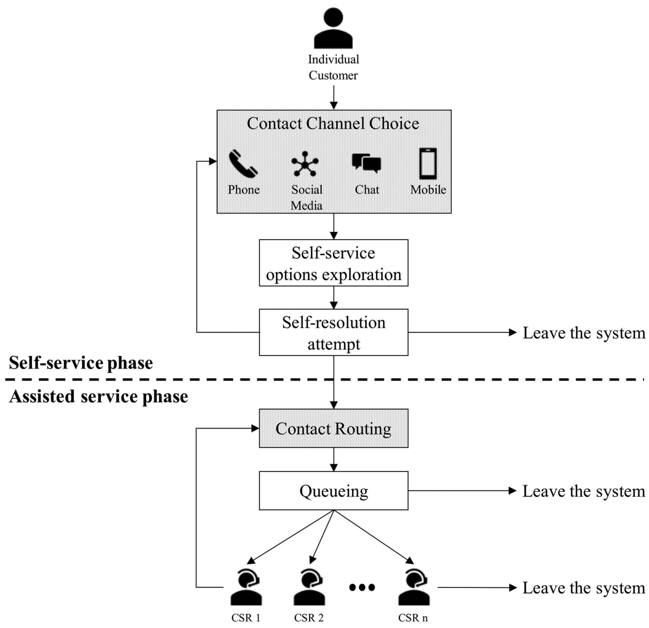

FIGURE 2. Conceptual design of a customer support system consists of self-service and assisted service phases.

This section reviews opportunities for data-driven technology infusion in a generic multichannel customer support system process depicted in Fig. 2. The self-service phase includes customer support requests within a selected contact channel where customers attempt to resolve their query through automated means. If unsuccessful, the customer can opt to request customer service representative (CSR) assistance, seek another communication channel, or abandon the system. Contact distribution combines routing and queuing activities to connect and match customers to CSR agents with the right skills [28]. The system accounts for intraagent transfers, and customers either abandon the system while waiting in queue or after receiving assistance.

A. SELF-SERVICE PHASE

Companies constantly adapt to new trends and customer demands. In particular, smartphones created a new opportunity to connect and relate to customers by expanding the use of self-service environments to lower operational costs while also targeting higher customer satisfaction, retention, and loyalty [29]. Interactive voice response (IVR) allows customers to interact over the phone without CSR assistance [30]. Web platforms allow customers to track orders, review items, cancel unshipped items, and request returns or refunds [31]. Advanced text mining and natural language processing techniques allows chatbots to digest customer data and provide effective and personalized service [32], [33].

Shifting demand from higher-cost assisted channels to self-service channels is a strategic decision to improve customer experience and reduce operations costs [30]. Moazeni and Andrade [15] showed that 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.

B. ASSISTED SERVICE PHASE

Customers who cannot solve queries in the self-service channel may request CSR assistance. *Contact routing* directs customers to CSRs with the appropriate skillset and is an essential step for effective and efficient service. An accurate routing process limits transfers between agents to reduce time in ineffective service, resulting in faster service and higher customer satisfaction [34].

The mechanisms that match customers to agents combine routing and queuing processes. First-in, first-out (FIFO) is a simple queuing discipline; however, more complex service systems with heterogeneous servers and customers often require rules to prioritize certain customers. Skill-based routing allows specialized or cross-trained agents to handle customers based on their type of query [17], [35]. Customers initially 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 email messages [36].

Call centers optimize routing mechanisms with dynamic programming by incorporating agent and customer data to improve performances [17]. Chan et al. [28] proposed a dynamic model for contact routing that combines the agent and customer heterogeneity to match a type of call and a group of agents. Mehrotra et al. [17] investigated dynamic routing strategies to match call types with agent skills, subject to their effectiveness for each call type. Finally, Ibrahim et al. [37] incorporated agent and call type heterogeneity, and time-dependent agent performance to model interdependent service times.

C. DATA-DRIVEN CHANNEL ALLOCATION

Multisource information and data-driven methods can anticipate the reason, time, and channel by which customer queries arrive at a customer support system, and allocate them to the most appropriate channel [15], [38]. For instance, a web channel may be more suitable to solve standardized services [39] while 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 benefit service quality and operational efficiency.

Data-driven technologies can leverage customer attributes from recorded interactions to predict future interactions. Jerath et al. [38] defineed a stochastic function of customer information needs to model 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 multiple factors influence the adoption and use of various channels including customer demographics, the type of information structure available in each communication venue, and the type of query [29], [40], [41]. Furthermore, Jerath et al. [38] emphasized the potential use of more detailed data from telephone conversation transcripts and usage patterns of other channels to fully develop and estimate the model.

D. DATA-DRIVEN CONTACT ROUTING

Enhanced contact routing benefits from customer and query classification using attributes gathered in the self-service phase and from historical customer multichannel interaction data. The high volume, variety, and speed at which data is generated and processed requires 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 [42].

