한국경영과학회지
제39권 제3호
2014년 9월
http://dx.doi.org/10.7737/JKORMS.2014.39.3.001

고객수 상태에 따른 서비스를 제공하는 M/G/1/K 대기체계에 관한 소고

최두일[1] · 김보근[2] · 이두호[3†]
[1]한라대학교 교양교직과정부, [2]삼성 SDS, [3]한국전자통신연구원 SW컨텐츠연구소

A Note on the M/G/1/K Queue with Two-Threshold Hysteresis Strategy of Service Intensity Switching

Doo Il Choi[1] · Bo Keun Kim[2] · Doo Ho Lee[3]
[1]Division of Liberal arts and Teaching, Halla University, [2]Samsung SDS,
[3]SW·Content Research Laboratory, ETRI

■ Abstract ■

We study the paper Zhernovyi and Zhernovyi [Zhernovyi, K.Y. and Y.V. Zhernovyi, "An $M^{\theta}/G/1/m$ system with two-threshold hysteresis strategy of service intensity switching," *Journal of Communications and Electronics*, Vol.12, No.2(2012), pp.127-140]. In the paper, authors used the Korolyuk potential method to obtain the stationary queue length distribution. Instead, our note makes an attempt to apply the most frequently used methods : the embedded Markov chain and the supplementary variable method. We derive the queue length distribution at a customer's departure epoch and then at an arbitrary epoch.

Keywords : M/G/1/K Queue, Queue Length Dependent Service Rate, Thresholds, Finite Buffer

1. Introduction

In this paper, the finite capacity queueing system with threshold based service is analyzed. Queueing systems with finite buffers arise in a wide variety of applications such as computer

논문접수일 : 2014년 03월 17일 논문게재확정일 : 2014년 06월 10일
논문수정일(1차 : 2014년 05월 07일)
† 교신저자 enjdh@etri.re.kr

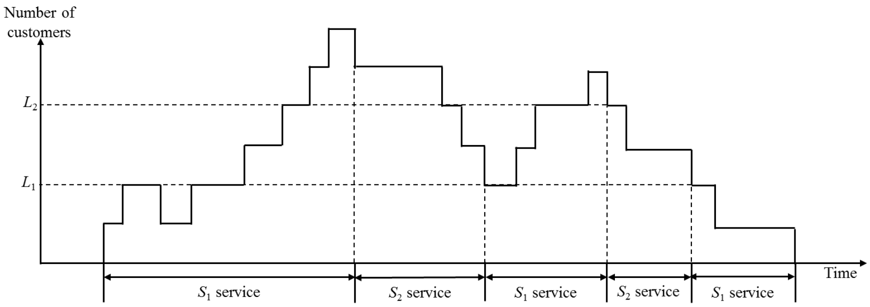
systems, telecommunication networks, production lines and so on. While operating systems with a queue, in which the arrivals at the systems and the services of customers occur randomly, some customers may suffer long delay. This may finally cause the situation that the delay requirement of users is not satisfied. One of the solutions to this problem is to control the service characteristics, typically times. It means the more customers in the system, the faster services provided. For example, a queueing system with variable service times depending upon the number of customers in the system operates as follows : when the number of customers in the system is less than the predetermined value, namely threshold, customers are served in ordinary service times. Meanwhile, when the number of customers exceeds the threshold, the service rate is increased to a certain value to expedite reducing waiting times.

Recently, Zhernovyi and Zhernovyi [1] analyzed the queueing model with threshold based service times. In their work, authors used the Korolyuk potential method to obtain the stationary queue length distribution. The purpose of this note is to apply two different methods that are the most frequently used by queueing analysts: the embedded Markov chain and the supplementary variable method. The embedded Markov chain is useful and intuitive when we establish the mathematical equations for modeling. The supplementary variable method in our note plays an auxiliary role of obtaining the stationary queue length distribution. Our study may help readers more easily understand Zhernovyi and Zhernovyi's [1] model.

The remainder of this paper is organized as follows. Section 2 describes our queueing model. In Section 3, the main results, queue length distributions at a customer's departure epoch and at an arbitrary epoch are given, respectively.

