J. Appl. Math. & Informatics Vol. 31(2013), No. 1 - 2, pp. 147 - 154
Website: http://www.kcam.biz

ANALYSIS OF QUEUEING MODEL WITH PRIORITY SCHEDULING BY SUPPLEMENTARY VARIABLE METHOD

DOO IL CHOI

Abstract. We analyze queueing model with priority scheduling by supplementary variable method. Customers are classified into two types (type-1 and type-2) according to their characteristics. Customers of each type arrive by independent Poisson processes, and all customers regardless of type have same general service time. The service order of each type is determined by the queue length of type-1 buffer. If the queue length of type-1 customer exceeds a threshold L, the service priority is given to the type-1 customer. Otherwise, the service priority is given to type-2 customer. Method of supplementary variable by remaining service time gives us information for queue length of two buffers. That is, we derive the differential difference equations for our queueing system. We obtain joint probability generating function for two queue lengths and the remaining service time. Also, the mean queue length of each buffer is derived.

AMS Mathematics Subject Classification : 60K25, 68M20.
key words and phrases : queueing model, priority scheduling, joint probability generating function, differential difference equations.

1. Introduction

Queueing system is a mathematical model to characterize a waiting-line situation where the customers arrive at the facility, join the queue if it is not immediately served and leave the facility after being served. Here, the arrivals and the service of customers occur randomly. These queueing systems arise in a wide variety of applications such as computer systems, telecommunication networks including home networks and ubiquitous system. Many customers (traffics) in these systems require the differentiated Quality of Service (QoS). Thus, when the customers with different service requirements may be accommodated in the system, these customers need to be handled differently [1,2]. These situations finally must be modeled by a queueing system with priority [1].

Received July 3, 2012. Revised October 9, 2012. Accepted November 5, 2012.

Recently, with applications in telecommunication networks, there has been renewed the interest for queueing system with priority scheduling. Representatively, to support the delay-sensitive customer (real-time traffic such as voice) with stringent delay requirement, the Head of Line (HOL) priority (or nonpreemptive priority) scheduling scheme has been applied [2,3]. In HOL priority scheduling scheme, the delay-sensitive customer can satisfy their delay requirement sufficiently. On the other hand, the nondelay-sensitive customer (nonrealtime traffic such as data) may be suffering by more many delay. This finally may cause the information to be no use at the destination. To satisfy the Quality of Service (QoS) of loss and waiting of delay-sensitive and nondelay-sensitive traffic simultaneously, the dynamic priority models have been proposed and analyzed [4,5]. One representative model is just Queue length threshold (QLT) scheduling policy [6,7,8]. There also are overload control schemes to satisfy QoS of traffics and to prevent congestion of network [9,10].

In this paper, in order to support the customers with different service requirements, we classify customers into two types called type-1 customers and type-2 customers. There are two buffers with infinite capacities to accommodate each type customers. Arrivals of type-1 and type-2 customers are assumed to be Poisson processes with rates λ_1 and λ_2, respectively. The type-1 and type-2 customers are served by a single server with general service time. That is, the service times (S) of customers regardless of type are independent and identically distributed with probability density function $b(\cdot)$. Let $\hat{b}(s) \triangleq \int_0^\infty e^{-st}b(t)dt$ be the Laplace transform of the service time S. The service of customers in each buffer is based on the first-come first-service (FCFS). The service priority of each buffer is determined by queue length of the type-1 buffer. Concretely, we place one threshold L on type-1 buffer. If the queue length for type-1 customers is less than or equal to the threshold L, the type-2 customers are served. Otherwise, the type-1 customers are served. If one of the buffers is empty, the customers in other buffer are served. Here, the type-1 customers can be considered as the nondelay-sensitive traffic and the type-2 customers can be considered as the delay-sensitive traffic.

This is the same model with those of Knessl, Choi and Tier [5]. However, they assumed the service time of customers to be the exponential distribution. It has many restriction applying to general situation. We analyze the queueing system with general service time. By using the method of supplementary variable by remaining service time, we derive the differential difference equations. We obtain joint probability generating function for two queue length and the remaining service time. From these results, the queue lengths of each buffers also can be derived.