Studies on natural language processing (NLP) improve contact routing processes based on unstructured data collected in real-time during service delivery [43], [44], [45], [46]. NLP techniques seek to make the call routing process more convenient for customers using speech-enabled systems rather than closed-menu systems. The challenge, therefore, is to understand the query from the customer's own description, rather than using touch-tone or speech-enabled menus [16].

In addition to query type and training, contact center data can also help pair customers and CSRs based on behavioral profile similarity. Social homophily is the process by which individuals form bonds based on shared characteristics [47]. Ali [48] proposed a contact routing model that used several machine learning methods to maximize customer satisfaction based on underlying customer and CSR demographics, psychographics, and historical performance data.

E. RESEARCH OBJECTIVE

Companies can achieve a competitive advantage by 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 [49].

However, when evaluating significant changes to the contact center system design, historical data are 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. 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

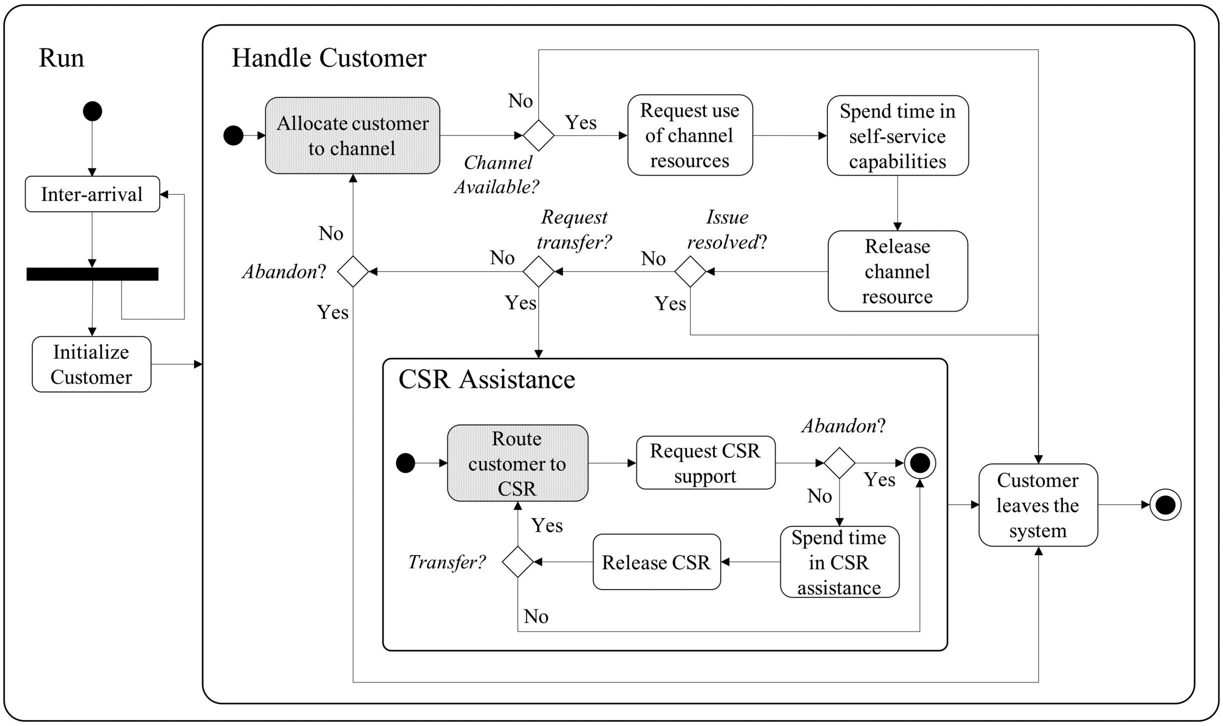

FIGURE 3. From a process-interaction worldview, the simulation model consists of activities that handle the generation, channel allocation, and routing of customers through the self-service and assisted service phases of a multichannel customer support system.

many times [50]. Relying on synthetic data from simulation models is ideally suited to describe real-life tradeoffs.

To address these concerns, this study develops a simulation-based assessment of a multidimensional tradeoff between operational management, system-level strategies, and cost savings from reduced labor. Results can inform the relative impact of alternative data-driven technologies and determine what level of performance is required at the operational management level to achieve desired performance improvement at the customer service system level.

III. SIMULATION MODEL DEVELOPMENT

Simulation provides an analysis method to investigate how lower-level operational decisions (e.g., introduction of a new data-driven technology) affect system-level performance such as service level, answer time, and customer satisfaction.

This section models a customer support system to investigate three data-driven technology infusion 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.

We define four customer support system configurations and simulate the system-level effects from changes to operational management by varying performance levels of the channel allocation platform (2) and contact routing mechanism (3).

A. PROCESS-INTERACTION ACTIVITY DIAGRAM

The activity diagram in Fig. 3 describes the customer support service system as a discrete event simulation from a process-interaction worldview with Run, Handle Customer, CSR assistance activities. The gray boxes indicate the processes influenced by the data-driven technologies of interest: channel allocation platform, and contact routing mechanism.

The 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.

