Advances in Economics, Business and Management Research, volume 159
Fifth International Conference on Economic and Business Management (FEBM 2020)

Study on the Workforce Scheduling and Routing Strategies of Heterogeneous Agents in Call Centers

Mengchen Wang, Xiuli Wang*

School of Economics and Management, Nanjing University of Science and Technology, Nanjing, Jiangsu 210000, China
**Corresponding author. Email: wangdu0816@163.com*

ABSTRACT
The performance of call center operations is usually associated with two factors, agent scheduling shifts and calls routing strategy, and they are dependent to some extent. In order to decrease the operating cost greatly, call centers need to properly allocate human resources and determine appropriate routing strategies. Since agent heterogeneity exists in the call center service system, this paper no longer assumes that agent service rates are all the same. We construct an integer linear programming of the scheduling problem for call centers with agent heterogeneity, and combine the use of a discrete-event simulation model with an artificial bee colony algorithm to solve the model. We evaluate the performance of the call center through discrete-event simulation experiments, and the artificial bee colony algorithm takes this performance metric as a judgment condition for satisfying the constraints, and iteratively obtains a better scheduling solution. Finally, we compare the effects of three routing strategies on the control of labor costs in enterprise scheduling. The actual data of the call center proves that considering the scheduling scheme of heterogeneous agents and the fastest server first routing can significantly reduce labor cost, and achieve effective service management.
Keywords: *Agent heterogeneity, workforce scheduling, routing strategies, simulation, artificial bee colony algorithm*

1. INTRODUCTION

A call center is an organization in which agents provide services to business customers via telephone, and is widely used in various service and manufacturing industries [1]. The call center is a labor-intensive industry, with over 70% of costs spent on human resources. How to effectively utilize human resources is critical to call centers [2].

Standard queuing models are often used to analyze call center performance, and these models assume that there is no difference in service rates for each agent within a group ("i.e., agent homogeneity") [3]. In fact, this assumption cannot be justified because agents have different experience and skills, and each agent handles customer requests at different speeds ("i.e., agent heterogeneity") [4].

In recent years, some researchers have studied the call center scheduling problem. Excoffier et al. [5] studied that shift scheduling problem under the uncertainty of call arrival rates prediction and constructed a large mixed linear planning model to solve the problem. Hu et al. [6] studied the scheduling problem in call centers under the consideration of shift type constraints, and designed heuristic algorithms to solve the large-scale scheduling problem. Some scholars have further studied call center with agent heterogeneity. Thomas [7] studied the causes of agent heterogeneity, proposed a routing strategy based on agent experience, and exploited the heterogeneity to improve the long-term performance of the call center. Current studies mostly focus on a certain part of the scheduling problem and there are few studies that combine the scheduling problem with the routing strategy. Su et al. [8] proposed a dynamic ensemble coverage algorithm to solve the multi-skilled scheduling routing problem, which could quickly assess the scheduling performance. However, their study does not take into account agent heterogeneity, and there is a difference with the actual situation.

Scheduling and routing problems are dependent at some level. Agents scheduling plans restrict routing decisions, which in turn has an impact on the scheduling scheme. The scheduling problem optimization schemes can obtain the set of agents on duty, and the routing strategy arranges service calls for a given set of agents. In this paper, we study the actual scenarios of scheduling and routing in the call center, and conduct research on heterogeneous agent scheduling optimization models and routing strategies.

Section 2 defines the problem and provides its integer linear programming *(ILP)* formulation. Section 3 describes the artificial bee colony algorithm to solve the *ILP* formulation. Section 4 presents computational results. Section 5 analyzes and compares the performance of three routing strategies and the paper is concluded in Section 6.

2. PROBLEM FORMULATION AND MODELING

2.1. Problem Statement

In this paper, we take a single day as the planning horizon, the day is divided into periods (e.g., 30 minutes or one hour each), and determine the shifts of heterogeneous agents under the constraints of satisfying the specified service level, so that the labor cost is minimal. We further compare the impact of each routing policy on the labor cost of scheduling to find the optimal routing strategy.
A call center is an example of a queuing system with an important performance metric is the Telephone Service Factor *(TSF)*, which is usually used as the percentage of inbound calls answered within the given time [9]. Given the TSF value, the Erlang-A model is usually used to calculate the labor requirement for the interval, but this method ignores the phenomenon of different agent service rates. Therefore, we will construct a simulation model of a call center with agent heterogeneity, and calculate the service level and labor demands of the call center through discrete-event simulation experiments, which does not require much approximation to the actual data and brings it closer to reality.

