The Tenth International Symposium on Operations Research and Its Applications (ISORA 2011)
Dunhuang, China, August 28–31, 2011

Analysis of a Two-Phase Queueing System with Impatient Customers and Multiple Vacations

Dequan Yue[1] Wuyi Yue[2] Xiuju Li[3]

[1]Department of Statistics, College of Science, Yanshan University, Qinhuangdao 066004, China
[2]Department of Intelligence and Informatics, Konan University, Kobe 658-8501, Japan
[3]College of Liren, Yanshan University, Qinhuangdao 066004, China

Abstract In this paper, we consider a two-phase queueing system with impatient customers and multiple vacations. Customers arrive at the system according to a Poisson process. They receive the first essential service as well as a second optional service. Arriving customers may balk with a certain probability and may depart after joining the queue without getting service due to impatience. Lack of service occurs when the server is on vacation or busy during the first phase of service. We analyze this model and derive the probability generating functions for the number of customers present in the system for various states of the server. We further obtain the closed-form expressions for various performance measures including the mean system sizes for various states of the server, the average rate of balking, the average rate of reneging, and the average rate of loss.

Keywords Queueing systems; impatient customers; multiple vacations; probability generating functions; mean system sizes.

1 Introduction

In real life, many queueing situations arise in which there may be a tendency for customers to be discouraged by a long queue. As a result, a customer may either decide not to join the queue (i.e. balk) or depart after joining the queue without getting service due to impatience (i.e. renege). Queueing systems with impatient customers appear in many real life situations such as those involving impatient telephone switchboard customers, hospital emergency rooms handling critical patients, and inventory systems that store perishable goods. There is growing interest in the analysis of queueing systems with impatient customers. This is due to their potential application in communication systems, call centers, production-inventory systems and many other related areas, see for instance [1]-[3] and the references therein.

Queueing systems with impatient customers have been studied by a number of authors. There is an extensive amount of literature based on this kind of model and we refer the reader to [4]-[9] and references therein. In all this papers, the source of impatience has always been taken to be either a long wait already experienced upon arrival at a queue, or a long wait anticipated by a customer upon arrival.

Recently, Altman and Yechiali [10], [11] analyzed queueing models where customers become impatient only when the servers are on vacation and unavailable for service. In other words, the cause of the impatience is the absence of the server. The M/M/1, M/G/1 and M/M/c queues were investigated in [10], and the M/M/∞ queue was investigated in [11]. Yue and Yue [12] studied an M/M/c/N queue with synchronous vacations by considering customers' impatience under balking and reneging. Perel and Yechiali [13] analyzed queueing system operating in a 2-phase (fast and slow) Markovian random environment. When in the slow phase, customers became impatient due to a slow service rate.

In all the papers mentioned above, it is assumed that only one phase service is provided by servers. However, in many real service systems, the server may provide a second optional service. To the best of our knowledge, for two-phase queueing systems with vacations, there is no literature which takes customers' impatience into consideration. In this paper, we consider a two-phase queueing system with multiple vacations and impatient customers.

The paper is organized as follows: In Section 2, we describe the model. In Section 3, we carry out the equilibrium analysis of the system state and derive the probability generating functions of the number of customers present in the system for various states of the server. We further obtain the closed-form expressions for various performance measures including the mean system sizes for various states of the server, the average rate of balking, the average rate of reneging, and the average rate of loss. Conclusions are given in Section 4.

2 Model Descriptions

In this section, we consider an M/M/1 queueing system with impatient customers and multiple vacations, in which the customers receive two-phase service.

Customers arrive according to a Poisson process at rate λ. They receive the first essential service (FES) as well as the second optional service (SOS) from a single server, who serves the customers on a first-come first-served (FCFS) basis. The FES is needed by all arriving customers. As soon as the FES of a customer is completed, then with a probability θ, the customer may opt for the SOS, or else with a probability $1-\theta$, it may opt to leave the system, in which case another customer at the head of the queue (if any) is taken up for its FES. The service times of the FES and the SOS have exponential distributions with parameters μ_1 and μ_2, respectively.