The Handle Customer activity processes self-service and assisted service phases. During the self-service stage, a platform allocates the customer to a channel. The customer checks the channel availability (i.e., due to scheduled or unscheduled maintenance) and attempts to resolve the query. The customer either leaves the system if the query is solved or abandons without requesting a CSR assistance.

The CSR assistance activity processes a customer request for additional human-agent support. First, the contact routing mechanism directs the customer to an agent who may or may not have the skills to address the query due to imperfect classification. 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 reclassified and transferred to new agent with the appropriate skill set.

B. MODEL DETAILS AND ASSUMPTIONS

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

Individual customer: Customers arrive following a Poisson process with rate λ and have one of two query types $q \in \{A, B\}$ sampled from a discrete uniform distribution. Customer patience (time until abandoning a call) follows an exponential distribution with mean time θ, independently generated each time they are put "on-hold."

Contact channel: Customers have the option of two contact channels, Telephone (T) and Web (W), directed by a channel allocation platform. The random variable $S \sim Bernoulli(\gamma_1)$ determines whether the allocated channel $c \in \{T, W\}$ is best $(S = 1)$ based on highest contact resolution probability ρ_c and capability to solve query q. After allocation, the random variable $Y \sim Bernoulli(\rho_c)$ denotes whether query was successfully resolved $(Y = 1)$ after a customer uses channel c. The random variable $X \sim Bernoulli(\varrho_c)$ defines whether an unresolved query in channel c leads to a CSR assistance request $(X = 1)$ or abandonment $(X = 0)$. 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 $(V = 1)$ at the time a customer seeks support. 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 matches the customer to an appropriate agent $(Z = 1)$ where γ_2 is referred to as the *contact routing accuracy probability*. We assume the automated contact routing platform has an adaptive capability to allocate customers to the CSR queue and follows an FIFO rule.

CSR assistance: We assume heterogeneous CSRs with no cross-training and service time dependent on the agent skill set and contact routing. The service time for customers matched to an agent lacking the skill set to resolve their query follows an 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 follows an exponential distribution with rate ν_q. Finally, all CSRs solve customers queries with a fixed contact resolution probability $\varphi \in [0, 1]$.

Table 1 shows the complete list of model input parameters and associated values further discussed in Section III-E.

TABLE 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	%

C. SYSTEM PERFORMANCE METRICS

Contact centers use several metrics to assess operational performance (e.g., see [51]). Besides controlling the customer service system performance, metrics are also important for operational planning, such as workforce scheduling based on a target performance. Metrics serve as constraints imposed on optimization to minimize the 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 meet a determined service quality. Definition and measurement of service quality is widely discussed in the service industry [14]. For simplicity, we use the service level (SL) to measure service quality where SL is the percentage of customers that wait less than a target duration to be served by a CSR. As the workforce can represent up to 70% of the budget in call centers [51], 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 $i = \{1, 2, \ldots, I\}$ 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

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

where $\Gamma(t)$ is an indicator function which returns 1 if $t \leq \tau$, i.e., the waiting time does not exceed the threshold τ, and otherwise returns 0.

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

$$SL_q(\tau) = \frac{\sum_{i \in \mathrm{G_q}} \Gamma(t_i)}{|G_q|} \qquad \forall q \in \{A, B\} \tag{2}$$

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:

$$\begin{aligned} \min \quad & \sum_{q\in\{A,B\}} N_q \\ \text{subject to} \quad & SL(\tau) \geq \phi \\ & SL_q(\tau) \geq \phi_q \qquad \forall q \in \{A,B\} \\ & N_q \geq 0 \qquad \forall q \in \{A,B\}. \end{aligned} \tag{3}$$

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

D. MODEL IMPLEMENTATION

The simulation model is implemented in Python 3.7 using SimPy (v. 3.0.11), an open-source process-oriented discrete-event simulation library [52], and the source code is available on a GitHub repository.[1]

Four object-oriented classes based on the activity diagram include: Customer, Channel, Agent, and Contact Center. The first three define element attributes while the last configures the simulation. 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, list of capabilities, contact resolution probability, and channel availability. The Agent class defines the CSR training type, average service time, contact resolution probability, and number of agents. Both Channel and Agent classes are defined as shared resources. Finally, the Contact Center class conducts the simulation by generating demands and executing the Handle Customer function that implements the processes illustrated in the activity diagram.

E. MODEL VERIFICATION AND VALIDATION

Model verification compares simulation results to a steady-state Erlang A model to confirm behavior under nominal conditions. Verification tests vary the arrival rate λ, average service time μ, number of servers N, and patience θ to generate results for 2000 min following a 1000 min warm-up to satisfactorily compare the abandon rate, average time in the system, and service level to those of Erlang A.

Validation evaluates whether the simulation model resembles the real-world system of interest. Our simulation model reflects key attributes of real customer support systems such as multiple communication channels with distinct capabilities, capacity, and efficiency, heterogeneous customers and agents, and multiple contact types. Literature emphasizes heterogeneous customers and agents, multiple query types, and a multichannel environment to connect analysis conclusions to managerial recommendations [36], [38], [53]. Our model relies on typical contact center requirements, assumptions from related literature, and real data gathered from a major U.S. insurance company (see [15] for details).

