Research on staffing problem of call center based on discrete event simulation

Xiangyun Kong[a], Xiuli Wang[b,*], Yue Xu[c]

School of Economics and Management, Nanjing University of Science and Technology, Nanjing 210000, China

[a]1552491269@qq.com, [b]wangdu0816@163.com, [c]xuyue_njust@163.com

Keywords: Call center, Staffing, Discrete event simulation.

Abstract: Staffing is the basis for scheduling in call center. The commonly used Erlang A model assumes that service time follows an exponential distribution, while the actual service time of some call centers is closer to a lognormal distribution. In order to determine more reasonable staffing when service time is fitted to a lognormal distribution, in this paper, discrete event simulation is used to solve the problem of the minimum number of agents under a certain service level constraint. Then, through the case application, the minimum number of agents per hour from 8:00 to 22:00 in a day is obtained. The change trend of the minimum number of agents with the number of calls is analyzed. This is of great reference value to the staffing of the call center.

1. Introduction

Call center is an important channel for enterprises to communicate with customers and is widely used in many fields. Many enterprises have set up call centers in order to provide effective service to customers. According to the customer's psychology, the customer wants to be served quickly, but for call center, in order to quickly respond to the customer's call, it needs enough agents to serve. About 70% of operating costs is attributable to personnel costs in call center, therefore, it is particularly important to make the number of agents as small as possible under the condition of a certain service level. In addition, staffing is the basis of scheduling in call center, it is of great practical significance to allocate the number of agents as economically and efficiently as possible on the premise of meeting requirement of service level. Where service level is defined as the proportion of the number of calls that respond in 30 seconds to the total number of calls.

Call center is often modeled as queuing systems, which are used to estimate system performance in order to calculate personnel requirements to achieve expected service performance. The most common queuing model is Erlang C model [1, 2]. It is assumed that, in a steady state conditions, the arrival process is Poisson distribution, the service time is exponential distribution, customer and server is independent. It did not consider the customer's patience. Many researchers [1, 2] advocate the Erlang A model, which took customer's patience and abandonment into account on the basis of Erlang C, it’s more practical. Both Erlang C and Erlang A models assume that the service time is exponentially distributed, according to the actual data of call center, in some of call centers, the actual service time distribution is not fitting for the exponential distribution. Rouba Ibrahim et al.[3] used actual data to fit the service time distribution, not exponential distribution. Lawrence Brown et al.[4] made a statistic analysis on the data of call center, found that the distribution of service time can better fitting with lognormal distribution. Luo yimei et al. [5] collected and analyzed the service time data of a call center, by using Anderson Darling test method to compare and analyze the statistical distribution of four possible service time, it is concluded that lognormal distribution can better reveal the distribution characteristics of service time.

This paper analyzes service time data of a small and medium-sized call center and finds that the service time is fitted as a lognormal distribution. In order to solve the staffing problem when the service time distribution is lognormal distribution, a discrete event simulation method is proposed to calculate number of agents. Discrete event simulation has also been applied by many researchers [6,7]. Therefore, it is scientific to use discrete event simulation.

DOI: 10.25236/msied.2018.017

2. Problem description

Within a day, number of calls is different at different times of the day. Our data set comes from small and medium-sized call centers that provide services for electronic products. We selected a large group and counted number of calls in every hour on a given day, as shown in Fig. 1. We can see that there are two peaks of call arrival in a day, the first peak is around 12:00, and the second peak is around 15:00 in the afternoon. Due to the change of inbound calls, number of agents is also different for each period of time.

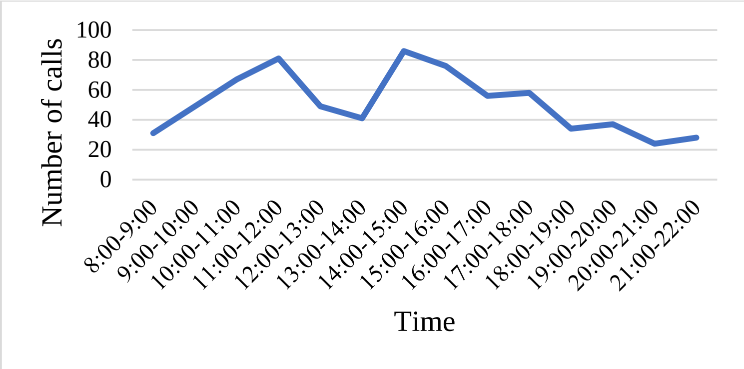

Fig. 1 Number of calls in every hour for that day

In fact, the distribution of service time is not exponential, but lognormal. We conducted K-S test for exponential distribution and lognormal distribution of service time respectively, and the P values were 0.002 and 0.2, respectively. Therefore, the service time is lognormal distribution, but not exponential distribution.