When the server finishes serving a customer and finds the system empty, the server leaves for a vacation of random length V. On return from a vacation if the server finds more than one customer waiting, it serves the customer at the head of the queue for its FES and continues to serve in this manner until the system is empty. Otherwise, it immediately goes for another vacation. The vacation time V has an exponential distribution with parameter η.

A customer who on arrival finds the server is on vacation (or busy carrying out the FES of a customer), either decides to join the queue with probability b_0 (or b_1), or balk with probability $1-b_0$ (or $1-b_1$). After joining the queue, in the case that the server is on vacation (or busy with the FES of a customer), each customer will wait a certain length of time T_0 (or T_1) for service to begin before they become impatient and leave the queue without being served. These times T_0 and T_1 are assumed to be distributed exponentially with parameters ξ_0 and ξ_1, respectively.

We assume that all the random variables defined above are independent.

Remark 1. If $\theta = 0$, $\xi_1 = 0$, $\mu_2 = 0$ and $b_0 = b_1 = 1$, the current model reduces to the M/M/1 model with impatient customers and multiple vacations that was studied by Altman and Yechiali [10].

3 Stationary Analysis

In this section, we present a stationary analysis for the model described in the last section. We first derive the probability generating functions of the number of customers present in the system for various states of the server. Then, we derive the closed-form expressions for various performance measures including the mean system size for various states of the server, the average rate of balking, the average rate of reneging and the average rate of loss.

3.1 Balance Equations and Generating Functions

Let L denote the number of customers in the system, and let J denote the status of the server, which is defined as follows: $J = 0$ denotes that the server is on vacation, $J = 1$ denotes that the server is busy with a FES, and $J = 2$ denotes that the server is busy with a SOS. Then, the pair (J, L) defines a continuous-time Markov process with state space $\Omega = \{0,0\} \cup \{(j,n) : j = 0,1,2; n = 1,2,...\}$.

Let $P_{jn} = P\{J = j, L = n\}$, $j = 0,1,2$; $n = 0,1,2,...$, denote the steady state probabilities of the system. Then, the set of balance equations is given as follows:

$$\lambda b_0 P_{00} = \xi_0 P_{01} + \mu_1(1-\theta)P_{11} + \mu_2 P_{21}, \tag{1}$$

$$(\lambda b_0 + \xi_0 + \eta)P_{0n} = \lambda b_0 P_{0n-1} + (n+1)\xi_0 P_{0n+1},\ n \geq 1, \tag{2}$$

$$\begin{aligned}[\lambda b_1 + \mu_1 + (n-1)\xi_1]P_{1n} = \eta P_{0n} + \lambda b_1 P_{1n-1} + [\mu_1(1-\theta) + n\xi_1]P_{1n+1} \\ + \mu_2 P_{2n-1},\ n \geq 1,\end{aligned} \tag{3}$$

$$(\lambda + \mu_2)P_{2n} = \mu_1 \theta P_{1n} + \lambda P_{2n-1},\ n \geq 1, \tag{4}$$

$$P_{00} + \sum_{n=1}^{\infty}(P_{0n} + P_{1n} + P_{2n}) = 1 \tag{5}$$

where $P_{10} \equiv 0$ and $P_{20} \equiv 0$.

Define the probability generating functions (PGFs) as

$$Q_j(z) = \sum_{n=1}^{\infty} P_{jn} z^n,\ |z| \leq 1,\ j = 0,1,2.$$

Then, $Q_0(z)$, $Q_1(z)$ and $Q_2(z)$ are the probability generating functions of the number of customers present in the system when the server is on vacation, busy with a FES, and busy with a SOS, respectively.