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

TABLE 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

Contact center customer arrival rates vary depending on the month of the year, day of the week, time of the day, and contact channel type [19], [54], [55]. 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.

Customer arrivals typically follow a Poisson process [56], [57], [58], [59]. To determine the average patience time we assume a rule-of-thumb of twice the average service time [53]. 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 [60]. Finally, based on insurance company data, we assume customers have an 80% probability to request assistance in nonresolved interactions in self-service systems.

The target acceptable waiting time in a queue that defines the service level is usually in the range of 20-30 s. For example, [56], [57], [61] assume a 20 s threshold, while the insurance company data (explored in [15] and [62]) contain attributes that flag transactions answered within 30 s. Contact centers often target service level of 80% [28], [56], [57], [61]. Therefore, the desired service quality level in our simulations is set to be greater than 80% defined by a 30-s threshold.

IV. ASSESSMENT OF DATA-DRIVEN METHODS

A. EXPERIMENT DESIGN

Simulation runs investigate four customer service support system configurations in Table 2 based on available communication channels and contact resolution capabilities.

Configuration 1 includes the single contact channel “Telephone” and can handle two query types. Other configurations add a new channel “Web,” which has a higher contact resolution capacity and is, therefore, more efficient. Configurations 2 and 3 consider two-channels with the Web channel intended to solve only one query type. Configuration 4 assumes both channels can solve both query types. We choose two channels and two query types to represent the simplest model that affords for such system design choices.

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 min. For each run, the first 1000 min are used as a warm-up period followed by 2000 min of data recording.

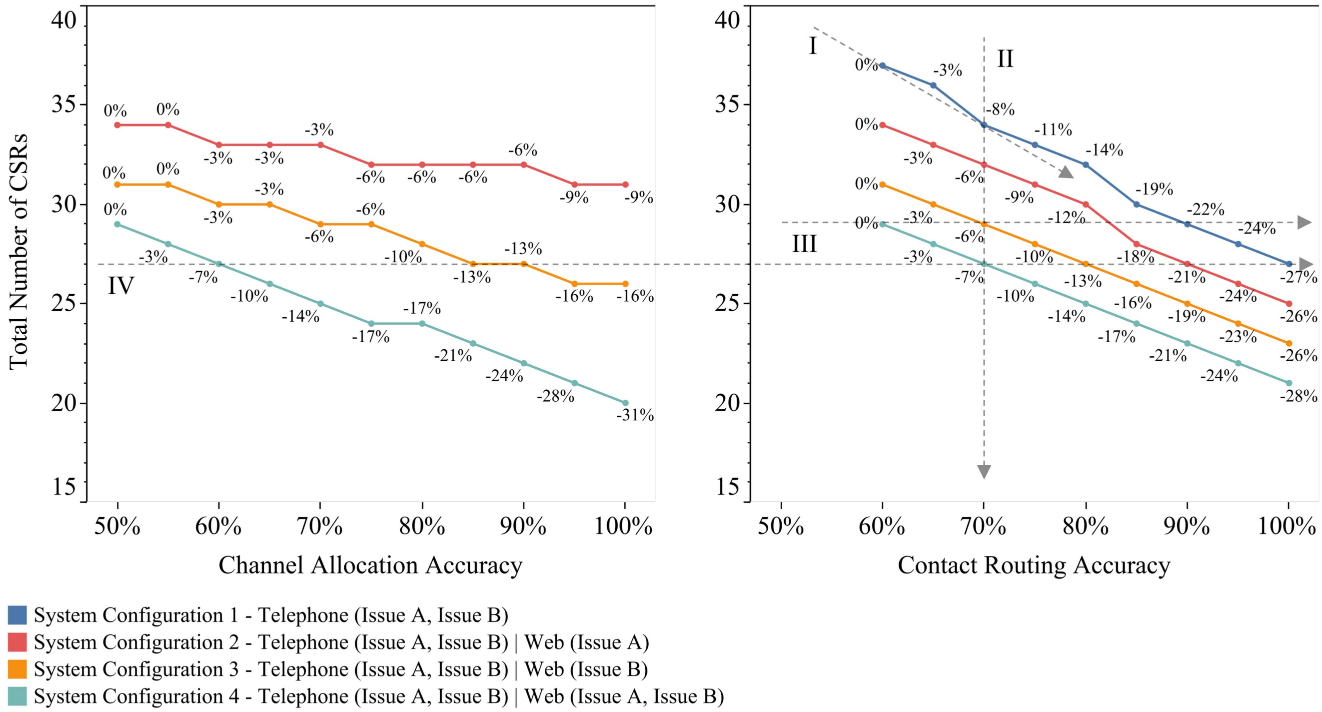

FIGURE 4. **Simulation results: Total number of customer service representatives necessary to maintain a 80% service level for distinct combinations of channel allocation accuracy, contact routing accuracy, and system configurations.**

B. SIMULATION RESULTS AND DISCUSSION

Fig. 4 summarizes the simulation results. The left and right panels correspond to scenarios varying the channel allocation accuracy and the contact routing accuracy, respectively. Lines show the number of CSRs required to solve the staffing problem for each configuration in Table 2 per (3). Labels show the relative percentage change in agents necessary to maintain the target service level for each scenario.

