

The Staffing Problem of the N-Design Multi-Skill Call Center Based on Queuing Model

Chun-Yan Li ^{1, 2}

¹ Zhijiang College of Zhejiang University of Technology
Hangzhou, 310024, China

² School of Economics and Management, Yanshan
University
Qinhuangdao, 066004, China
llccyy1980@126.com

De-Quan Yue

School of Economics and Management, Yanshan
University
Qinhuangdao, 066004, China

Abstract—This paper studied the staffing problem of a multi-skill call center in N-design. In this model, there are two types of customers and two groups of servers who have different skills. We proposed a new state space division method, according to the different state of each agent groups, dividing the infinite number of system states into several finite state sets, and obtained the state-transition rates. Then, we established equations for the steady-state probabilities of the system. Finally, we obtained the computational formula for the service level and we established the staffing calculation model for optimal number of the agents in each group, we adopted the implicit enumeration method to find the solution.

Keywords—Queuing model; Multi-skill call center; Steady-state probabilities; Service level; Staffing problem.

I. INTRODUCTION

Recent years, the rapid developments in technology and the business environment have dramatically increased the need for call center service systems. A call center is defined as a service system in which agents or servers serve callers or customers over the telephone, fax, email, etc. This comprehensive information service system relies on computer telephone integration technology. The call centers are widely used in communication, finance, e-business, emergency centers and so on.

Call centers running well, must be up to the balance between service costs and service level. Call centers must guarantee that at any time the presence of a suitable number of agents and a suitable amount of skills according to the forecasted level of demand, this problem is usually referred to as the staffing problem. As the importance and complexity of modern call centers grow, there is a proliferation of literature relating to them, typically focusing on queuing models. In a queuing model of a call center, the call agents and calls correspond to servers and customers, respectively. For important related surveys, we refer to Koole and Mandelbaum [1] and Gans et al. [2]. An introduction to staffing problems with relevant bibliographic references can be found in Aksin et al. [3].

Most of the early literature on queuing models of call centers assumes that call centers consist of a single type of incoming call and a single pool of homogeneous agents. Such call centers are referred to as single-skill call centers. However, as telecommunications and information technology have advanced over the years, call centers often handle several types of calls distinguished by the required skills for delivering service. Accordingly, multi-skill call centers have emerged and have recently been studied in the literature. A multi-skill call center handles several types of calls, and each agent has different skills. A typical example is an international call center where incoming calls are in different languages, see Gan et al. [2].

Perry and Nilsson [4] consider a V-design in which two classes of calls are served by a single group of crossed-trained agents. The model represents a single group of operators that provides directory assistance, as well as serving toll/assist calls. Bhulai and Koole [5] proposed scheduling policies and show that the policy is optimal for equal service time distributions and a very good approximation otherwise. Gans and Zhou [6] also studied the same V-design model, they using the Linear Programming Approach, Their optimal strategy is a critical value presupposition strategy. Other researches about V-design model see e.g. [7], [8], [9], [10] and the references therein.

Stanford and Grassman considered a queuing model of the N-design call center which has a fixed priority strategy [11]. Shumsky proposed an approximate analysis of this model [12]. In an N-design call center, two groups of agents handle two types of calls. Group 1 serves types 1 and 2; and group 2 serves types 2 only. In the actual call center system, many have two types of customers. Such as bilingual call center, which have Chinese inbound and foreign inbound two types of incoming calls, this requires the agents master two languages in both English and Chinese at the same time, but the cost of this kind agents is relatively high. So in the call center, there are only some agents master two languages, the other part only master one language. Such there are two groups of agents, one group with two kinds of skill, the other group only has one skill.

In this paper we study the staffing problem of the N-design multi-skill call center. By the analysis above, the current research on the N-design call center is few, and the staffing problem calculation mostly with the method of linear programming and simulation. One of the important calculations of the service level is obtained by simulation method, but the simulation method needs longer computing time, and there is no specific calculation formula. So this paper studies the staffing problem of the N-design multi-skill call center using the method of queuing model.

