UDC 519.872

Svitlana Pustova

MODELING CALL CENTER OPERATION WITH TAKING INTO ACCOUNT REPEATED ATTEMPTS OF SUBSCRIBERS

NAU Department of Computerized Systems
E-mail: p82004@ukr.net

Mathematical model for call center operation analysis was developed with taking into account repeated subscribers' attempts. The flow of repeated attempts of subscribers influences on the parameters of the quality of service of call center. On the basis of Monte-Carlo method the algorithm of statistical model of call center operation is suggested. This model makes allowance of non-exponentially distributive laws as well as exponentially distributive laws, what makes it more adequate to realistic call centers.

Розроблено математичну модель функціонування центру обробки викликів з урахуванням повторних дзвінків абонентів, які впливають на якісні параметри функціонування центру. На основі методу Монте-Карло запропоновано алгоритм статистичної моделі функціонування системи, яка дозволяє використати як експоненціальні, так і неекспоненціальні закони розподілу та є більш адекватною до реальних центрів обробки викликів.

Problem statement

Call centers are recently obtaining growing public recognition as an effective instrument of interaction between companies and customers [1]. For many companies, such as airlines, hotels, retail banks, trading, insurance, telecommunication and other companies, call centers provide a primary link between customer and service provider and are indispensable means of an effective customer relationship management.

Call center is the common term for a telephone-based human-service operation. A call center provides tele-services in which the subcribers and the service operators are remote from each other. The operators, who sit in cubicles, constitute the physical embodiment of the call center. They serve subcribers over the phone, while facing a computer terminal that outputs and inputs subscriber data. The subscribers are only virtually present: they are either being served or they are delayed in, what's called, tele-queues. Those waiting to be served share a virtual queue, invisible to each other and the operators serving them, waiting and accumulating impatience until one of two things happens – an agent is allocated to serve them, or they abandon the queue, plausibly due to impatience that has built up to exceed their anticipated worth of the service [2].

Call centers can be categorized along many dimensions. The functions that they provide are highly varied: from customer service, help desk, and emergency response services, to tele-marketing and order taking. They vary greatly in size and geographic dispersion, from small sites with a few agents that take local calls – for example at a medical practice – to large national or international centers in which hundreds or thousands of agents may be on the phone at any time.

Furthermore, the latest telecommunications and information technology allow a call center to be the virtual embodiment of a few or many geographically dispersed operations.

These range from small groups of very large centers that are connected over several continents – for example, in the USA, Ireland, and India – to large collections of individual agents that work from their homes.

Contact centers are the contemporary successors of call centers. In addition to phone services, they interface with customers with the help of the Internet, e-mail, chat and fax.

Call centers are the preferred and prevalent way for many companies to communicate with their customers. The call center industry is thus vast and rapidly expanding in terms of both workforce and economic scope. Call centers have become an important part of modern business and that's why the development of a valid mathematical model is a top priority task for call center owners.

Research and publications analysis

The instrument of call center mathematical modeling is the means of the queuing theory [3], with the help of which call center analytical model is built. Over the last 20 years retrial queuing systems are being intensively studying.

These models adequately describe call center operation indeed.

It should be admitted that the only monograph on the theory of retrial queuing systems is the paper written by Falin and Templeton [4]. At the same time Aguir [5] and other researchers of call center statistical data show that disregarding the retrial phenomenon in call centers can lead to huge distortions in subsequent forecasting and staffing analysis.

Objectives

The aim of the research is to describe call center operation process, to develop analytical and statistical models of call center operation process adequate to real system and to analyze achieved results.

Call center operation process

Call center operators that serve subscriber calls are often called agents. Call center operation process can be seen as shown in the fig. 1.