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

$$r_1 = \frac{N(\gamma_1, \gamma_2 = 0.6, \Omega_k)}{N(\gamma_1 = 0.5, \gamma_2 = 0.6, \Omega_k)} - 1 \tag{4}$$

and

$$r_2 = \frac{N(\gamma_1 = 0.5, \gamma_2, \Omega_k)}{N(\gamma_1 = 0.5, \gamma_2 = 0.6, \Omega_k)} - 1 \tag{5}$$

be the percent changes in the total number of CSRs shown, respectively, on the left and on the right charts of Fig. 4.

Dashed lines I, II, III, and IV illustrate how simulation results can guide operations manager decisions when considering adopting data-driven technologies in the customer support system. Four hypothetical scenarios discuss recommendations based on simulation results.

Scenario I aims to reduce operational costs by developing more efficient contact routing software with a Telephone channel. Line I shows the effect of improving contact routing accuracy on the number of CSRs necessary to maintain service quality, allowing a more clear financial trade between nonrecurring software development and recurring operations expenses. For example, an increase from 60% to 70% in routing accuracy allows a 8% reduction from 37 to 34 agents.

After deciding to develop data-driven technologies that improve contact routing accuracy to 70%, scenario II seeks opportunities to reduce labor to 27 CSRs. Enhancing contact classification and routing optimization algorithms to match customers to agents above 70% accuracy may be technically complex and financially prohibitive. Therefore, investing in additional channels becomes an attractive alternative. Line II illustrates how the company can meet this target by developing Web channel capabilities while maintaining the same contact routing performance.

Scenario III seeks 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 goal.

1) Develop contact routing to reach 90% accuracy (unlikely due to intrinsic complexity, e.g., customer inputs using closed-menu and speech-enabled systems, chatbots, and query classification mechanisms).
2) Develop contact routing to reach 82.5% accuracy and add a Web channel capable of solving type A queries.
3) Develop contact routing to reach 70% accuracy and add a Web channel capable of solving type B queries.
4) Develop a Web channel capable of solving both types of queries.

Finally, scenario IV analyzes how to achieve a 27% workforce reduction (10 CSRs) by enhancing contact routing accuracy, adding new channels, or adopting data-driven technologies to match customers to the most efficient channel. Line IV guide shows the following six alternative decisions.

1) Develop contact routing to reach 100% accuracy (unlikely due to intrinsic complexity).
2) Develop contact routing to reach 90% accuracy and add a Web channel capable of solving type A queries.
3) Develop contact routing to reach 80% accuracy and add a Web channel capable of solving type B queries.
4) Develop contact routing to reach 70% accuracy and add a Web channel capable of solving both types of queries.
5) Develop channel allocation to reach 85% accuracy and add a Web channel capable of solving type B queries.
6) Develop channel allocation to reach 60% accuracy and add a Web channel that can solve both types of queries.

Results show the addition of a Web channel leads to considerable staffing cost reduction, even without modifying the channel allocation and contact routing accuracy. Apart from the scenario considering both Telephone and Web channels capable of solving both query types, improvements in contact routing accuracy leads to larger cost reduction compared to the channel allocation accuracy. The combination of adding a new channel and improvements to customer-channel allocation or contact routing boosts its benefits.

Staffing reductions present distinct patterns depending on the system design and Web channel configuration. Although there is a mostly linear relationship between accuracy and number of agents, the rate at which the costs decrease is different for each scenario because the mean service time for query B is twice as long as query A. 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. Results also show discrete changes, as the workforce optimization only allows integer values. 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.

C. IMPLICATIONS FOR OPERATIONS MANAGEMENT

Fig. 4 enables a multidimensional tradeoff analysis between operational and strategic issues, by allowing more clear comparison of recurring cost savings to nonrecurring development costs. While this is a notional case not intended to represent a real company, it demonstrates a simulation-based methodology to assess the tradeoffs of data-driven methods in customer service support systems for operations management. Open source libraries such as SimPy allow 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. 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 recommendations suggest significant reductions in operating costs from developing new technology, it is important to highlight the limitations 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 15 min. Contact volume experiences high volatility throughout the day, which requires a dynamic workforce scheduling to achieve optimal operating costs. Therefore, more detailed models should incorporate mechanisms to account for regular variation to allow a more realistic assessment of the advantages of adopting the aforementioned 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 intraagent 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.

V. CONCLUSION

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, email, 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 article 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.

REFERENCES

[1] S. Ransbotham, D. Kiron, P. Gerbert, and M. Reeves, "Reshaping business with artificial intelligence," *MIT Sloan Manage. Rev. Boston Consulting Group*, vol. 59, no. 1, pp. 1–17, 2017.

[2] M. C. Cohen, "Big data and service operations," *Prod. Operations Manage.*, vol. 27, no. 9, pp. 1709–1723, 2018.