The rest of the paper is organized as follows. In Section 2, we describe the model of the multi-skill call center in N-design. In Section 3, we present a new method of solving the model by dividing the system's state space into several sets of states. Then, we obtain the equilibrium equations and the steady-state probabilities. In Section 4, we obtain the specific computational formula of the service level and the computational procedure of the staffing problem. Section 5 concludes the paper and highlights some future research.

II. SYSTEM MODEL

This paper studies a classical N-design multi-skill call center queue model, the model has two types of calls and two server groups. The model is shown in Fig. 1.

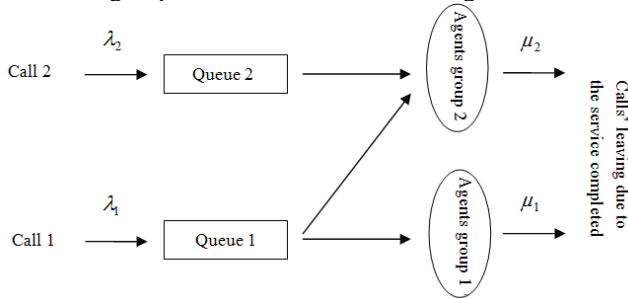


Figure 1. The N-design multi-skill call center queue model.

Arrival process: There are two types of calls (or customers), Call 1 and Call 2. The two types of calls arrive according to a Poisson process with rates λ_1 and λ_2 , respectively. There are two queues, Queue 1 and Queue 2, which consist of customers of Call 1 and Call 2, respectively. We assume that customers have enough patience and the calls through the call center's selection system can be accurately classified.

Service process: There are two groups of server, Group 1 with N_1 servers, Group 2 with N_2 servers. Group 1 is of specialized servers who can only serve customers of Call 1, Group 2 is of flexible servers who can serve customers of both Call 1 and Call 2. The service times of servers in Group 1 and 2 are all exponentially distributed with means $\frac{1}{\mu_1}$ and $\frac{1}{\mu_2}$, respectively.

Routing policy: The routing policy of the model is based on skills and the importance of the two different types of calls. It is assumed that Call 1 is important than Call 2. In other words, Call 1 has non-preemptive priority. When a server in Group 2 completes service, if there are

customers of Call 1 waited in Queue 1 this server will service a customer waited in Queue 1, otherwise this server will serve a customer waited in Queue 2. When a server in Group 1 completes service, he will serve a customer of Call 1 waited in Queue 1, otherwise he will be idle.

Queuing discipline: There are infinite waiting spaces for both queues. For the same type calls, they are served in First Come First Served (FCFS) discipline. The queues of Call 1 and Call 2 are independent of each other.

III. THE CALCULATION OF THE STEADY-STATE PROBABILITY

In this section, we firstly define the states of the model. Then, we derive the state-transition rates by using results of an M/M/c/c queuing system. Finally, we obtain equations for the steady-state probabilities of the system. It is assumed that the system is stationary.

A. The Division of the State Space

Let $n_1(n_2)$ be the number of customers waiting for service including those being serviced by servers of Group 1 (2). There are 7 state sets in the system, let S_i , ($i=1,2,\dots,7$) denote the specific state set, define S_1 is the state set that the agents in Group 1 are the idle state ($n_1 < N_1$), and the agents in Group 2 are the idle state ($n_2 < N_2$) too. S_2 is the agents in Group 1 are the idle state ($n_1 < N_1$), but the agents in Group 2 are the just full state ($n_2 = N_2$). S_3 is the agents in Group 1 are the idle state ($n_1 < N_1$), but the agents in Group 2 are the busy state ($n_2 > N_2$). S_4 is the agents in Group 1 are the just full state ($n_1 = N_1$), but the agents in Group 2 are the idle state ($n_2 < N_2$). S_5 is the agents in Group 1 are the just full state ($n_1 = N_1$), and the agents in Group 2 are the just full state ($n_2 = N_2$) too. S_6 is the agents in Group 1 are the just full state ($n_1 = N_1$), but the agents in Group 2 are the busy state ($n_2 > N_2$). S_7 is the agents in Group 1 are the busy state ($n_1 > N_1$), and the agents in Group 2 are the busy state ($n_2 > N_2$) too. The state transition diagram is shown in Figure 2.