2. Model Description

The customers arrive according to a Poisson process with rate λ. The customers are accommodated in buffer if the server is not available. The buffer is assumed to be finite with capacity K for real applications. The customers arriving when the buffer is full are blocked and lost. There is a single server. The customers are served by First-Come-First-Service based on their arrival order. The service time of customers has different service times according to the number of customers in the system. Concretely, there are two thresholds L_1 and $L_2 (L_1 \le L_2)$ on buffer. The server is initially idle and starts to work with the service time S_1 on a customer's arrival. If the number of customers in the system is equal to or greater than the threshold L_2 at the service initiation, the customers are served by the service time S_2. This service time also is continued until the number of customers in the system reduces the threshold L_1. If the number of customers in the system is equal to or smaller than L_1 at the service initiation, the customers are served by the service time S_1. This process is repeated. For $r \in \{1, 2\}$, the service time S_r has the distribution function G_r, mean μ_r, and the Laplace transform $G_r^*(s)$. We assume $\mu_2 \le \mu_1$ because customers must be served faster when there are many customers in the system. In [Figure 1], we present one sample path of our queueing model for clear understanding.



[Figure 1] An Example of the Sample Path

3. Analysis

3.1 Queue Length Distribution at Departure Epochs

In this section, we derive the queue length distribution at departure epochs of customers. We define the period in which the service time of customers is generated by the service time S_1 as the underload period, and define the period in which the service time of customers is generated by the service time S_2 as the overload period. Let us introduce the following notations :

$\tau_n =$ the nth customer's departure epoch, $n \geq 1$, $\tau_0 = 0$.

$N_n =$ the number of customers in the system right after τ_n.

$$\xi_n = \begin{cases} 1, & \text{if the system is in the underload period right after } \tau_n, \\ 2, & \text{if the system is in the overload period right after } \tau_n. \end{cases}$$

Then, the process $\{(N_n, \xi_n), n \geq 0\}$ form a Markov chain with finite state space

$$\{(0,1),\ (0,2),\ (1,1),\ (1,2), \cdots, (K-1,1),\ (K-1,2)\}.$$

Note that if $N_n \leq L_1$, then $\xi_n = 1$, and if $N_n \geq L_2$, then $\xi_n = 2$. We define the steady-state probability of the Markov chain $\{(N_n, \xi_n), n \geq 0\}$ as follows :

$$x_{k,r} = \lim_{n \to \infty} \Pr\{N_n = k,\ \xi_n = r\},\ 0 \leq k \leq K-1,\ r = 1, 2$$

Note that $x_{0,2} = 0$ for $0 \leq k \leq L_1$, $x_{k,1} = 0$ for $L_2 \leq k \leq K-1$. In order to derive $x_{k,r}$, we introduce the following probabilities :

$$\alpha_n^r = \Pr\{n \text{arrivals of customers during the service time } S_r\}$$
$$= \int_0^\infty \frac{e^{-\lambda x}(\lambda x)^n}{n!} dG_r(x), \quad r = 1, 2,$$
$$\overline{\alpha}_n^r = \sum_{k=n}^{\infty} \alpha_k^r.$$

We also introduce the matrices :

$$A_k = \begin{pmatrix} \alpha_k^1 & 0 \\ 0 & 0 \end{pmatrix},\ 0 \leq k \leq L_2 - 1, \quad A'_k = \begin{pmatrix} 0 & \alpha_k^1 \\ 0 & 0 \end{pmatrix},\ k \geq L_2,$$

$$B_k = \begin{pmatrix} \alpha_k^1 & 0 \\ 0 & \alpha_k^2 \end{pmatrix},\ 0 \leq k \leq L_2 - L_1 - 1,\ B_0' = \begin{pmatrix} \alpha_0^1 & 0 \\ \alpha_0^2 & 0 \end{pmatrix},$$

$$B_k' = \begin{pmatrix} 0 & \alpha_k^1 \\ 0 & \alpha_k^2 \end{pmatrix},\ k \geq L_2,\ C_k = \begin{pmatrix} 0 & 0 \\ 0 & \alpha_k^2 \end{pmatrix},\ k \geq 0,$$

and

$$\overline{A}_n' = \begin{pmatrix} 0 & \overline{\alpha}_n^1 \\ 0 & 0 \end{pmatrix}, \quad \overline{A}_n' = \begin{pmatrix} 0 & \overline{\alpha}_n^1 \\ 0 & \overline{\alpha}_n^2 \end{pmatrix}, \quad \overline{C}_n = \begin{pmatrix} 0 & 0 \\ 0 & \overline{\alpha}_n^2 \end{pmatrix}.$$