[3] T. H. Davenport and R. Ronanki, "Artificial intelligence for the real world," *Harvard Bus. Rev.*, vol. 96, no. 1, pp. 108–116, 2018.

[4] N. Hanifa, M. Zaidin, M. Nurazri, M. Diah, H. Yee, and S. Sorooshian, "Quality management in industry 4.0 era," *J. Manage. Sci.*, vol. 4, no. 3, pp. 2250–1819, 2014.

[5] D. Yaffe-Bellany, "Would you like fries with that? McDonald's already knows the answer," The New York Times, 2019.

[6] C. Brocchi, D. Grande, K. Rowshankish, T. Saleh, and A. Weinberg, "Designing a data transformation that delivers value right from the start," *McKinsey on Payments*, vol. 11, no. 28, pp. 52–59, 2018.

[7] B. S. Sahay, "Multi-factor productivity measurement model for service organisation," *Int. J. Product. Perform. Manage.*, vol. 54, no. 1, pp. 7–22, 2005.

[8] J. Heizer and B. Render, *Operations Management: Sustainability and Supply Chain Management*, 11th ed. Upper Saddle River, NJ, USA: Pearson Education, Inc., 2014.

[9] J. Chen, Y. Tao, H. Wang, and T. Chen, "Big data based fraud risk management at Alibaba," *J. Finance Data Sci.*, vol. 1, no. 1, pp. 1–10, 2015.

[10] M. Andrews, X. Luo, Z. Fang, and A. Ghose, "Mobile ad effectiveness: Hyper-contextual targeting with crowdedness," *Marketing Sci.*, vol. 35, no. 2, pp. 218–233, 2015.

[11] S. Moazeni, B. Defourny, and M. J. Wilczak, "Sequential learning in designing marketing campaigns for market entry," *Manage. Sci.*, vol. 66, no. 9, pp. 1–20, 2019.

[12] M. Saberi, O. K. Hussain, and E. Chang, "Past, present and future of contact centers: A literature review," *Bus. Process Manage. J.*, vol. 23, no. 3, pp. 574–597, 2017.

[13] Deloitte, "Global contact center survey," Deloitte, London, U.K., White paper, 2017. [Online]. Available: https://www2.deloitte.com/content/dam/Deloitte/us/Documents/strategy/us-con-2017-global-contact-center-survey.pdf

[14] R. Andrade, S. Moazeni, and J. E. Ramirez-Marquez, "A systems perspective on contact centers and customer service reliability modeling," *Syst. Eng.*, vol. 23, no. 2, pp. 221–236, 2020.

[15] S. Moazeni and R. Andrade, "A data-driven approach to predict an individual customer's call arrival in multichannel customer support centers," in *Proc. IEEE Int. Congress Big Data*, 2018, pp. 66–73.

[16] S. Rustamov, E. Mustafayev, and M. A. Clements, "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 Eng.*, vol. 8, no. 1, pp. 61–68, 2018.

[17] V. Mehrotra, K. Ross, G. Ryder, and Y. P. Zhou, "Routing to manage resolution and waiting time in call centers with heterogeneous servers," *Manuf. Service Operations Manage.*, vol. 14, no. 1, pp. 66–81, 2012.

[18] C. Taylor, "Combine the power of IVR and WFO analytics to evolve agent and customer experience," Forrester Res., Inc., Cambridge, MA, USA, White paper. 2017. [Online]. Available: https://www.verint.com/wp-content/uploads/Forrester-Combine-the-Power-of-IVR-WFO-Analytics-to-Evolve-Agent-and-Cust-Experience-1.pdf

[19] Z. Aksin, M. Armony, and V. Mehrotra, "The modern call-center: A multi-disciplinary perspective on operations management research," *Prod. Operations Manage.*, vol. 16, no. 6, pp. 665–688, 2007.

[20] A. A. Seada and A. B. Eltawil, "Modeling and analysis of workforce management decisions in modern call centers," *in Proc. Int. Conf. Ind. Eng. Oper. Manage.*, 2015, pp. 1–10.

[21] R. Geissbauer, J. Vedso, and S. Schrauf, "Industry 4.0: Building the digital enterprise," PwC, London, U.K., White paper, 2016. [Online]. Available: https://www.pwc.com/gx/en/industries/industries-4.0/landing-page/industry-4.0-building-your-digital-enterprise-april-2016.pdf

[22] S. J. Berman, "Digital transformation: Opportunities to create new business models," *Strategy Leadership*, vol. 40, no. 2, pp. 16–24, 2012.

[23] E. Kurti and D. Haftor, "Barriers and enablers of digital business model transformation," in *Proc. Eur. Conf. Inf. Syst. Manage.*, 2015, pp. 262–268.

[24] C. Matt, T. Hess, and A. Benlian, "Digital transformation strategies," *Bus. Inf. Syst. Eng.*, vol. 57, no. 5, pp. 339–343, 2015.

[25] T. V. Leipzig et al., "Initialising customer-orientated digital transformation in enterprises," *Procedia Manuf.*, vol. 8, pp. 517–524, 2017.