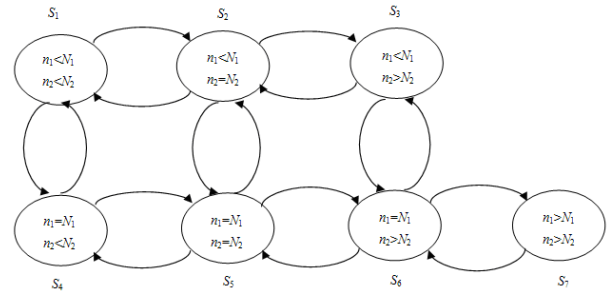


Figure 2. The state sets transition diagram

B. The Calculation of the State Set Transition Rate

The states transition occurs due to the arrival of a call or the completion of a service.

- (1) The change of states due to arrivals of calls: Consider the state S_1 , The diagram of the transition from state S_1 to state S_4 is shown in Fig. 3.

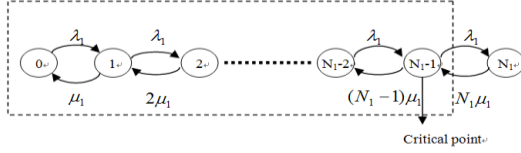


Figure 3. Diagram of the state-transition from state S_1 to state S_4

The trigger for the transfer from state S_1 to state S_4 is a call of Type 1. If a call of Type 1 arrives at the epoch when $n_1 \leq N_1 - 2$, the state S_1 will not be changed. If a call of Type 1 arrives at the epoch when $n_1 = N_1 - 1$, then the transition will occur from S_1 to S_4 . Thus, we can obtain the transition rate q_{1-4} from state S_1 to state S_4 as follows.

$$q_{1-4} = \lim_{\Delta t \rightarrow 0} \frac{P_{S_1, S_4}(\Delta t)}{\Delta t} = \lambda_1 P(n_1 = N_1 - 1) \quad (1)$$

where $P_{S_1, S_4}(\Delta t)$ is the probability that the state process is in state S_4 at time $t + \Delta t$ provided that the process is in state S_1 at time t , and $P(n_1 = N_1 - 1)$ is the probability that there are $N_1 - 1$ calls of Type 1 needing to be serviced by the agents in Group 1.

Note that if the process is in the state S_1 then the number of calls of either Type 1 or Type 2 is less than the number of the agents either in Group 1 or Group 2, i.e., $n_1 < N_1$ and $n_2 < N_2$. In this case, each queue behaves like an M/M/c/c loss queuing system. Thus, using the results of the M/M/c/c loss queuing system, we have

$$P(n_1 = N_1 - 1) = \frac{1}{(N_1 - 1)!} \left(\frac{\lambda_1}{\mu_1} \right)^{N_1 - 1} \bigg/ \sum_{j=0}^{N_1} \frac{\left(\frac{\lambda_1}{\mu_1} \right)^j}{j!}. \quad (2)$$

Similar analysis gives the other transition rates q_{i-j} caused by the arrival of calls which are given as follows:

$$\begin{aligned} q_{1-2} &= P(n_2 = N_2 - 1) \times \lambda_2, \quad q_{2-3} = \lambda_2, \quad q_{6-7} = \lambda_1, \quad q_{5-6} = \lambda_1 + \lambda_2, \\ q_{1-4} = q_{2-5} = q_{3-6} &= P(n_1 = N_1 - 1) \times \lambda_1, \quad q_{4-5} = P(n_2 = N_2 - 1) \times (\lambda_1 + \lambda_2), \\ \text{where} \end{aligned}$$

$$P(n_2 = N_2 - 1) = \frac{1}{(N_2 - 1)!} \left(\frac{\lambda_2}{\mu_2} \right)^{N_2 - 1} \bigg/ \sum_{j=0}^{N_2} \frac{\left(\frac{\lambda_2}{\mu_2} \right)^j}{j!}. \quad (3)$$

- (2) The state-transition due to the completion of a service of a call: Consider the state S_2 , The diagram of the transition from state S_2 to state S_1 is shown in Fig. 4.