2. Analysis

Let $N_1(t)$ and $N_2(t)$ be the number of type-1 and type-2 customers in each buffer at time t, respectively. We analyze our model with remaining service time as supplementary variable. Let $X(t)$ be the remaining service time of the customer currently in service. We also introduce

$$\xi(t) = \begin{cases} 1, & \text{if the server is busy at time } t, \\ 0, & \text{if the server is idle at time } t. \end{cases}$$

We assume $(\lambda_1 + \lambda_2)E(S) < 1$ and consider only stable system.
Define

$$p_{m,n}(x)dx = \lim_{t\to\infty} \Pr\{N_1(t) = m, N_2(t) = n, X(t) \in (x, x + dx), \xi(t) = 1\}, \quad m, n \geq 0$$
$$p_0 = \lim_{t\to\infty} \Pr\{\xi(t) = 0\}.$$

We then have the following ballance equation and the system of differential difference equations:

$$(\lambda_1 + \lambda_2)p_0 = p_{0,0}(0). \tag{1.1}$$

For $m \geq L + 1$, $n \geq 1$,

$$\begin{aligned} -\frac{dp_{m,n}(x)}{dx} =& -(\lambda_1 + \lambda_2)p_{m,n}(x) + \lambda_1 p_{m-1,n}(x) \\ &+ \lambda_2 p_{m,n-1}(x) + b(x)p_{m+1,n}(0). \end{aligned} \tag{1.2}$$

For $m \geq L + 1$, $n = 0$,

$$-\frac{dp_{m,0}(x)}{dx} = -(\lambda_1 + \lambda_2)p_{m,0}(x) + \lambda_1 p_{m-1,0}(x) + b(x)p_{m+1,0}(0). \tag{1.3}$$

For $m = L$, $n \geq 1$,

$$\begin{aligned} -\frac{dp_{L,n}(x)}{dx} =& -(\lambda_1 + \lambda_2)p_{L,n}(x) + \lambda_1 p_{L-1,n}(x) + \lambda_2 p_{L,n-1}(x) \\ &+ b(x)p_{L+1,n}(0) + b(x)p_{L,n+1}(0). \end{aligned} \tag{1.4}$$

For $m = L$, $n = 0$,

$$\begin{aligned} -\frac{dp_{L,0}(x)}{dx} =& -(\lambda_1 + \lambda_2)p_{L,0}(x) + \lambda_1 p_{L-1,0}(x) \\ &+ b(x)p_{L+1,0}(0) + b(x)p_{L,1}(0). \end{aligned} \tag{1.5}$$

For $0 < m < L$, $n \geq 1$,

$$-\frac{dp_{m,n}(x)}{dx} = -(\lambda_1 + \lambda_2)p_{m,n}(x) + \lambda_1 p_{m-1,n}(x)$$

$$+ \lambda_2 p_{m,n-1}(x) + b(x) p_{m,n+1}(0). \tag{1.6}$$

For $0 < m < L$, $n = 0$,

$$\begin{aligned} -\frac{dp_{m,0}(x)}{dx} =& - (\lambda_1 + \lambda_2) p_{m,0}(x) + \lambda_1 p_{m-1,0}(x) \\ &+ b(x) p_{m+1,0}(0) + b(x) p_{m,1}(0). \end{aligned} \tag{1.7}$$

For $m = 0$, $n \geq 1$,

$$\begin{aligned} -\frac{dp_{0,n}(x)}{dx} =& - (\lambda_1 + \lambda_2) p_{0,n}(x) \\ &+ \lambda_2 p_{0,n-1}(x) + b(x) p_{0,n+1}(0). \end{aligned} \tag{1.8}$$

For $m = 0$, $n = 0$,

$$\begin{aligned} -\frac{dp_{0,0}(x)}{dx} =& - (\lambda_1 + \lambda_2) p_{0,0}(x) + b(x) p_{1,0}(0) \\ &+ b(x) p_{0,1}(0) + (\lambda_1 + \lambda_2) b(x) p_0. \end{aligned} \tag{1.9}$$

and the normalization condition

$$\sum_{m,n=0}^{\infty} \int_0^{\infty} p_{m,n}(x) dx + p_0 = 1. \tag{1.10}$$

Introduce the following probability generating functions:

$$\begin{aligned} G(x, z, w) &\triangleq \sum_{n=0}^{\infty} \sum_{m=L}^{\infty} p_{m,n}(x) z^{m-L} w^n, \\ H_j(x, w) &\triangleq \sum_{n=0}^{\infty} p_{j,n}(x) w^n, \qquad 0 \leq j \leq L. \end{aligned}$$