Multiplying each equation for n in Eq. (2) by z^n, and summing all possible values of n and re-arranging terms, we get

$$\xi_0(1-z)Q_0'(z) - [\eta + \lambda b_0(1-z)]Q_0(z) = -\lambda b_0 P_{00} z + \xi_0 P_{01} \tag{6}$$

where $Q_0'(z) = \frac{d}{dz} Q_0(z)$. Similarly, using Eq. (3), we obtain

$$\xi_1(1-z)Q_1'(z) - f(z)Q_1(z) + \frac{\mu_2}{z} Q_2(z) + \eta Q_0(z) = \mu_1(1-\theta)P_{11} + \mu_2 P_{21} \tag{7}$$

where

$$f(z) = \lambda b_1(1-z) + (\xi_1 - \mu_1)(\frac{1}{z} - 1) + \frac{\mu_1 \theta}{z} \tag{8}$$

and $Q_1'(z) = \frac{d}{dz} Q_1(z)$. Similarly, using Eq. (4), we get

$$Q_2(z) = \frac{\mu_1 \theta}{\lambda(1-z) + \mu_2} Q_1(z). \tag{9}$$

By solving the differential equations (6) and (7), we can obtain $Q_0(z)$ and $Q_1(z)$. The following theorems give the solutions of Eqs. (6) and (7).

Theorem 1. The probability generating function $Q_0(z)$ of the number of customers present in the system when the server is on vacation is expressed as follows:

$$Q_0(z) = \frac{\lambda b_0}{\xi_0} e^{\frac{\lambda b_0}{\xi_0} z} (1-z)^{-\frac{\eta}{\xi_0}} F(z) P_{00} \tag{10}$$

where

$$F(z) = \int_0^z \left(\frac{A}{B} - x \right) e^{-\frac{\lambda b_0}{\xi_0} x} (1-x)^{\frac{\eta}{\xi_0} - 1} dx \tag{11}$$

with

$$A = \int_0^1 e^{-\frac{\lambda b_0}{\xi_0} x} x (1-x)^{\frac{\eta}{\xi_0} - 1} dx, \tag{12}$$

and

$$B = \int_0^z e^{-\frac{\lambda b_0}{\xi_0} x} (1-x)^{\frac{\eta}{\xi_0} - 1} dx. \tag{13}$$

Proof. For $z \neq 1$, Eq. (6) can be written as follows:

$$Q_0'(z) + \left[-\frac{\lambda b_0}{\xi_0} - \frac{\eta}{\xi_0(1-z)} \right] Q_0(z) = -\frac{1}{\xi_0} (\lambda b_0 P_{00} z - \xi_0 P_{01}) (1-z)^{-1}. \tag{14}$$

Multiplying both sides of Eq. (14) by $e^{-\frac{\lambda b_0}{\xi_0} z} (1-z)^{\frac{\eta}{\xi_0}}$, we get

$$\frac{d}{dz} \left[e^{-\frac{\lambda b_0}{\xi_0} z} (1-z)^{\frac{\eta}{\xi_0}} Q_0(z) \right] = -\frac{1}{\xi_0} (\lambda b_0 P_{00} z - \xi_0 P_{01}) e^{-\frac{\lambda b_0}{\xi_0} z} (1-z)^{\frac{\eta}{\xi_0} - 1}. \tag{15}$$

By solving Eq. (15) we can obtain $Q_0(z)$ which is given by Eq. (10). The details are omitted.

Theorem 2. The probability generating function $Q_1(z)$ of the number of customers present in the system when the server is busy with a FES is expressed as follows:

$$Q_1(z) = \frac{\lambda b_0}{\xi_1} e^{\frac{\lambda b_1}{\xi_1} z} z^u (\lambda(1-z)+\mu_2)^{-v} H(z) P_{00} \tag{16}$$

where

$$H(z) = \int_0^z e^{-\frac{\lambda b_1}{\xi_1} x} x^{-u} (1-x)^{-(1+\frac{\eta}{\xi_0})} (\lambda(1-x)+\mu_2)^v G(x) dx \tag{17}$$

with

$$G(z) = \left(1-\frac{A}{B}\right)(1-z)^{\frac{\eta}{\xi_0}} - \frac{\eta}{\xi_0} e^{\frac{\lambda b_0}{\xi_0} z} F(z) \tag{18}$$

and

$$u = 1 - \frac{\mu_1(1-\theta)}{\xi_1} - \frac{\mu_1\mu_2\theta}{\xi_1(\lambda+\mu_2)}, \quad v = \frac{\lambda\mu_1\theta}{\xi_1(\lambda+\mu_2)}. \tag{19}$$

Proof. The details of the proof are omitted.