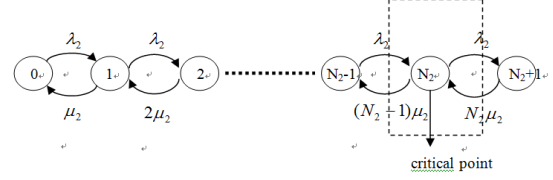


Figure 4. Diagram of the state-transition from state S_2 to state S_1

The trigger of the transfer from state S_2 to state S_1 is due to a service completion of a call of Type 2. If the state process is in state S_2 , then all N_2 agents are busy. Thus, the transition rate from state S_2 to S_1 is $q_{2-1} = N_2 \mu_2$. Similar analysis gives the other transition rates q_{i-j} caused by the service completion are given as follows

$$\begin{aligned} q_{2-1} &= q_{5-4} = N_2 \mu_2, \quad q_{4-1} = q_{5-2} = q_{6-3} = N_1 \\ q_{7-6} &= P(n_1 = N_1 + 1) \times N_1 \mu_1, \quad q_{3-2} = q_{6-5} = \\ &P(n_2 = N_2 + 1) \times N_2 \mu_2 \end{aligned}$$

where the probabilities of $P(n_1 = N_1 + 1)$ and $P(n_2 = N_2 + 1)$ can be obtained by using the results of the M/M/c queuing system which are given as follows:

$$P(n_1 = N_1 + 1) = \frac{1}{N_1 \cdot N_1!} \left(\frac{\lambda_1}{\mu_1} \right)^{N_1 + 1} P_0^1, \quad (4)$$

$$P_0^1 = \left[\sum_{j=0}^{N_1 - 1} \frac{\left(\frac{\lambda_1}{\mu_1} \right)^j}{j!} + \frac{N_1 \left(\frac{\lambda_1}{\mu_1} \right)^{N_1}}{N_1! \left(N_1 - \frac{\lambda_1}{\mu_1} \right)} \right]^{-1}, \quad (5)$$

The probability $P(n_2 = N_2 + 1)$ is similar to $P(n_1 = N_1 + 1)$, it is the probability that there are $N_2 + 1$ customers in the M/M/ N_2 queuing system with arrival rate and service rate .

C. The Equations for the Steady-state Probabilities

By the former analysis we can get 16 state transition rates, thus we can obtain the equilibrium equations of system, as follows:

$$\begin{aligned} P(S_1)(q_{1-2} + q_{1-4}) &= P(S_2)q_{2-1} + P(S_4)q_{4-1}, \\ P(S_2)(q_{2-1} + q_{2-3}) &= P(S_1)q_{1-2} + P(S_3)q_{3-2}, \\ P(S_3)(q_{3-2} + q_{3-6}) &= P(S_2)q_{2-3} + P(S_6)q_{6-3}, \\ P(S_4)(q_{4-1} + q_{4-5}) &= P(S_1)q_{1-4} + P(S_5)q_{5-4}, \\ P(S_5)(q_{5-2} + q_{5-4} + q_{5-6}) &= P(S_2)q_{2-5} + P(S_4)q_{4-5} + P(S_6)q_{6-5}, \\ P(S_6)(q_{6-3} + q_{6-5} + q_{6-7}) &= P(S_3)q_{3-6} + P(S_5)q_{5-6} + P(S_7)q_{7-6}, \\ P(S_7)q_{7-6} &= P(S_6)q_{6-7}, \\ \sum_{i=1}^7 P(S_i) &= 1. \end{aligned}$$

From the above equations we can obtain all of the steady-state probabilities $P(S_i)$, $i=1,2,\dots,7$. Because the

equations are linear, using MATLAB software, the calculating speed can be guaranteed.