From (1.2), (1.3), (1.4) and (1.5), we obtain

$$\begin{aligned} -\frac{\partial G(x,z,w)}{\partial x} =& - (\lambda_1 + \lambda_2) G(x,z,w) + \lambda_1 z G(x,z,w) + \lambda_1 H_{L-1}(x,w) \\ &+ \lambda_2 w G(x,z,w) + \frac{b(x)}{z} [G(0,z,w) - H_L(0,w)] \\ &+ \frac{b(x)}{w} [H_L(0,w) - p_{L,0}(0)]. \end{aligned}$$

The above equation can be rewritten into the following linear differential equation:

$$\frac{\partial G(x,z,w)}{\partial x} + \{\lambda_1 z + \lambda_2 w - (\lambda_1 + \lambda_2)\} G(x,z,w)$$

$$= \left[(\frac{1}{z} - \frac{1}{w}) H_L(0,w) - \frac{1}{z} G(0,z,w) + \frac{p_{L,0}(0)}{w} \right] b(x) \\ - \lambda_1 H_{L-1}(x,w). \tag{2.1}$$

The solution of equation (2.1) is given by

$$\begin{aligned} G(x,z,w) = & \left[(\frac{1}{w} - \frac{1}{z}) H_L(0,w) + \frac{G(0,z,w)}{z} - \frac{p_{L,0}(0)}{w} \right] \\ & \times \int_x^\infty e^{\{\lambda_1 z + \lambda_2 w - (\lambda_1+\lambda_2)\}(t-x)} b(t) dt \\ & + \lambda_1 \int_x^\infty e^{\{\lambda_1 z + \lambda_2 w - (\lambda_1+\lambda_2)\}(t-x)} H_{L-1}(t,w) dt. \end{aligned} \tag{2.2}$$

The unknown function $G(0,z,w)$ is found by setting $x = 0$ in the equation (2.2). After simple manipulation we obtain

$$\begin{aligned} & G(0,z,w) \\ & = \frac{1}{z - \hat{b}(\lambda_1 + \lambda_2 - \lambda_1 z - \lambda_2 w)} \left[\frac{z-w}{w} \hat{b}(\lambda_1 + \lambda_2 - \lambda_1 z - \lambda_2 w) H_L(0,w) \right. \\ & \quad - \frac{z}{w} \hat{b}(\lambda_1 + \lambda_2 - \lambda_1 z - \lambda_2 w) p_{L,0}(0) \\ & \quad \left. + \lambda_1 z \int_0^\infty e^{-\{\lambda_1 + \lambda_2 - \lambda_1 z - \lambda_2 w\} t} H_{L-1}(t,w) dt \right]. \end{aligned} \tag{2.3}$$

Let $z = z_1(w)$ be the solution of the equation $z = \hat{b}(\lambda_1 + \lambda_2 - \lambda_1 z - \lambda_2 w)$ within the unit circle $|z| = 1$, where $|w| \leq 1$. Then, we obtain

$$\begin{aligned} H_L(0,w) = & \frac{z_1(w)}{(z_1(w) - w)} p_{L,0}(0) \\ & + \frac{\lambda_1 w}{(z_1(w) - w)} \int_0^\infty e^{-\{\lambda_1 + \lambda_2 - \lambda_1 z_1(w) - \lambda_2 w\} t} H_{L-1}(t,w) dt. \end{aligned} \tag{2.4}$$