Then, the transition probability matrix $\mathbf{Q}$ of the Markov chain $\{(N_n, \xi_n), n \geq 0\}$ is given by

$\mathbf{Q}=$

$$\begin{pmatrix}
A_0 & A_1 & A_2 & \cdots & A_{L_1-1} & A_{L_1} & A_{L_1+1} & \cdots & A_{L_2-1} & A'_{L_2} & A'_{L_2+1} & \cdots & A'_{K-3} & A'_{K-2} & \bar{A}'_{K-1} \\
A_0 & A_1 & A_2 & \cdots & A_{L_1-1} & A_{L_1} & A_{L_1+1} & \cdots & A_{L_2-1} & A'_{L_2} & A'_{L_2+1} & \cdots & A'_{K-3} & A'_{K-2} & \bar{A}'_{K-1} \\
0 & A_0 & A_1 & \cdots & A_{L_1-2} & A_{L_1-1} & A_{L_1} & \cdots & A_{L_2-2} & A'_{L_2-1} & A'_{L_2} & \cdots & A'_{K-4} & A'_{K-3} & \bar{A}'_{K-2} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & A_1 & A_2 & A_3 & \cdots & A_{L_2-L_1+1} & A'_{L_2-L_1+2} & A'_{L_2-L_1+3} & \cdots & A'_{K-L_1-1} & A'_{K-L_1} & \bar{A}'_{K-L_1+1} \\
0 & 0 & 0 & \cdots & A_0 & A_1 & A_2 & \cdots & A_{L_2-L_1} & A'_{L_2-L_1+1} & A'_{L_2-L_1+2} & \cdots & A'_{K-L_1-2} & A'_{K-L_1-1} & \bar{A}'_{K-L_1} \\
0 & 0 & 0 & \cdots & 0 & B'_0 & B_1 & \cdots & B_{L_2-L_1-1} & B'_{L_2-L_1} & B'_{L_2-L_1+1} & \cdots & B'_{K-L_1-3} & B'_{K-L_1-2} & \bar{B}'_{K-L_1-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & B_1 & B'_2 & B'_3 & \cdots & B'_{K-L_2-1} & B'_{K-L_2} & \bar{B}'_{K-L_2+1} \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & C_0 & C_1 & C_2 & \cdots & C_{K-L_2-2} & C_{K-L_2-1} & \bar{C}_{K-L_2} \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & C_0 & C_1 & \cdots & C_{K-L_2-3} & C_{K-L_2-2} & \bar{C}_{K-L_2-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & C_1 & C_2 & \bar{C}_3 \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & C_0 & C_1 & \bar{C}_2 \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & & C_0 & \bar{C}_1
\end{pmatrix}$$

The steady-state probability vector $\boldsymbol{x}$ of the Markov chain $\{(N_n, \xi_n), n \geq 0\}$ is given by solving following simultaneous equations : $\boldsymbol{x}\mathbf{Q} = \boldsymbol{x}$ and $\boldsymbol{x}(1, 1, \cdots, 1)^T = 1$.

3.2 Queue Length Distribution at an Arbitrary Time

In this subsection we derive the probability distribution of the queue length at an arbitrary time. Let $N(t)$ be the number of customers in the system at time t, and

$$\psi_n = \begin{cases} 1, & \text{if the system is in the underload period at time } t, \\ 2, & \text{if the system is in the overload period at time } t. \end{cases}$$

Define the stationary probabilities :

$$y_n = \lim_{n \to \infty} \Pr\{N(t) = n\}, \ 0 \leq n \leq K.$$

First, by the key renewal theorem, we have

$$y_0 = \frac{1}{E} x_{0,1} \ \frac{1}{\lambda},$$

where $E = x_{0,1}(\lambda^{-1} + \mu_1) + \Sigma_{n=1}^{L-1} x_{n,1}\mu_1 + \Sigma_{n=1}^{K-1} x_{n,2}\mu_2$, the mean inter-departure time of customers. Next, for $1 \leq n \leq K$, we derive the probabilities y_n by using the supplementary variable method. Let $\tilde{T}(\hat{T})$ be the elapsed (remaining, respectively) service time for the customer in service. Furthermore, we define the stationary joint probability distribution of the number of customers in the system and the remaining service time for the customer in service :

$$\begin{aligned} \alpha_{n,r}(x)dx &= \lim_{t \to \infty} \Pr\{N(t) = n, \psi(t) \\ &= r, x < \hat{T} \leq x + dx\}, \ 1 \leq n \leq K, \ r = 1, 2 \end{aligned}$$