2.2. Formulation

Define the following parameters and decision variable notations:
Sets
N : agents.
I : periods.
J : possible schedules.
μ : service rates.
State variables
C_n : daily cost of agent *n*.
α : specified service level.
a_{ij} : equals 1 if period *i* is included in shift *j* and otherwise, equals 0.
u_i : labor demands for time period *i*.
E^i : the unit matrix of period *i* $(n*n)$.
S^i : service rate set for on-duty staffs in period *i*, $S^i = E^i * \mu$.
K_j : number of staffs on duty on shift *j*.
Y^i : number of staffs on duty in time period *i*.
$T^i(S^i,Y^i)$: based on the staffing level Y and the corresponding agents service rate set S, the service level is calculated by discrete-event simulation in period *i*.
Decision variables
x_n : equals 1 if agent *n* is on duty and otherwise, equals 0.
E^i_{nn} : equals 1 if agent *n* is working in time period *i* and otherwise, equals 0.
y_{nj} : equals 1 if agent *n* is on duty on shift *j* and otherwise, equals 0.
An integer linear program model for the call center agent scheduling optimization problem is constructed as follows.
Integer Linear Program (ILP)

$$Z = \min \sum_{n=1}^{N} x_n C_n \quad (1)$$

s.t.

$$\sum_{j=1}^{J} y_{nj} \le 1 \quad \forall n \in N \quad (2)$$

$$\sum_{j=1}^{J} y_{nj} = x_n \quad \forall n \in N \quad (3)$$

$$\sum_{n=1}^{N} E^i_{nn} = Y^i \quad \forall i \in I \quad (4)$$

$$\sum_{j=1}^{J} y_{nj} a_{ij} = E^i_{nn} \quad \forall n \in N, i \in I \quad (5)$$

$$\sum_{j=1}^{J} a_{ij} K_j \ge u_i \quad \forall i \in I \quad (6)$$

$$T^i(S^i, Y^i) \ge \alpha \quad \forall n \in N \quad (7)$$

$$x_n, y_{nj} \in \{0,1\}, Y_j, Y_i \in \mathbf{Z}^+ \cup \{0\} \quad (8)$$

The objective function is to minimize the total labor cost. Constraint (2) indicates maximum of one shift each agent work per day. Constraint (3), (4) and (5) describe the relationship between the variables, respectively. Constraint (6) represents a limit on the number of agents in an interval required to meet the target service level. Constraint (7) indicates the level of service attained by the on-duty agents answering calls in each interval to satisfy the target service level constraint. Constraint (8) defines the non-negativity and integer conditions for program variables.

2.3. Simulation Model

This paper establishes a discrete event simulation optimization model for a call center based on the following assumptions:
①Assuming that the arrival rate of calls is a stable Poisson distribution, denoted by λ_t .
② Assuming that service times are independent and exponential, denoted by $1/\mu_n$. Given the heterogeneity of the agents, μ_n is not identical.
③ Assuming that arriving calls are first routed to the available agent with longest idle time.

④ Assuming that callers have finite patience, which is exponentially distributed, denoted as β . If wait time exceeds patience time, the caller will abandon. We use MATLAB to construct a discrete-event simulation model to calculate the service level of the call center. Each simulation result is based on the mean of 1000 simulations repeated for a sufficiently long period of time, and the confidence interval of the simulation results is based on a 95% confidence level.

In the simulation optimization, the solution obtained by the optimization algorithm is used as an input for the simulation model, and the performance metric of the simulation output are used to guide the optimization algorithm to find a new solution, and the new solution generated by the optimization algorithm is used as an input for the next simulation. The process is repeated until the termination conditions are met. The implementation process is shown in Figure 1.

3. CALL CENTER SCHEDULING OPTIMIZATION ALGORITHM

Considering the differences in the service rate of each agent in a call center, for better arrangement and management, we propose to schedule the agents in a group as the smallest unit and use an artificial bee colony algorithm to achieve the scheduling solution. Artificial Bee Colony *(ABC)* is a swarm intelligence random search algorithm proposed to find the optimal food resource by exchanging information between three types of bees to find their behaviour.

3.1. Design of Artificial Bee Colony Algorithm

3.1.1. Coding

This paper uses integer coding, in Table 1 genes “1-7” represent shift numbers. And “0” means that the agent does not work that day.