Eqs. (10) and (16) express $Q_0(z)$ and $Q_1(z)$ in terms of P_{00}. Also, from Eq. (9), $Q_2(z)$ is a function of $Q_1(z)$. Thus, once P_{00} is calculated, $Q_0(z)$, $Q_1(z)$ and $Q_2(z)$ are completely determined. We derive the probabilities P_{00} and the mean system sizes for various states of the server in the next section.

3.2 Mean System Sizes

In order to derive the mean system sizes for various states of the server, we first derive the probability P_{00}.

Let $z = 1$ in Eqs. (10), (16) and (9), we obtain

$$Q_0(1) = \frac{\lambda b_0}{\eta}\left(1-\frac{A}{B}\right) P_{00}, \tag{20}$$

$$Q_1(1) = \frac{\lambda b_0}{\xi_0} e^{\frac{\lambda b_1}{\xi_1}} \mu_2^{-v} H(1) P_{00} \tag{21}$$

and

$$Q_2(1) = \frac{\mu_1\theta}{\mu_2} Q_1(1). \tag{22}$$

Noting that $Q_j(1) = \sum_{n=1}^{\infty} P_{jn}$, $j = 0,1,2$, and using Eq. (5), we have

$$P_{00} + Q_0(1) + Q_1(1) + Q_2(1) = 1. \tag{23}$$

Substituting Eqs. (20), (21) and (22) into Eq. (23), we get

$$P_{00} = \left\{ 1 + \frac{\lambda b_0}{\xi_0}\left(1 - \frac{A}{B}\right) + \frac{\lambda b_0}{\xi_0}\left(1 + \frac{\mu_1\theta}{\mu_2}\right) e^{\frac{\lambda b_1}{\xi_1}} \mu_2^{-\nu} H(1) \right\}^{-1}. \tag{24}$$

Let $E(L_j) = Q_j'(1) = \sum_{n=1}^{\infty} nP_{jn}$, $j = 0,1,2$, denote the mean system sizes for various states of the server.

Theorem 3. The mean system size $E(L_0)$ when the server is on the vacation is given by

$$E(L_0) = \frac{\lambda b_0}{\xi_0 + \eta}\left[1 + \left(1 - \frac{A}{B}\right)\frac{\lambda b_0}{\xi_0}\right] P_{00}. \tag{25}$$

The mean system size $E(L_1)$ when the server is busy with a FES is given by

$$E(L_1) = \frac{\eta}{\xi_1}E(L_0) + \frac{1}{\xi_1}\left[(\lambda b_1 + \xi_1 - \mu_1)Q_1(1) + \lambda Q_2(1)\right]. \tag{26}$$

The mean system size $E(L_2)$ when the server is busy with a SOS is given by

$$E(L_2) = \frac{\lambda}{\mu_2}Q_2(1) + \frac{\mu_1\theta}{\mu_2}E(L_1). \tag{27}$$

Proof. The details of the proof are omitted.

Thus, the mean system size can be calculated by

$$E(L) = E(L_0) + E(L_1) + E(L_2). \tag{28}$$

3.3 Other Performance Measures

In this section, we present other important performance measures including the average rate of balking, denoted by R_{balk}, the average rate of reneging, denoted by R_{reneg}, and the average rate of loss, denoted by R_{loss}.

When the system is in state (jn), $j = 0,1$, then the instantaneous rate of balking is $\lambda(1-b_j)$, $j = 0,1$. Thus, the average rate of balking is given by

$$\begin{aligned} R_{balk} &= \sum_{n=0}^{\infty}\lambda(1-b_0)P_{0n} + \sum_{n=0}^{\infty}\lambda(1-b_1)P_{1n} \\ &= \lambda(1-b_0)(P_{00} + Q_0(1)) + \lambda(1-b_1)Q_1(1) \end{aligned} \tag{29}$$

where $Q_0(1)$, $Q_1(1)$ and P_{00} are given by Eqs. (20), (21) and (24), respectively.