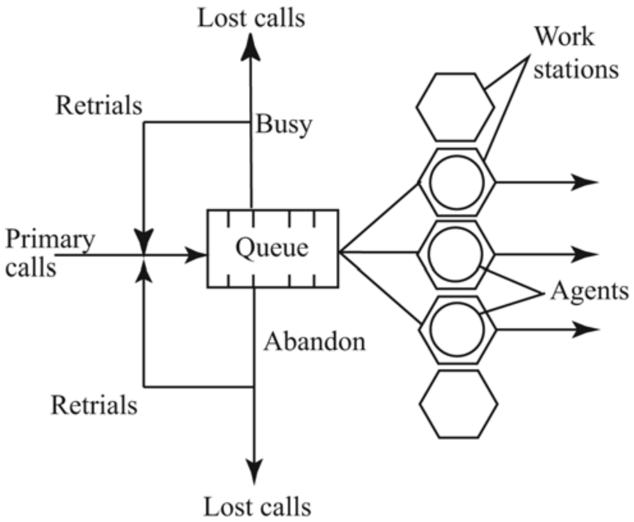

Fig. 1. Call center operation scheme

At the input of call center with c agents the primary subscriber calls arrive with certain rate. When a subscriber call arrives, it will be served immediately if a server is available. If all servers are busy with other calls, the subscriber will be blocked from entering a system. Then the subscriber will be put on hold, and will be asked to wait until a server becomes available. Some subscribers are patient enough to wait for a server to become available, while others will hang-up or abandon. They may either leave the system (then the call is said to be lost) or make a retrial call (i.e. to produce a source of repeated calls). The pool of subscribers, that are waiting to repeat their call, are said to form the source of retrial calls – the orbit. An orbit may be viewed as a queue, which nevertheless differs from an ordinary queue. First of all, a retrial call can't be served by an agent just after he/she completes the previous call. It can be handled only if subscriber retries for service again and accesses a free agent at the time of arrival. Subscribers can make these repeated attempts after random-distributed amount of time.

In the fig.1 it is clearly seen that retrials form one of the components of the input flow of calls; that's why their ignorance can greatly influence on the main call center operation parameters, especially during the optimization call center sizing in order to ensure certain service level.

Hence, call center operation is considered as multi-channel retrial queuing system in which agents provide service to subscribers, with input flows of primary and repeated calls, waiting places and abandons.

Analytical model

Call center operation process can be viewed as retrial queuing systems with abandons (fig. 2).

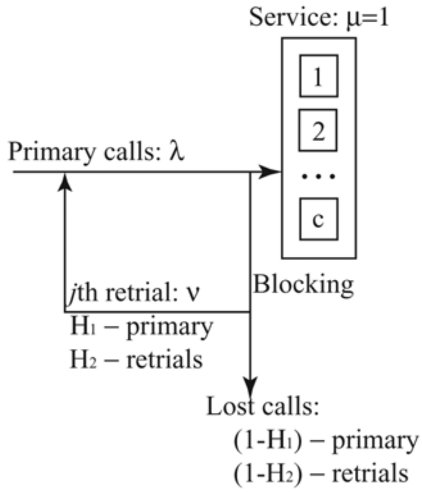

Fig. 2. Call center as M/M/c retrial queue with abandons

For the sake of simplicity call center operation process is being viewed without waiting places, considering that subscribers waiting to the service form the orbit. The adequacy of this model is explained in such way that call center management policy is generally based on disconnection of the subscribers with asking to repeat their calls again later in order not to keep them waiting for service. Consider a calling subscriber leaves the system after some unsuccessful attempts. Let H_j be a probability that after j-th unsuccessful attempt a subscriber will make another $(j+1)$-th one. The set of probabilities $\{H_j, j \geq 1\}$ is called the persistence function. Assume that the probability of call retrial after unsuccessful attempt does not depend on the quantity of previous repeated attempts (i.e. $H_2 = H_3 = ...$).

Consider a queuing system with c servers, at which a Poisson flow of primary calls arrives with the rate λ. If there is a free server at the time of a call arrival, the call is served immediately and leaves the system. Otherwise, if all servers are busy, the call leaves the system with the probability $(1-H_1)$ without service and forms a source of repeated calls with the probability H_1. Every such source produces a Poisson flow of repeated attempts with the rate ν. If an arriving retrial call finds a free server, it is served and leaves the system, while the source that produced this retrial call disappears.

Otherwise, if all servers are busy at the repeated call arriving time, the source leaves the system with probability $(1-H_2)$ without being served and forms the orbit with probability H_2 until being served or leaving system without service. Let service times be exponentially distributed with the rate μ. Without loss of generality we may assume that $\mu=1$. Suppose intervals between primary and repeated calls and service times are mutually independent.

The system functioning can be described as a bivariate process $(C(t),N(t))$, where $C(t)$ is the number of busy servers, $N(t)$ is the number of sources of repeated calls at time t. Thus the process $(C(t),N(t))$ is Markovian defined on the state space $S=\{0,1,...,c\}\times\{0,1,...\}$. According to the above assumptions the transition diagram is defined as shown in the fig. 3. From a practical point of view the most important parameters of the quality of service to subscribers are:

– the stationary blocking probability:

$B=\lim_{t\to\infty} P\{C(t)=c\}$;

– the average number of calls that are in the orbit:

$N=\lim_{t\to\infty} \mathrm{E}N(t)$;

– the average number of busy servers:

$Y=\lim_{t\to\infty} \mathrm{E}C(t)$;

– the probability that a call will be lost:

$L=1-\dfrac{Y\mu}{\lambda}$.