IV. THE STAFFING PROBLEM

In this section, we first consider the calculation of the service level. Then, we consider the staffing problem.

A. The Calculation of the Service Level

The service level is defined as the percentage of the serviced calls in a given fixed waiting time, denoted by P/T. Actually, the 80/20 principle is a general rule, that is to say at least 80 percent of the calls should be serviced within a 20 second waiting time. We can derive the service level using the steady-state probabilities.

Let P_{sl}^1 and P_{sl}^2 be the probabilities that the call of Type 1 and Type 2 is serviced in a fixed time T_1 and T_2 , respectively.

Consider a call of Type 1, here we assume that the service level of call 1 is defined as the probability of the calls not be serviced in T_1 time, call 1 has a queue only happen in state sets S_7 . Due to call 1 is important than call 2, if a server in group 2 completed a service, he will choose call 1 in the queue to serve immediately, so in the state set S_7 , the service rate of call 1 is $N_1 \mu_1$, thus we can get, in T_1 time, the number of the calls could be served is an integer not greater than $T_1 \times (N_1 \mu_1 + N_2 \mu_2)$. By the above analysis we can get the service level probability of call 1 in T_1 time is:

$$P_{sl}^1 = 1 - P(S_7) \times P(n_1 \geq N_1 + N_2 + [T_1 \times (N_1 \mu_1 + N_2 \mu_2)]) \quad (6)$$

Where

$$P(n_1 \geq N_1 + N_2 + [T_1 \times (N_1 \mu_1 + N_2 \mu_2)]) = \sum_{i=N_1+N_2+[T_1 \times (N_1 \mu_1 + N_2 \mu_2)]}^{\infty} P(n_1 = i) \quad (7)$$

$$P(n_1 = i) = \frac{1}{N_1^{i-N_1} \cdot N_1!} \left(\frac{\lambda_1}{\mu_1} \right)^i P_0 \quad (8)$$

$$P_0 = \left[\sum_{j=0}^{N_1-1} \frac{\left(\frac{\lambda_1}{\mu_1} \right)^j}{j!} + \frac{N_1 \left(\frac{\lambda_1}{\mu_1} \right)^{N_1}}{N_1! \left(N_1 - \frac{\lambda_1}{\mu_1} \right)} \right]^{-1} \quad (9)$$

For a call of Type 2, it is more complicated than call 1, but the analysis method is the same, call 2 have a queue

happens in the state set S_3 , S_6 and S_7 , therefore the service level probability of call 2 in T_2 time is:

The probability calculation is similar to call 1.

TABLE 1. THE NUMERICAL RESULTS OF THE SERVICE LEVELS P_{sl}^1 AND P_{sl}^2 .

N	N_1	N_2	P_{sl}^1	P_{sl}^2
37	12	25	0.9994	0.7348
36	11	25	0.9830	0.6402
41	11	30	0.9910	0.6643
46	11	35	0.9972	0.7032
50	15	35	0.9992	0.8326
50	25	25	1.0000	0.8419
50	20	30	1.0000	0.8704
135	65	70	1.0000	0.7352
161	61	100	0.9461	0.4922
170	80	90	1.0000	0.8236

The actual call center generally classified according to the number of the agents. Our calculation formula of the service level is suitable for various scale of call center. Because this paper adopted the division of state set method, there are 7 state set, and 7 steady-state probabilities, the solution of the steady-state probability has nothing to do with the agents number N , So theoretically, we can solve any scale call center. There are some numerical results of the service levels P_{sl}^1 and P_{sl}^2 for different scale call centers in table 1. In multi-skill call center, the scale is the total number of two agent groups. Let N denote the total number of the agents.