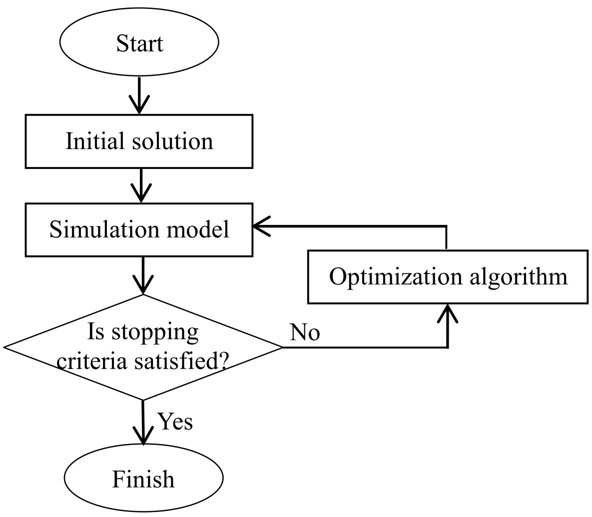

Figure 1 The flowchart of optimization algorithm

3.1.2. Food source initialization

During the initialization phase, determine the number of populations, the number of bees, and the values of the control parameters. Assuming the food source dimension is *n*, *n* in this paper represents the number of agents. The food source is generated as follows.

$$X_{ij} = Lb_j + (Ub_j - Lb_j) * rand(0,1) \quad (9)$$

Subject to satisfying the constraints, X_i is a source, a solution to the scheduling problem.

3.1.3. Employed bees phase

Each employed bee generates a new solution X_{newi} by searching the neighbourhood for a known solution X_i , as follows.

$$X_{new\,i}(j) = X_{ij} + (X_{ij} - X_{kj}) * \varepsilon \quad (10)$$

Where ε is the random number in the interval [-1,1].Calculate the adaptation of new food source and compare it with the old one, retain better adapted food sources based on the greedy principle.

3.1.4. Observed bees phase

The observed bee selects a food resource according to the probability calculated from equation (11) and produces a new food resource according to equation (10).

$$P_i = \frac{f_i}{\sum_{n \in N} f_n} \quad (11)$$

Calculate the adaptation of the new resource and update food resources according to the greedy principle.

3.1.5. Detective bees phase

If a food resource is still unable to find a better solution after the number of searches in the vicinity of a food resource reaches the limit times, the food resource is abandoned. The associated employed bees become detective bees and generate a new food resource according to equation (9).

The above process is repeated until the termination condition is met. The information contained in the food resource at this point in the decoding is the better scheduling solution.

Table 1 Example of gene sequences

No. of agents	1	2	3	…	16	17
Gene	0	1	5	…	3	0

4. COMPUTATIONAL EXPERIMENTS

The call center divides agents into several service groups, with varying numbers of agents in each group. Table 2 shows a list of call center shift-types and shifts.

4.1. Instance Calculation

We used an example to illustrate the model and algorithm specific calculation process and results. The call center operates from 7 a.m. to 10 p.m. and covers 16 periods (length 1 hour).The distribution of inbound calls for the day periods is shown in Figure 2.

The service level indicator TSF specifies that 85% of calls must be answered within 30 seconds. Based on the actual data of this service group, we use the *R* to fit the analysis, and the customer's patience time β takes a value of 2 minutes and the agent service rates parameter μ_i takes a value in the range of 0.3-0.5. There are 17 agents in this group, and their service rates are: 0.50,0.48,0.46,0.47,0.45,0.42,0.43,0.4,0.44,0.35,0.37,0.36,0.38,0.365,0.3,0.32 and 0.31, respectively.

Classifying the service rates of agents by level, the daily cost of agents is a function of their service level. The hourly cost of a staff member with the lowest service rate is standardized to 100, and the cost of a staff member with service capability level m is $1+(m-1)*p$, where $p=0.15$, as shown in Table 3.

Table 2 List of call center shifts and types

Shift Types	Shifts	Working Hours	Meal Times
Morning Shift	A1(01)	7:00-16:00	11:00-12:00
	A2(02)	8:00-17:00	11:00-12:00
	A3(03)	9:00-18:00	12:00-13:00
	A4(04)	10:00-19:00	12:00-13:00
Evening Shift	B1(05)	11:00-20:00	16:00-17:00
	B2(06)	12:00-21:00	16:00-17:00
	B3(07)	13:00-22:00	17:00-18:00

Table 3 List of daily cost of agents

Service rate	Capability level (*m*)	Cost(per day)
$0.3\le\mu<0.35$	1	100
$0.35\le\mu<0.4$	2	115
$0.4\le\mu<0.45$	3	130
$\mu\ge 0.45$	4	145

Mis-application of traditional staffing procedure to call centers with agent heterogeneity can lead to problems with staffing and scheduling schemes. Therefore, we combine the consecutive binary search method with discrete-event simulation model to calculate the demand of manpower in a period. The calculated distribution of the minimum manpower requirements is shown in Figure 2.