And the Laplace transform of $\alpha_{n,r}(x)$ is given by $\alpha^*_{n,r}(s) = \int_0^\infty e^{-sx}\alpha_{n,r}(x)dx$. In order to derive the queue length distribution at an arbitrary time, we must know the number of arrivals of customers during the elapsed service time. So we also define the following joint probability $\beta_r(n, x)$:

$$\beta_r(n,x)dx = \lim_{t\to\infty} \Pr\{n \text{ arrivals of customers during } \tilde{T},\ \psi(t) = r,\ x < \tilde{T} \le x+dx\},$$
$$n \ge 0,\ r = 1, 2,$$

and the Laplace transform of $\beta_r(n, x)$ is given by $\beta_r^*(n, x) = \int_0^\infty e^{-sx}\beta_r(n, x)dx$. By conditioning the queue length at last service completion epoch before time t, $\alpha_{n,r}^*(s)$ satisfies the following equations for $1 \le n \le K$:

$$\alpha_{n,1}^*(s) = \frac{\mu_1}{E}\left[x_{0,1}\beta_1^*(n-1, s) + \sum_{k=1}^{\min\{n, L_2-1\}} x_{k,1}\beta_1^*(n-k, s)\right],$$

$$\alpha_{n,2}^*(s) = \frac{\mu_2}{E}\left[\sum_{k=L_1+1}^{n} x_{k,2}\beta_2^*(n-k, s)\right]1_{\{n > L_1\}}.$$

By the same method as in Choi et al. [2], $\beta_r^*(n, s)$ is given as follows :

$$\beta_r^*(n, s) = \frac{1}{\mu_r}\left[\sum_{k=0}^{n}\alpha_k^r R_{n-k}(s) - G_r^*(s)R_n(s)\right],\ r = 1, 2$$

where $R_n(s) = (s-\lambda)^{-1}\{\lambda(\lambda-s)^{-1}\}^n$. Finally, substituting $\beta_r^*(n, s)$ into above equations, and putting $s=0$, we obtain the stationary queue length probabilities at an arbitrary time for $1 \le n \le K-1$:

$$\begin{aligned} y_n &= \alpha_{n,1}^*(0) + \alpha_{n,2}^*(0) \\ &= \frac{1}{\lambda E}\left[x_{0,1}\left(1-\sum_{k=0}^{n-1}a_k^1\right) + \sum_{k=1}^{\min\{n, L-1\}} x_{k,1}\left(1-\sum_{l=0}^{n-k}a_l^1\right)\right. \\ &\quad \left. + \sum_{k=L_{1+1}}^{n} x_{k,2}\left(1-\sum_{l=0}^{n-k}a_l^2\right)1_{\{n > L_1\}}\right], \end{aligned}$$

and

$$y_K = 1 - \sum_{n=0}^{K-1} y_n.$$

Thus, by using the stationary queue length distribution $\{y_n, n \ge 0\}$, we obtain the following performance measures :

(a) The loss probability (P_{loss}) :

$$P_{loss} = y_K$$

(b) The mean queue length :

$$M = \sum_{i=1}^{K} i y_i$$

(c) By Little's law, we obtain the mean waiting time in the system :

$$W = \frac{M}{\lambda(1-P_{loss})}.$$

4. Conclusion

In this paper, we analyzed the M/G/1/K queueing model with two-threshold hysteresis strategy of service intensity switching and suggested to use the embedded Markov chain and the supplementary variable method to obtain the queue length distribution of our model. Our method is much simpler than that of Zhernovyi and Zhernovyi [1] with respect to obtaining queue length distribution.

References

[1] Zhernovyi, K.Y. and Y.V. Zhernovyi, "An *M θ/G/1/m* system with two-threshold hysteresis strategy of service intensity switching," *Journal of Communications and Electronics*, Vol.12, No.2(2012), pp.127-140.

[2] Choi, D.I., T.S. Kim, and S. Lee, "Analysis of a queueing system with a general service scheduling function, with applications to telecommunication network traffic control," *European Journal of Operational Research*, Vol.178, No.2(2007), pp.463-471.