In this case, calculation formulas of various service performance indicators of Erlang A are not applicable. The simulation method can model the service time and various random factors. Therefore, under the condition of lognormal distribution of service time, this paper applies the simulation model to solve the problem of the minimum number of agents when expected service level is met.

3. Simulation model and method

3.1 Simulation model

This paper studies such a queuing model, which is described as follows: Call arrival is a Poisson distribution with parameter λ, so arrival interval is an exponential distribution with parameter λ. Customers arrive in the system and queue up for service. System has m agents, service time S is a lognormal distribution, $\ln S \sim N(\mu, \delta^2)$. Incoming calls enter the system. If at least one server is idle, it will be served immediately. Otherwise, it will enter the queue and wait for service. In the process of waiting for service, customers will leave the system due to their impatience, and customer's patience time P is an exponential distribution with parameter θ. The processes involved in the system are independent of each other and the rules of service is First come First served (FCFS). The routing rule is the first to route to the agent with the longest idle time.

3.2 Discrete event simulation

Discrete event simulation model is widely used in research of various queuing systems. The steps of this simulation experiment include: drawing the working flow chart of the system, determining the arrival, service, patience and queuing models which constitute the simulation model of discrete event system, writing a running program and execute it on a computer. Because it involves many random numbers produced by discrete time simulation, only one simulation experiment cannot accurately describe on the performance of queuing system, therefore, at the same time of input parameters, 1000 simulations are carried out, and then, we calculate the average service level for 1000 simulations, which is the real service level. The simulation time unit is seconds.

4. Case application

We selected a large group from data set and conducted statistical analysis on the data of the group from 8:00 to 22:00 on March 27, 2013. Firstly, we calculate number of calls per hour, as shown in the Table 1. Secondly, we conducted distribution fitting and parameter estimation for the service time of agents. The service time of agents follows lognormal distribution, that is, $\ln S \sim N(4.9792, 0.9484^2)$. Finally, distribution test and parameter estimation of customer's patience time are carried out. The patient time of customers is an exponential distribution with a parameter of θ=0.02135.

Table 1 Number of calls per hour

Time	Number of calls	Time	Number of calls
8:00-9:00	31	15:00-16:00	76
9:00-10:00	49	16:00-17:00	56
10:00-11:00	67	17:00-18:00	58
11:00-12:00	81	18:00-19:00	34
12:00-13:00	49	19:00-20:00	37
13:00-14:00	41	20:00-21:00	24
14:00-15:00	86	21:00-22:00	28

We use discrete event simulation to calculate service levels corresponding to different number of agents in each period. Therefore, we can obtain the minimum number of agents per period when service level is not less than 80%, as shown in the Table 2 and Fig.2.

Table 1 Number of agents and service level per hour

Time	Number of agents	Service level	Time	Number of agents	Service level
8:00-9:00	3	0.8350	15:00-16:00	6	0.8536
9:00-10:00	4	0.8115	16:00-17:00	5	0.8712
10:00-11:00	5	0.8161	17:00-18:00	5	0.8633
11:00-12:00	6	0.8272	18:00-19:00	3	0.8032
12:00-13:00	4	0.8156	19:00-20:00	4	0.8976
13:00-14:00	4	0.8741	20:00-21:00	3	0.8990
14:00-15:00	6	0.8012	21:00-22:00	3	0.8638