[26] D. Kiel, J. M. Muller, C. Arnold, and K.-I. Voigt, "Sustainable industrial value creation: Benefits and challenges of industry 4.0," *Int. J. Innov. Manage.*, vol. 21, no. 8, pp. 1–34, 2017.

[27] F. Mueller, H. Naujoks, H. Singh, G. Schwarz, A. Schwedel, and K. Thomson, "Global digital insurance benchmarking report 2015 - pathways to success in a digital world," Bain & Company Inc., Boston, MA, USA, White paper, 2015. [Online]. Available: https://media.bain.com/Images/GLOBAL-DIGITAL-INSURANCE-2015.pdf

[28] W. Chan, G. Koole, and P. L'Ecuyer, "Dynamic call center routing policies using call waiting and agent idle times," *Manuf. Service Oper. Manage.*, vol. 16, no. 4, pp. 544–560, 2014.

[29] A. Kumar and R. Telang, "Does the web reduce customer service cost? empirical evidence from a call center," *Inf. Syst. Res.*, vol. 23, no. 3-part-1, pp. 721–737, 2012.

[30] T. Tezcan and B. Behzad, "Robust design and control of call centers with flexible interactive voice response systems," *Manuf. Service Operations Manage.*, vol. 14, no. 3, pp. 355–484, 2012.
[31] M. F. Keblis and M. Chen, "Improving customer service operations at Amazon.com," *Interfaces*, vol. 36, no. 5, pp. 433–445, 2006.
[32] A. Scherer, N. V. Wünderlich, and F. V. Wangenheim, "The value of self-service: Long-term effects of technology-based self-service usage on customer retention," *MIS Quart.*, vol. 39, no. 1, pp. 177–200, 2015.
[33] T. Verhagen, J. Van Nes, F. Feldberg, and W. Van Dolen, "Virtual customer service agents: Using social presence and personalization to shape online service encounters," *J. Comput.-Mediated Commun.*, vol. 19, no. 3, pp. 529–545, 2014.
[34] M. Tang, B. Pellom, and K. Hacioglu, "Call-type classification and unsupervised training for the call center domain," in *Proc. IEEE Workshop Autom. Speech Recognit. Understanding*, 2003, pp. 204–208.
[35] L. F. Ali III, "A call center simulation study: Comparing the reliability of cross-trained agents to specialized agents," Master's thesis, Graduate Sch., Univ. Tennessee, Knoxville, TN, USA2010.
[36] M. Ahghari and B. Balcioglu, "Benefits of cross-training in a skill-based routing contact center with priority queues and impatient customers," *IIE Trans.*, vol. 41, no. 6, pp. 524–536, 2009.
[37] R. Ibrahim, P. L'Ecuyer, H. Shen, and M. Thiongane, "Inter-dependent, heterogeneous, and time-varying service-time distributions in call centers," *Eur. J. Oper. Res.*, vol. 250, no. 2, pp. 480–492, 2016.
[38] K. Jerath, A. Kumar, and S. Netessine, "An information stock model of customer behavior in multichannel customer support services," *Manuf. Service Oper. Manage.*, vol. 17, no. 3, pp. 368–383, 2015.
[39] K. K. Boyer, R. Hallowell, and A. V. Roth, "E-services- operating strategy—a case study and a method for analyzing operational benefits," *J. Operations Manage.*, vol. 20, no. 2, pp. 172–188, 2002.
[40] M. Xue, L. M. Hitt, and P. T. Harker, "Customer efficiency, channel usage, and firm performance in retail banking," *Manuf. Service Operations Manage.*, vol. 9, no. 4, pp. 535–558, 2007.
[41] D. Campbell and F. Frei, "Cost structure, customer profitability, and retention implications of self-service distribution channels: Evidence from customer behavior in an online banking channel," *Manage. Sci.*, vol. 56, no. 1, pp. 4–24, 2010.
[42] A. Schoeller and R. Heffner, "Design your contact center from the outside in," Forrester Research, Inc., Cambridge, MA, USA, White paper, 2014.
[43] J. Chu-Carroll and B. Carpenter, "Vector-based natural language call routing," *Comput. Linguistics*, vol. 25, no. 3, pp. 361–388, 1999.
[44] I. Zitouni, H.-K. J. Kuo, and C.-H. Lee, "Boosting and combination of classifiers for natural language call routing systems," *Speech Commun.*, vol. 21, pp. 647–661, 2003.
[45] N. Tyson and V. Matula, "Improved lsi-based natural language call routing using speech recognition confidence scores," in *Proc. 2nd IEEE Int. Conf. Comput. Cybern.*, 2004, pp. 409–413.
[46] R. Sarikaya, G. E. Hinton, and A. Deoras, "Application of deep belief networks for natural language understanding," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 22, no. 4, pp. 778–784, Apr. 2014.
[47] B. W. Domingue, D. W. Belsky, J. M. Fletcher, D. Conley, J. D. Boardman, and K. M. Harris, "The social genome of friends and schoolmates in the national longitudinal study of adolescent to adult health," *Proc. Nat. Acad. Sci. United States Amer.*, vol. 115, no. 4, pp. 702–707, 2018.
[48] A. R. Ali, "Intelligent call routing: Optimizing contact center throughput," in *Proc. 11th Int. Workshop Multimedia Data Mining*, 2011, pp. 1–9.
[49] S. Erevelles, N. Fukawa, and L. Swayne, "Big data consumer analytics and the transformation of marketing," *J. Bus. Res.*, vol. 69, no. 2, pp. 897–904, 2016.
[50] S. L. Rosen, C. P. Saunders, and S. K. Guharay, "A structured approach for rapidly mapping multilevel system measures via simulation metamodeling," *Syst. Eng.*, vol. 18, no. 1, pp. 87–101, 2015.
[51] N. Gans, G. Koole, and A. Mandelbaum, "Telephone call centers: Tutorial, review, and research prospects," *Manuf. Service Operations Manage.*, vol. 5, no. 2, pp. 79–141, 2003.
[52] A. Meurer et al., "SymPy: Symbolic computing in python," *PeerJ Comput. Sci.*, vol. 3, 2017, Art. no. e103.
[53] A. Mandelbaum and S. Zeltyn, "Staffing many-server queues with impatient customers: Constraint satisfaction in call centers," *Operations Res.*, vol. 57, no. 5, pp. 1189–1205, 2009.
[54] R. Andrade and S. Moazeni, "Feature selection for voice self-service transfer rate in insurance call centers," presented at INFORMS Annu. Meeting Conf., Houston, TX, USA, 2017.
[55] J. W. Taylor, "Density forecasting of intraday call center arrivals using models based on exponential smoothing," *Manage. Sci.*, vol. 58, no. 3, pp. 534–549, 2012.
[56] M. T. Cezik and P. L'Ecuyer, "Staffing multiskill call centers via linear programming and simulation," *Manage. Sci.*, vol. 54, no. 2, pp. 310–323, 2008.
[57] A. N. Avramidis, W. Chan, M. Gendreau, P. L'Ecuyer, and O. Pisacane, "Optimizing daily agent scheduling in a multiskill call center," *Eur. J. Oper. Res.*, vol. 200, no. 3, pp. 822–832, 2010.
[58] P. Khudyakov, P. D. Feigin, and A. Mandelbaum, "Designing a call center with an IVR (interactive voice response)," *Queueing Syst.*, vol. 66, no. 3, pp. 215–237, 2010.
[59] R. Ibrahim and P. L'Ecuyer, "Forecasting call center arrivals: Fixed-effects, mixed-effects, and bivariate models," *Manuf. Service Operations Manage.*, vol. 15, no. 1, pp. 72–85, 2012.
[60] N. Gans, N. Liu, A. Mandelbaum, H. Shen, and H. Ye, "Service times in call centers: Agent heterogeneity and learning with some operational consequences," in *Borrowing Strength: Theory Powering Applications—A Festschrift for Lawrence D. Brown*. Institute of Mathematical Statistics, Beachwood, OH, USA, 2010, pp. 99–123.
[61] A. Pot, S. Bhulai, and G. Koole, "A simple staffing method for multiskill call centers," *Manuf. Service Oper. Manage.*, vol. 10, no. 3, pp. 421–428, 2008.
[62] S. Moazeni, "Predicting individual-level call arrival from online account customer activity," in *Proc. IEEE Int. Conf. Big Data*, 2018, pp. 1624–1631.