This paper is tested on a computer with an Intel Core^{TM} i3 CPU and 4GB of RAM. We use MATLAB 2019b to program the ABC algorithm to solve the problem. The computational results are shown in Table 4.

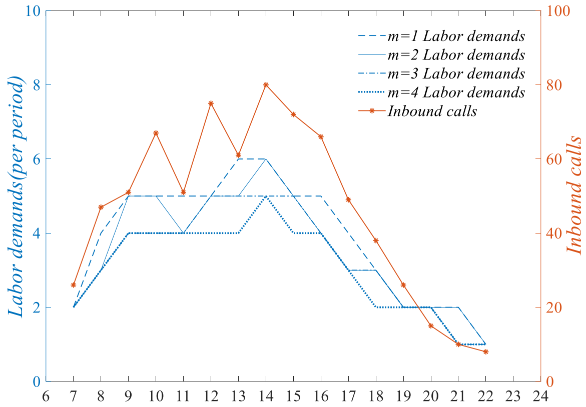

Figure 2 Inbound calls of periods and calculated labor demands one day

4.2. Experimental Analysis

Next, we analyze the impact of the agent homogeneity assumption on scheduling labor cost control. The group remain the same, with a total labor cost of $2,120 and an average daily cost of $125 per agent. Assuming that the number of people on duty under homogeneous agents is the same as under heterogeneous agents and each period is required to meet the specified service level, the average service rate $\bar{\mu}$=0.385 for homogeneous agents is calculated by discrete-event simulation experiments. Since the agents are homogenous, we use Erlang-A model to calculate the manpower requirement for the periods and this can be used as a constraint to meet the TSF requirement. As a result, the scheduling scheme is shown in Table 5.

The scheduling schemes under both the homogeneous and heterogeneous agents' scenarios require 8 people to satisfy the service level constraint, but the total cost of scheduling is $1,000 for the former and only $935 for the latter. The result shows that considering the heterogeneous scheduling of agents can significantly reduce costs, and the call center can maintain the specified service level, which can achieve effective service management.

There are peak and trough periods in the call center during the day. The above service group is divided into two groups with an average service rate, the faster service group is , and the slower group is . According to the above scheduling

scheme, the skilled agents in N_1 are scheduled for peak call hours and the unskilled agents are scheduled for trough call hours, resulting in 1000 scheduling schemes to calculate service level values.
Defined service level deviation = (calculated service level by simulation-target service level) $\div$ target service level $*100\%$.
The calculation results are shown in Figure 3.
Figure 3 indicates that arranging skilled agents in the call center during peak hours a substantial amount of excess capacity may be created in other periods due to shift constraints and there is a problem of overstaff. On the other hand, arranging unskilled staffs in the call center of time during trough periods, there is a problem of understaff in the call center and poor performance.
Therefore, arranging the agents with high and low service rate to work in one shift can effectively utilize human resources, minimize labor costs, and ensure that each period can meet the specified service level.

5. ROUTING STRATEGY ANALYSIS

The routing problem is a control problem that involves assigning incoming calls to specific agents, thus, changes in routing strategies make scheduling schemes vary, the performance of a routing strategy has a significant impact on operating costs and service levels. Next, we use MATLAB to construct call center simulation models with different routing strategies to analyze and compare the advantages and disadvantages of each routing strategy from the perspective of labor cost control.

5.1. Routing Strategy

In this section we evaluate three routing strategies; namely:
①Longest Idle Routing (LIR):find the agent who has been idle for the longest time. This is a standard routing scheme.
②Fastest Server First (FSF): find the available agent with the highest service rate to handle the call.
③ Experienced Based Routing (EBR): find the least experienced agent when the call center is in low volume mode and find the most experienced agent when the call center is in high volume mode.

5.2. Experimental Results and Analysis

We evaluate the impact of three routing strategies on scheduling labor cost through several examples of different scales. We use MATLAB to construct simulation models for call centers under different routing strategies. The comparison results are shown in Table 6.
Table 6 reveals that (i) When the size of the call center is small, the difference between the scheduling numbers and labor costs under the three routing strategies is small, and the difference is relatively obvious when the call center gradually increases in size. (ii) When the group size and the incoming call arrival rate are the same, the fastest server first *(FSF)* routing policy can reduce the scheduling labor cost.