From the table we can observe that, the service level calculation formula can calculate any scale of call center, and at the same time, for multi-skill call center, the total number of agents is the same, but every agent group number is different, the service level of call center also has a great influence. So we have to find the optimal numbers of every agent group, to achieve the service level, and at the same time the cost of the system is minimal.

B. The Calculation of the Staffing Problem

Let C_1 and C_2 be the costs for each server in Group 1 and Group 2, respectively. In order to minimize the cost, we try to find the optimal number of servers N_1 and N_2 subject to the constraint conditions. The optimization of a staffing problem can be expressed as follows:

$$\begin{aligned} \min \quad & Z = C_1 N_1 + C_2 N_2 \\ \text{s.t.} \quad & P_{sl}^1 \geq \alpha_1 \\ & P_{sl}^2 \geq \alpha_2 \\ & a_i \leq N_i \leq b_i \\ & N_i, a_i, b_i \in \mathbb{Z}^+, i = 1, 2 \end{aligned} \quad (11)$$

$$P_{sl}^2 = 1 - P(S_3) \times P(n_2 \geq N_2 + [T_2 \times N_2 \mu_2]) - P(S_6) \times P(n_2 \geq N_2 + [T_2 \times N_2 \mu_2]) - P(S_7) \quad (10)$$

where μ_1 and μ_2 are the given service rates of the call of Type 1 and the call of Type 2, Z^+ denote a set of positive integers.

This is a nonlinear integer programming problem, the objective function is a linear function, there are two variables N_1 and N_2 need to be sure. Constraints are inbound calls to satisfy a certain service level, that is to say, the service level of call 1 is greater than or equal to α , the service level of call 2 is greater than or equal to β . The calculate formula of service level see above (6) and (10), from the formula we can see, the constraint is highly nonlinear, Since there are only two variables, we can solve this problem very well by the traditional method. We use the implicit enumeration method, by MATLAB software, to find the optimal numbers of every agent group.

The implicit enumeration method was proposed by Balas E in 1965, it is compared with the exhaustive method (complete enumeration method). Its basic characteristic is only need to investigate part of the variable combinations of the problem. The key of using the implicit enumeration method to solve the staffing problem is to reduce the number of enumeration. According to this model, in order to reduce the number of enumeration, first we use the constraint conditions that the service level of call 1 and call 2 are greater than a certain value to eliminate some variable combinations. When computing, the variable combinations test in turn with these constraints. If constraints did not pass, get rid of this set of variables. Then from the satisfy combinations to find out the combination of the minimum cost, then the combination is the optimal number of agents we want to find.

The specific algorithm is described below:

- (1) Initialize the parameters,
- (2) Give all variable combinations of N_1 and N_2 ,
- (3) Use constraints to get rid of the combinations that does not satisfy the constraints,
- (4) Then from the satisfy combinations to find out the minimum cost.

In order to verify the above algorithm's effectiveness for solve the staff problem, There are some numerical results in the table 2, the table below shows the optimal number of agents in the cases of different agent group costs. The parameter settings are as follows:

$$T_1=20, \quad T_2=30,$$

TABLE 2. THE NUMERICAL RESULTS OF THE OPTIMAL N_1 AND N_2 .

C_1	C_2	N_1	N_2	N	cost
2	2	14	31	45	90
5	1	13	36	49	101
5	2	14	31	45	132
10	1	12	46	58	166
10	2	13	36	49	202
10	3	14	31	45	233
2	5	15	30	45	180
2	6	15	30	45	210
2	10	15	30	45	330

We can see from the table 2, the costs C_1 and C_2 influence the number of agents in each group, in order to meet the requirements of the service level, the costs C_1 and C_2 different, the number of agents in each group also different, and the agent group 2 needs more agents.