RODRIGO CAPORALI DE ANDRADE received the bachelor's degree in mining engineering from the Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, in 2013, and the M.E. degree in financial engineering and the Ph.D. degree in engineering management from the Stevens Institute of Technology, Hoboken, NJ, USA, in 2016 and 2020, respectively.

From 2020 to 2022 he was a Senior Quantitative Risk Analyst with Bunge, Chesterfield, MO, USA, where, he has been a Senior Manager of Risk Analytics since 2022.

PAUL T. GROGAN (Senior Member, IEEE) received the B.S. degree in engineering mechanics and astronautics from the University of Wisconsin, Madison, WI, USA, in 2008, and the S.M. degree in aeronautics and astronautics and the Ph.D. degree in engineering systems from the Massachusetts Institute of Technology, Cambridge, MA, USA in 2010 and 2014, respectively.

From 2016 to 2022, he was an Assistant Professor with the School of Systems and Enterprises, Stevens Institute of Technology, Hoboken, NJ, USA, where he has been an Associate Professor since 2023.

SOMAYEH MOAZENI (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in mathematics in 2005 from the Amirkabir University of Technology, Tehran, Iran, and the M.Math degree in combinatorics and optimization and the Ph.D. degree in computer science from the University of Waterloo, Waterloo, ON, Canada in 2006 and 2012, respectively.

From 2014 to 2016, she was an Assistant Professor with the School of Systems and Enterprises and from 2017 to 2021 an Assistant Professor with the School of Business, Stevens Institute of Technology, Hoboken, NJ, USA, where she has been an Associate Professor since 2021.