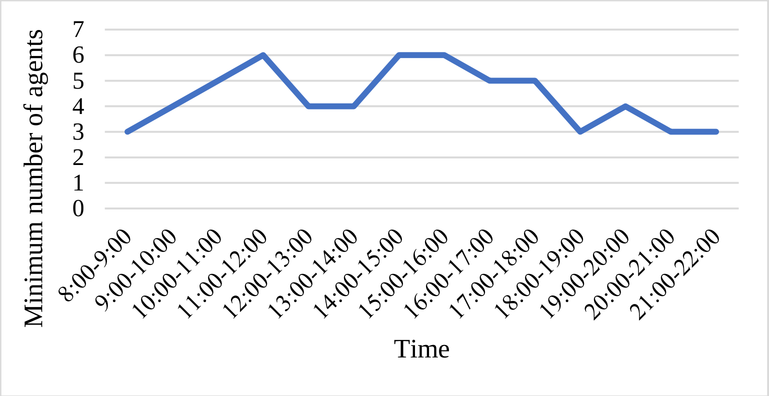

Fig. 2 Minimum number of agents per hour

As can be seen from the results, when the service level is not less than 80%, there are also two peaks in the number of agents corresponding to the two peaks in the number of calls. One peak is around 12:00 and the other peak is around 15:00, and 6 agents are needed for service.

In order to better arrange actual staffing, we did a sensitivity analysis and got the number of agents corresponding to different number of calls under the condition that service level is no less than 80%, as shown in the Table 3 and Fig. 3. As can be seen, when the number of agents is 1 and the number of calls is 6, an increase of 13 calls requires an increase of 1 agent to meet the service level. When the number of agents is 7 and the number of calls is 101, an increase of 20 calls requires an increase of 1 agent to meet the service level. Therefore, when the number of agents is small, in order to meet 80% of service level, the number of calls has a great impact on the number of agents.

Table 2 Number of agents required for different number of calls

Number of calls	Number of agents	Number of calls	Number of agents
1-5	1	85-101	7
6-18	2	102-119	8
19-34	3	120-137	9
35-49	4	138-153	10
50-66	5	154-174	11
67-84	6	175-192	12

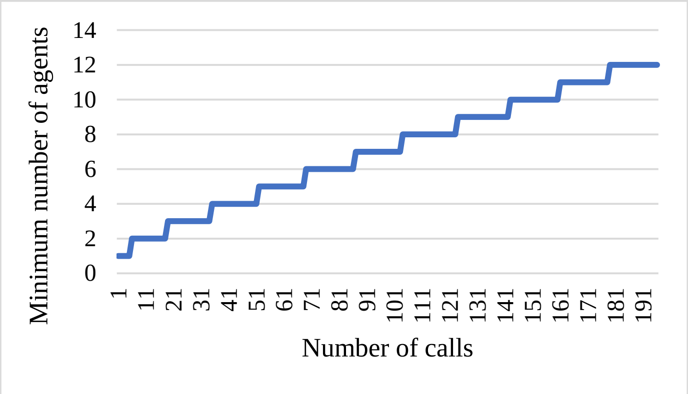

Fig. 3 The number of agents required for different number of calls

5. Summary

A large part of the cost of the call center is spent on labor costs, so it is very important to determine number of agents in call center. When the service time of the call center is lognormal distribution, in the case that the service level is not less than 80%, this paper adopts the discrete event simulation method to solve the problem of how many agents are needed at least per hour from 8:00 to 22:00. Finally, the change trend of the minimum number of agents with the number of calls is analyzed. This is of great reference value to the staffing of the call center.

References

[1] Gans N, Koole G, Mandelbaum A, Telephone call centers: Tutorial, review, and research prospects, J. Manufacturing Service Oper. Management. 5 (2003) 79–141.

[2] Brown L, Gans N, Mandelbaum A, et al. Statistical analysis of a telephone call center: A queueing-science perspective. J. Amer. Statist. Assoc. 100(2005):36–50.

[3] Rouba Ibrahim, Pierre L'Ecuyer, Haipeng Shen, et al. Inter-dependent, heterogeneous, and time-varying service-time distributions in call centers, J. European Journal of Operational Research. 250 (2016) 480-492.

[4] Lawrence Brown et al., Statistical Analysis of a Telephone Call Center, J. Journal of the American Statistical Association. 100 (2005) 36-50.

[5] Luo yimei, Zhu hongfang, He zhen, Study on the statistical distribution of call center service time, J. Industrial Engineering Journal. 9 (2006) 105-108.

[6] Noah Gans, Nan Liu, Avishai Mandelbaum, Haipeng Shen, and Han Ye, Service times in call centers: Agent heterogeneity and learning with some operational consequences, J. Institute of Mathematical Statistics. 6 (2010) 99–123.

[7] Thomas R. Robbins, Experience-Based Routing in Call Center Environments, J. Service Science. 7 (2015) 132–148.