Ergodicity

With the help of the Tweedie theorem [6] it was shown that the process $(C(t),N(t))$ is ergodic for any arrival rate if $H_2<1$.

For $H_2=1$ this process is ergodic if $\lambda H_1/\mu<c$. It was shown that this condition is sufficient and neccesary for ergodicity.

Explicit formulas for parameters of the $M/M/c$ retrial queue with abandons

Let

$p_{ij}(t)=P\{C(t)=i,\ N(t)=j\}$

be a probability that at the time t the system is in the state (i,j), i.e. i servers are busy and j calls are in the orbit.

Then for the steady-state condition the set of Kolmogorov equations is as followed:

$$(\lambda+i+j\nu)p_{ij}=\lambda p_{i-1,j}+(j+1)\nu p_{i-1,j+1}+$$
$$+(i+1)p_{i+1,j},\ 0\le i\le c-1,\ 0\le j<\infty,$$

$$(\lambda H_1+j\nu(1-H_2)+c)p_{cj}=\lambda p_{c-1,j}+(j+1)\nu p_{c-1,j+1}+$$
$$+\lambda H_1 p_{c,j-1}+(j+1)\nu(1-H_2)p_{c,j+1},\ i=c,\ 0\le j<\infty$$

and the normalizing condition

$$\sum_{i=0}^{c}\sum_{j=0}^{\infty}p_{ij}=1.$$

With the use of generating functions the next formulas for parameters of the system were obtained:

$$N=\frac{\lambda H_2+\lambda(H_1-H_2)B-H_2Y}{\nu(1-H_2)},\ \text{if } H_2<1,$$

$$Y=\lambda-\lambda(1-H_1)B,\ \text{if } H_2=1,$$

$$N=\frac{1+\nu}{\nu(c-\lambda H_1)}\Big[\lambda+\lambda^2-\mathrm{E}C^2(t)-\lambda(1-H_1)\times$$
$$\times\left(\lambda+c+1-\frac{\lambda H_1}{1+\nu}\right)B\Big],\ \text{if } H_2=1.$$

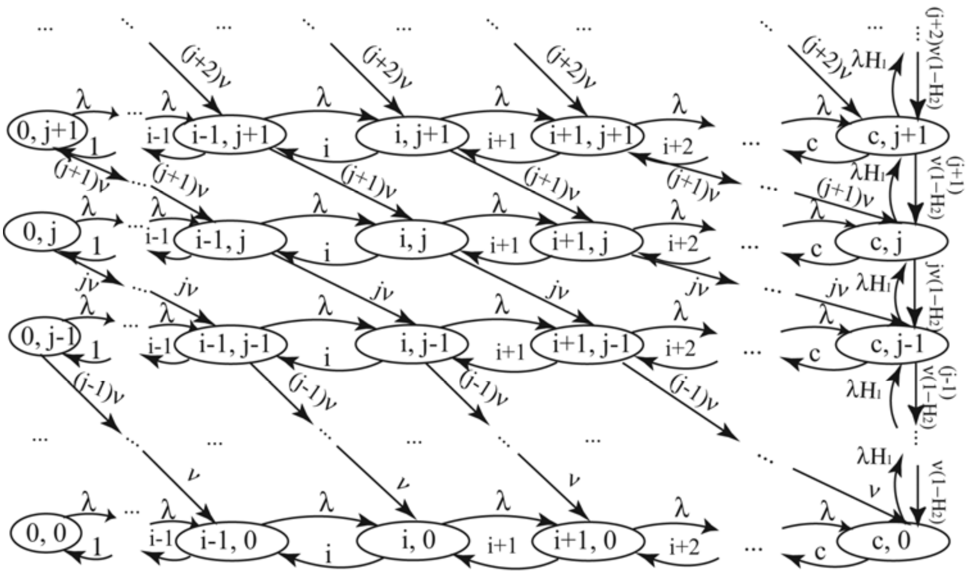

Fig. 3. The transition diagram of the $M/M/c$ retrial queue with abandons

And the next condition was also received:

$$Y = \lambda - \lambda(1 - H_1)B - \nu(1 - H_2)N_c ,$$

where $N_c = \mathrm{E}\{N(t); C(t) = c\}$.

The software algorithm for numerical calculation of stationary distribution of call center as *M/M/c* retrial queue with abandons and bounded orbit was suggested and implemented in a program.