When the system is in state $(0n)$, there are n waiting customers in the queue. Thus, the instantaneous rate of reneging is $n\xi_0$. However, when the system is in state $(1n)$, there are $n-1$ waiting customers in the queue. Thus, the instantaneous rate of reneging is $(n-1)\xi_0$. Therefore, the average rate of reneging is given by

$$\begin{aligned} R_{renege} &= \sum_{n=1}^{\infty} n\xi_0 P_{0n} + \sum_{n=0}^{\infty}(n-1)\xi_1 P_{1n} \\ &= \xi_0 E(L_0) + \xi_1[E(L_1) - Q_1(1)] \end{aligned} \tag{30}$$

where $Q_1(1)$, $E(L_0)$ and $E(L_1)$ are given by Eqs. (21), (25) and (26), respectively.

The average rate of loss (due to balking and reneging) is given by $R_{loss} = R_{balk} + R_{renege}$.

4 Conclusions

In this paper a two phase queueing model with impatient customers and multiple vacations is discussed. Our model considers customers' impatience under the balking and the reneging. The model investigated is more realistic than those existing ones since the customers may get impatience not only when the server is on vacation but also is busy carrying out the first essential service. The closed form expressions for various performance measures including the mean system sizes for various states of the server, the average rate of balking, the average rate of reneging, and the average rate of loss, are obtained.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (No. 71071133), and is supported in part by GRANT-IN-AID FOR SCIENCE RESEARCH (No. 21500086), Japan.

References

[1] T. Bonald and J. Roberts, Performance modeling of elastic traffic in overload, *ACM Sigmetrics Performance Evaluation Review*, 29, 342-343, 2001.

[2] N. Gans, G. Koole and A. Mandelbaum, Telephone call centers: Tutotial, review, and research prospects, *Manufacturing and Service Operations Management*, 5, 79-141, 2003.

[3] S. Benjaafar, J. Gayon and S. Tepe, Optimal control of a production-inventory system with customer impatience, *Operations Research Letters*, 38, 267-272, 2010.

[4] C. Palm, Methods of judging the annoyance caused by congestion, *Tele*, 4, 189-208, 1953.

[5] D. J. Daley, General customer impatience in the queue GI/G/1, *Journal of Applied Probability*, 2, 186-205, 1965.

[6] L. Takacs, A single-server queue with limited virtual waiting time, *Journal of Applied Probability*, 11, 612-617, 1974.

[7] F. Baccelli, P. Boyer and G. Hebuterne, Single-server queues with impatient customers, *Advances in Applied Probability*, 16, 887-905, 1984.

[8] O. J. Boxma and P. R. de Waal, "Multiserver queues with impatient customers," *ITC*, 14, 743-756, 1994.

[9] B. Van Houdt, R. B. Lenin and C. Blonia, Delay distribution of (im)patient customers in a discrete time D-MAP/PH/1 queue with age-dependent service times, *Queueing Systems*, 45, 59-73, 2003.

[10] E. Altman and U. Yechiali, Analysis of customers' impatience in queues with server vacations, *Queueing Systems*, 52, 261-279, 2006.

[11] E. Altman and U. Yechiali, Infinite-server queues with systems' additional task and impatient customers, *Probability in the Engineering and Informational Sciences*, 22, 477-493, 2008.

[12] D. Yue and W. Yue, Analysis of M/M/c/N queueing system with balking, reneging and synchronous vacations, In: Wuyi Yue et al. (Eds.), *Advanced in Queueing Theory and Network Applications*, New York: Springer, 165-180, 2009.

[13] N. Perel and U. Yechiali, Queues with slow servers and impatient customers, *European Journal of Operational Research*, 201, 247-258, 2010.