TABLE 3. THE NUMERICAL RESULTS OF THE OPTIMAL N_1 AND N_2 .

time					N_1	N_2	N	cost
07:00~07:30	26	0.8	48	2.5	39	23	62	209
07:30~08:00	22	0.7	67	2.2	35	34	69	241
08:00~08:30	41	0.8	87	2.8	58	35	93	314
08:30~09:00	58	0.6	83	2.3	108	44	152	500
09:00~09:30	38	0.5	75	2.3	85	37	122	403
09:30~10:00	45	0.5	85	2.6	99	38	137	449
10:00~10:30	39	0.6	76	2.4	72	37	109	364
10:30~11:00	44	0.7	64	1.6	74	47	121	410
11:00~11:30	24	0.5	52	2.5	55	24	79	261
11:30~12:00	23	0.5	52	2.2	52	27	79	264
12:00~12:30	28	0.5	41	2.4	63	23	86	281
12:30~13:00	20	0.7	84	1.2	32	74	106	392

There are some data of a certain logistics company call center in the table 3. In the logistics company, the agent group 1 serves the new customers, μ_1 and μ_2 are the arrival rate and the service rate of the new customers respectively. The agent group 2 serves both the registered customers and the new customers, μ_1 representative the arrival rate of registered customers, and μ_2 representative the service rate of the agent group 2. We selected the data of 12 time intervals to calculate the optimal agent numbers. Here we assume the cost of the two agent groups are C_1 and C_2 , and $T_1=20$, $T_2=30$, that is call 1 should be serviced within a 20 second waiting time, call 2 should be serviced within a 30 second waiting time.

The above numerical results is obtained in the cases that the costs of the two agent groups are not much difference, using our method, we can obtain the logistics call center's optimal demand for every agent group.

V. CONCLUSIONS

This paper has studied a queuing model of the N-design multi-skill call center. The system has two types of customers and two server groups. We divided the system's infinite number of system states into several finite state sets, obtained the transition rates of the state sets, the equilibrium equations and the steady-state probabilities of the system. We also obtained the computational formula of the service level and presented an optimization model of the staffing problem to calculate the optimal number of the agents in each groups. In future it will be useful to extend the analysis of our model to the Non-exponential models with general service times, impatient times and so on.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China (No. 71071133).

REFERENCES

- [1] Koole G., Mandelbaum A. Queueing models of call centers: an introduction. *Annals of Operations Research*, 113: 41-59 (2002).
- [2] Gans N, Koole G, Mandelbaum A. Telephone call centers: Tutorial, review, and research prospects. *Manufacturing & Service Operation Management*, 5(2): 79- 141 (2003)
- [3] Aksin, Z., Armony, M., Mehrotra, V. The modern call-center: a multi-disciplinary perspective on operations management research. *Production and Operations Management*, 16: 655-688(2007)
- [4] Perry, M., Nilsson A. Performance modeling of automatic call distributors: Assignable grade of service staffing. *Proceedings of the 14th International Switching Symposium*, 294-298 (1992)
- [5] Bhulai, S., Koole G. A queueing model for call blending in call centers. *IEEE Transactions on Automatic Control*, 48:1434-1438 (2003)
- [6] Gans N., Zhou Y. A call-routing problem with service-level constraints. *Operations Research*, 51(2): 255-271 (2003)
- [7] Armony, M, Maglaras C. On customer contact centers with a call-back option: Customer decisions, routing rules, and system design. *Operations Research*, 52(2):271-292 (2004)
- [8] Armony, M., Maglaras C. Contact centers with a call-back option and real-time delay information. *Operations Research*, 52 (4): 527-545 (2004)
- [9] Ormeci E., Burnetas A, Emmons H. Dynamic policies of admission to a two class system based on customer offers. *IEE Transactions*, 34:813-822 (2002)
- [10] Pekoz E., Optimal policies for multi-server non-preemptive priority queues. *Queueing Systems: Theory and Applications*, 42:91-101 (2002)
- [11] Stanford D, Grassman W. Bilingual server call centers. *Analysis of Communication Networks: Call centers, Traffic and Performance*, 2000
- [12] Shumsky R. Approximation and analysis of a queueing system with flexible and specialized servers. *OR Spektrum*, 26:307-330, (2003)