Therefore, we must know $p_{L,0}(0)$ and $H_{L-1}(t,w)$ to obtain $H_L(0,w)$ and finally $G(x,z,w)$. From (1.6), (1.7), (1.8) and (1.9), we also obtain the following linear differential equations:

$$\begin{aligned} & \frac{\partial H_m(x,w)}{\partial x} + \{\lambda_2 w - (\lambda_1 + \lambda_2)\} H_m(x,w) \\ & \quad = -\lambda_1 H_{m-1}(x,w) + b(x) \left[\frac{1}{w} (p_{m,0}(0) - H_m(0,w)) - p_{m+1,0}(0) \right], \\ & \qquad 0 < m < L. \end{aligned} \tag{2.5}$$

$$\frac{\partial H_0(x,w)}{\partial x} + \{\lambda_2 w - (\lambda_1 + \lambda_2)\} H_0(x,w)$$

$$= b(x)\left[\frac{1}{w}(p_{0,0}(0) - H_0(0,w)) - p_{1,0}(0) - (\lambda_1 + \lambda_2)p_0\right]. \tag{2.6}$$

The solutions of the linear differential equations (2.5) and (2.6) are given by

$$\begin{aligned} H_m(x,w) = &\left[\frac{1}{w}(H_m(0,w) - p_{m,0}(0)) + p_{m+1,0}(0)\right] \\ &\times \int_x^\infty e^{\{\lambda_2 w - (\lambda_1+\lambda_2)\}(t-x)} b(t)dt \\ &+ \lambda_1 \int_x^\infty e^{\{\lambda_2 w - (\lambda_1+\lambda_2)\}(t-x)} H_{m-1}(t,w)dt, \quad 0 < m < L. \end{aligned} \tag{2.7}$$

$$\begin{aligned} H_0(x,w) = &[p_{1,0}(0) + (\lambda_1 + \lambda_2)p_0 \\ &- \frac{1}{w}(p_{0,0}(0) - H_0(0,w))\Big] \int_x^\infty e^{\{\lambda_2 w - (\lambda_1+\lambda_2)\}(t-x)} b(t)dt. \end{aligned} \tag{2.8}$$

Setting $x = 0$ in the equations (2.7) and (2.8), we obtain

$$\begin{aligned} H_m(0,w) = &\frac{w}{w - \hat{b}(\lambda_1 + \lambda_2 - \lambda_2 w)}\left[\left(p_{m+1,0}(0) - \frac{p_{m,0}(0)}{w}\right)\hat{b}(\lambda_1 + \lambda_2 - \lambda_2 w)\right. \\ &\left. + \lambda_1 \int_0^\infty e^{\{\lambda_2 w - (\lambda_1+\lambda_2)\}t} H_{m-1}(t,w)dt\right], \qquad 0 < m < L. \end{aligned} \tag{2.9}$$

$$\begin{aligned} H_0(0,w) = &\frac{w}{w - \hat{b}(\lambda_1 + \lambda_2 - \lambda_2 w)}\Big[p_{1,0}(0) \\ &+ (\lambda_1 + \lambda_2)p_0 - \frac{p_{0,0}(0)}{w}\Big]\hat{b}(\lambda_1 + \lambda_2 - \lambda_2 w). \end{aligned} \tag{2.10}$$

Let w^* be the solution of the equation $w = \hat{b}(\lambda_1 + \lambda_2 - \lambda_2 w)$. That is, it can easily be proved by Rouche's theorem that $w = \hat{b}(\lambda_1 + \lambda_2 - \lambda_2 w)$ has an unique solution within an unit circle $|w| = 1$. Since $H_0(0,w)$ and $H_m(0,w)(0 < m < L)$ are analytic within and on unit circle $|w| = 1$, the denominator of the right side of equations (2.9) and (2.10) must be zero. Thus, we obtain

$$p_{m+1,0}(0) = \frac{p_{m,0}(0)}{w^*} - \frac{\lambda_1}{w^*}\int_0^\infty e^{\{\lambda_2 w^* - (\lambda_1+\lambda_2)\}t} H_{m-1}(t,w^*)dt. \tag{3.1}$$

$$p_{1,0}(0) = \frac{p_{0,0}(0)}{w^*} - (\lambda_1 + \lambda_2)p_0. \tag{3.2}$$

By (1.1), the equation (3.2) can be rewritten by

$$p_{1,0}(0) = (\lambda_1 + \lambda_2)p_0\left(\frac{1}{w^*} - 1\right). \tag{3.3}$$