As a result we obtained the next dependence shown in the fig. 4.

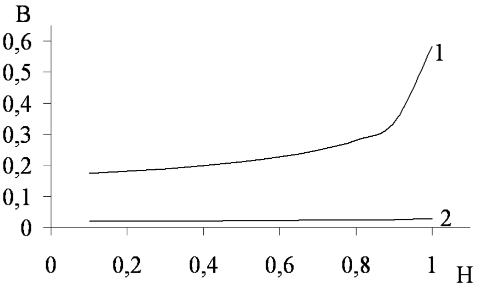

Fig. 4. The dependence of blocking probability B on the input flow intensity λ and repetition probability H, $\mu = 5$, $c = 10$:
1 – $\lambda = 9$; *2* – $\lambda = 5$

These results show high sensitivity of the blocking probability B to the repetition probability in the case of very persistent customers when H is close to one.

Statistical model

As it is quite complicated to obtain explicit formulas and results for retrial queuing systems with non-exponentially distributed times, the statistical model of call center is built.

Thus, on the basis of Monte-Carlo method [7] the algorithm of statistical model of call center operation is suggested.

This model makes allowance of non-exponentially distributive laws as well as exponentially distributive laws, what makes it more adequate to realistic call centers. It was shown that the statistical model converges to the analytical and the average relative error is 2–3%; this proves the model's adequacy.

As a result of the statistical model the dependence shown in the fig. 5 was obtained.

Call center square-root staffing rule

The suggested models can be used while estimating the optimal number of working agents for cost minimization, because the most part of the expenditure of call center maintenance is hold by salary of its agents. We should admit that there is a square-root staffing rule [8] for call centers:

$$C_{opt} = R + \beta\sqrt{R} ,$$

where C_{opt} is the optimal number of agents; β is a positive constant, which defines a service level; R is system load (e.g., for $M/M/1$ queue $R = \lambda/\mu$; λ is input flow intensity; $1/\mu$ is average service time).

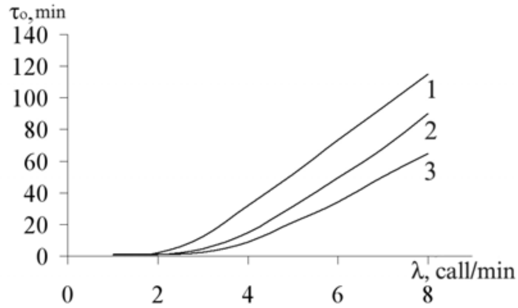

Fig. 5. The dependence of the average waiting time τ_0 in the orbit on the intensity of primary calls λ and quantity of agents c:
1 – $c = 10$; *2* – $c = 12$; *3* – $c = 14$

Conclusions

Hence, call center operation process is described in detail. Analytical and statistical models of call center operation are developed as retrial queuing systems. These models adequately describe call center operation, because they take into account the secondary, tertiary, etc. input repeated attempts of subscribers. The flow of repeated attempts influences on the characteristics of the quality of service of the call center. The adequacy of the statistical model was proved heuristically. The research results may be introduced while organizing a new call center or increasing the effectiveness of the existent one.

References

1. *Gans N., Koole G., Mandelbaum A.* Telephone call centers: Tutorial, review, and research prospects // Invited review paper by Manufacturing and Service Operations Management (M&SOM). – 2003. – 5 (2). – P. 79–141.
2. *Mandelbaum A., Sakov A., Zeltyn S.* Empirical analysis of a call center // Technical Report. – Technion, 2001. – 30 p.
3. *Гнеденко Б.В., Коваленко И.Н.* Введение в теорию массового обслуживания. – М.: Наука, 1987. – 336 с.
4. *Falin G.I., Templeton J.G.C.* Retrial queues. – London: Chapmen & Hall. – 1997. – 395 p.
5. *Aguir M. S., Karaesmen F., Aksin Z., Chauvet F.* The Impact of Retrials on Call Center Performance // OR Spectrum. – 2004. – 26(3). – P. 353–376.
6. *Tweedie R.L.* Sufficient conditions for regularity, recurrence and ergodicity of Markov processes // Math. Proc. of the Cambridge Philosophical Society. – 1975. – N78, p. 1. – P. 125–136.
7. *Ермаков С.М., Михайлов Г.А.* Курс статистического моделирования. – М.: Наука, 1976. – 320 с.
8. *Whitt W.* Understanding the efficiency of multi-server service systems // Management Sci. – 1992. – 38 (5). – P. 708–723.

The editors received the article on 5 October 2006.