Table 4 Experimental results

Agents	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17
Service rates	0.50	0.48	0.46	0.47	0.45	0.42	0.43	0.4	0.44	0.35	0.37	0.36	0.36	0.38	0.3	0.32	0.31
Cost	145	145	145	145	145	130	130	130	130	115	115	115	115	115	100	100	100
Shifts	0	0	0	0	2	0	7	0	5	0	1	0	3	0	5	1	3
Total staffs									8								
Total costs									935								

Table 5 Results of homogeneous agents schedule

Shifts	A1	A2	A3	A4	B1	B2	B3
Scheduling staffs	2	1	2	0	2	0	1
Total staffs				8			
Total costs				1000			

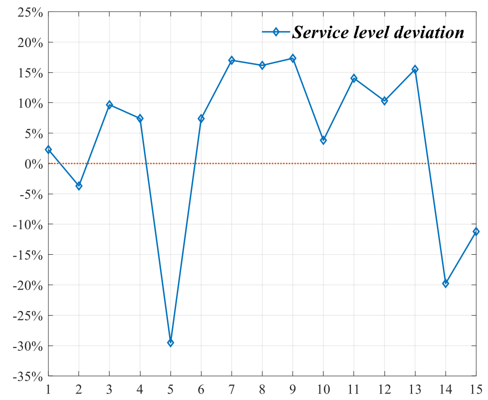

Figure 3 Service level deviation of one day

Table 6 Experimental results of different size problem examples

Examples	Group size	LIR		FSF		EBR	
		Avg.cost	Staffs	Avg.cost	Staffs	Avg.cost	Staffs
1	17	950	8	920	8	940	8
2	20	1071	9	1047	9	1068	9
3	27	1402	13	1374	12	1392	12
4	33	1988	17	1920	16	1960	16
5	40	3120	26	2832	24	3028	26
6	52	4248	36	3794	32	3980	34

6. CONCLUSION

This paper proposes a scheduling optimization model focusing on agent heterogeneity; we associate the discrete-event simulation model with the ABC algorithm to solve the scheduling problem. Next, we build models for three routing strategies respectively and analyze the impact of three routing strategies on the labor cost of scheduling. We consider the fluctuation of the service time of each agent, so constructing a discrete-event simulation model can maintain the stability of the system, and the proposed ABC algorithm can also solve the scheduling problem with the agent homogeneity, which can quickly obtain the optimal service arrangement. The study shows that that a scheduling scheme considering agent heterogeneity and the fastest server first routing can reduce the labor costs, utilize human resources more efficiently, and maintain a higher level of service quality than the scenario with agent homogeneity.

ACKNOWLEDGMENT

This work was supported by the Jiangsu Province Postgraduate Research Innovation and Practice Project (KYCX20_0388).

REFERENCES

[1] Brown L, Gans N, Mandelbaum A, Sakov A, Shen H, Zeltyn S, Zhao L. Statistical analysis of a telephone call center: a queueing-science perspective, American Statistical Association, 100(469) (2005) 36-50.DOI:10.1198/016214504000001808

[2] Gans N, Koole G, Mandelbaum A. Telephone call centers: tutorial, review, and research prospects, Manufacturing & Service Operations Management, 5(5) (2003) 79-141.DOI:10.1287/msom.5.2.79.16071

[3] Koole G, Mandelbaum A. Queueing models of call centers: an introduction, Annals of Operations

Research, 113(1-4) (2002) 41-59.DOI:10.1023/A:1020949626017

[4] Gans N, Liu N, Mandelbaum A, Shen H, Ye H. Service times in call centers: agent heterogeneity andlearning with some operational consequences, Journal of Immunology, 161(11) (2010) 6198-205.DOI:10.1214/10-IMSCOLL608

[5] Excoffier M, Gicquel C, Jouini O. A joint chance-constrained programming approach for call center workforce scheduling under uncertain call arrival forecasts, Computers & Industrial Engineering, 96 (2016) 16-30.DOI:10.1016/j.cie.2016.03.013

[6] Hu XW, Wang RC, Wang XL. Model and algorithm for the shift scheduling and rostering problem of call centers with shift-type constraints, Industrial Engineering and Management, 25(02) (2020) 51-58.DOI:10.19495/j.cnki.1007-5429.2020.02.008

[7] Robbins T R. Experience-based routing in call center environments, Informs, (2015). DOI:10.1287/serv.2015.0101

[8] Su Q, Zhao F. Study on the staffing algorithm for multi-skill call center, Systems Engineering-Theory&Practice, 35(1) (2015) 239-246.DOI: 10.12011/1000-6788(2015)1-239

[9] Stefan H, Kirsten H. Profit-oriented shift scheduling of inbound contact centers with skills-based routing, impatient customers, and retrials, OR Spectrum, 32(1) (2015) 109-134.DOI:10.1007/s00291-008-0141-8