Also, substituting (1.1) and (3.3) into equation (2.10), we obtain

$$H_0(0,w) = \frac{(w-w^*)\hat{b}(\lambda_1+\lambda_2-\lambda_2 w)}{w^*[w-\hat{b}(\lambda_1+\lambda_2-\lambda_2 w)]}(\lambda_1+\lambda_2)p_0. \tag{3.4}$$

In equation (3.1), with $m = 1$, we obtain $p_{2,0}(0)$ by equation (2.8) and (3.3). By applying the equations (2.7), (2.8), (2.9) and (2.10) repeatedly, we finally can obtain $p_{L,0}(0)$ and $H_{L-1}(t,w)$. Therefore, the joint probability generating function $G(x,z,w)$ is obtained. The probability of empty system p_0 also is determined by the normalization condition (1.10).

From above results, we can also obtain the mean queue length for each type customers:

(a) The mean queue length ($E(N_1)$) of type-1 customers

$$E(N_1) = \frac{\partial}{\partial z}\left[\int_0^\infty G(x,z,1)dx \cdot z^L\right]\Bigg|_{z=1} + \sum_{j=1}^{L-1} j \int_0^\infty H_j(x,1)dx.$$

(b) The mean queue length ($E(N_2)$) of type-2 customers

$$E(N_2) = \frac{\partial}{\partial w}\int_0^\infty G(x,1,w)dx\Bigg|_{w=1}.$$

3. Conclusion

In this paper, we considered the queueing model with priority scheduling. Instead of the method of embedded Markov chain, by using the method of supplementary variable by remaining service time, we obtained the probability generating function for two queue lengths and the remaining service time explicitly. The motivation for analyzing this queueing model was necessary of traffic control to support differentiated traffic streams in telecommunication networks. Thus, we expect the analysis result to be used for optimizing of the network with appropriate threshold value.

References

1. N. K. Jaiswal, *priority queues*, Academic Press , New York, (1968).
2. R. Chipalkatti, J. F. Kurose, and D. Towsley, *Scheduling policies for real-time and nonreal-time traffic in a statistical multiplexer*, Proceedings of the IEEE INFOCOM'89, Ottawa (1989), 774-783.
3. A. Sugahara, T. Takine, Y. Takahashi, T. Hasegawa, *Analysis of a nonpreemptive priority queue with SPP arrivals of high class*, Perf. Eval. 21 (1995), 215-238.
4. D. I. Choi and Y. Lee, *Performance analysis of a dynamic priority queue for traffic control of bursty traffic in ATM networks*, IEE Proc. - Commun. 148 (2001), No. 3 181-187.
5. C. Knessl, D. I. Choi and C. Tier, *A dynamic priority queue model for simultaneous service of two traffic types*, SIAM J. APPL. MATH. 63(2002), No. 2 398-422.
6. D. I. Choi, B. D. Choi and D. K. Sung, *Performance analysis of priority leaky bucket scheme with queue-length-threshold scheduling policy*, IEE Proc.- Comm. 145 (1998), No. 6 395-401.

7. B. D. Choi, D. I. Choi, *An analysis of M,MMPP/G/1 queue with QLT scheduling policy and Bernoulli schedule*, IEICE Tran. on Comm. E81-B (1998), No. 1 13-22.
8. D. S. Lee and B. Sengupta, *Queueing analysis of a threshold based priority scheme for ATM networks*, IEEE/ACM Tran. on Networking 1(1993), No. 6 709-717.
9. D. I. Choi, *Transient analysis of a queueing system with Markov-modulated Bernoulli arrivals and overload control*, Journal of Applied Mathematics and Computing 15(2004), No.1-2 405-414.
10. D. I. Choi, *Analysis of a queueing system with overload control by arrival rates* , Journal of Applied Mathematics and Computing 18(2005), No.1-2 455-464.

Doo Il Choi received Ph. D at Department of Mathematics KAIST. Since 1998 he has been at Halla University and his research field is queueing analysis and its applications to telecommunication networks.

Department of Applied Mathematics, Halla University, Heungup-li, Heungup-myon, Wonju-shi, Kangwon-do, 220-712, Korea.
e-mail: dichoi@halla.ac.kr