Nonlinear Dynamics and Systems Theory, **6**(3) (2006) 295–308

The Matrix-Geometric Solution of the M/E$_k$/1 Queue with Balking and State-Dependent Service

Dequan Yue[1], Chunyan Li[1] and Wuyi Yue[2]

[1] *Department of Statistics, College of Sciences, Yanshan University Qinhuangdao 066004, China*
[2] *Department of Information Science and Systems Engineering Konan University, Kobe 658-8501, Japan*

Received: August 26, 2005; Revised: January 10, 2006

Abstract: In this paper, we present an analysis for an M/E$_k$/1 queuing system with balking and state-dependent service. Customers are served with two different rates depending on the number of customers in the system. If a customer on arrival finds other customers in the system, it either decides to enter the queue or balks with a constant probability. We first formulate the queuing model as a quasi-birth and death (QBD) process. Then, we obtain the equilibrium condition of the system. By using the matrix geometric solution method, we obtain the explicit expressions for steady-state probability vector via the rate matrix $\boldsymbol{R}$. The computation of the rate matrix $\boldsymbol{R}$ is also discussed. Then, we derive explicitly some performance measures of the system. Based on these performance analysis, we develop a cost model to determine numerically the optimal cost and optimal critical value. Finally, we perform sensitivity analysis through numerical experiments.

Keywords: *Balking; state-dependent; matrix geometric solution; steady-state probability.*

Mathematics Subject Classification (2000): 60K25, 68M20.

1 Introduction

We consider an M/E$_k$/1 queuing system with balking and state-dependent service. Customers are served with two different service rates depending on the number of customers in the system. If a customer on arrival finds other customers in the system, it either decides to enter the queue or balk (does not enter) with a constant probability. Balking is not only a common phenomenon in queues arising in daily activities, but also in communication systems, production line systems and in various machine interferences or repair models (see [1]-[4] and references therein).

The queuing systems with balking, or reneging, or both have been studied by many researchers. Haight [5] is the first person who considered an M/M/1 queue with balking.

An M/M/1 queue with customer reneging was also proposed by Haight [6]. The combined effects of balking and reneging in an M/M/1 queue with limited waiting room and with unlimited waiting room have been investigated by Ancker and Gafarian [7], [8]. They obtained the steady-state probabilities and some performance measures of the system such as the mean number in the queue, the mean number in the system and the mean rate of customer loss.

Abou-EI-Ata [9] extended the model in [7] to study the state-dependent M/M/1/N queue with reneging and a general balk function, where the server has two service rates depending on the number of customers in the system. Some of its variations have been studied by several authors including, for example, Abou-EI-Ata and Kotb [10], Abou-EI-Ata et al. [11] and Abou-EI-Ata and Shawky [12].

Recently, Drekic and Woolford [13] studied a preemptive priority Markovian queue with state-dependent service and lower priority balking customers. They formulated the queueing model as a quasi-birth and death (QBD) process. By using the method of generalized eigenvalues, they established an explicit representation for the so-called rate matrix. They also obtained the steady-state joint distribution of the number of high and low priority customers in the system.

The state-dependent M/M/1 queue with balking was studied by Al-seedy and Kotb [14]. They obtained the transient solution of the state probabilities. Al-seedy [15] extended the model proposed by Abou-EI-Ata [9] to the state-dependent M/E$_k$/1/N queue with balking. By solving the steady-state probability-difference equations, Al-seedy [15] obtained some iterative expressions of the steady-state probabilities. However, these iterative expressions are too complex to obtain explicit expressions of the steady-state probabilities in general cases, and they could not derive explicitly some performance measures such as the distribution of the queue length and the expected number of customers in the system and in the queue. Even for a special case when the waiting room is unlimited (i.e., $N \to \infty$), it is difficult to obtain the explicit expressions of the steady-state probabilities from the iterative expressions.

In this paper, we study a state-dependent M/E$_k$/1 queue with balking and an unlimited waiting room. The rest of the paper is organized as follows. In Section 2, we formulate the queuing model as a QBD process and obtain the equilibrium condition of the system. In Section 3, by using a matrix-geometric solution method, we derive the explicit expression for steady-state probability vector. Also, we derive explicitly some performance measures of the system such as the expected number of the customers in the system and in the queue and the mean balking rate of the system. Based on these analyses, we develop a cost model to determine numerically the optimal cost and optimal critical value. In Section 4, we perform sensitivity analysis through numerical experiments. Conclusions are given in Section 5.

2 System Model and Equilibrium Condition

In this section, we first describe the system model. Then, we derive an infinitesimal generator of a QBD process of the system. Finally, we provide an equilibrium condition of the system.

2.1 Model assumptions

In this paper, we consider an M/E$_k$/1 queuing system with balking and state-dependent service rate. The assumptions of the system model are as follows:

(a) There is only one server in the system, and the server can only serve one customer at the same time. The capacity of the system is infinite. It is assumed that the service is independent of the arrival of the customers.
(b) Customers arrive at the system one by one according to a Poisson process with rate λ ($\lambda > 0$).
(c) A customer on arrival decides to join the queue or balk. If a customer on arrival finds some customers in the system, then it joins the queue with probability β and balks with probability $1-\beta$. If a customer on arrival finds no customer in the system, then he joins the system and will be serviced immediately.
(d) The customers are served on a first-come, first served (FCFS) discipline. Once service commences it always proceeds to completion. The service times are assumed to be distributed according to an Erlang distribution with mean k/μ_n and stage parameter k. The Erlang type k distribution is made up of k independent and identical exponential stages, each with mean $1/\mu_n$, given by

$$\mu_n = \begin{cases} k\mu_1, & n = 1, 2, ..., r, \\ k\mu_2, & n = r+1, r+2, \dots . \end{cases}$$

This means that the server has two rates say called "slow and fast" depending on the number of customers n in the system. When the number of customers n in the system is less than or equal to the critical value r, the server has slow service rate μ_1; otherwise, the server has fast service rate μ_2 $(0 < \mu_1 < \mu_2)$.

2.2 Infinitesimal generator of a QBD process

Let $N(t)$ denote the number of the customers in the system at time t, and $J(t)$ denote the service stage that the customer being served at time t $(t \geq 0)$. A customer goes into the first stage of the service (say stage k), then progresses through the remaining stages and must complete the last stage (say stage 1). The state space of the two dimensional process $\{(N(t), J(t)); t \geq 0\}$ is given by

$$S = \{(i, j);\ i = 0, 1, ..., j = 1, 2, ..., k\}.$$

All states of this two dimensional process are labelled in the lexicographic order as follows:

$$(0, 0); (1, 1), (1, 2), ..., (1, k); (2, 1), (2, 2), ..., (2, k); \dots .$$

By the probability analysis, we have the following infinitesimal generator of the process $\{(N(t), J(t)); t \geq 0\}$.

$$\boldsymbol{Q} = \left(\begin{array}{cccccccc} B_0 & \boldsymbol{C}_0 & & & & & & \\ \boldsymbol{A}_1 & \boldsymbol{B}_1 & \boldsymbol{C}_1 & & & & & \\ & \boldsymbol{A}_2 & \boldsymbol{B}_1 & \boldsymbol{C}_1 & & & & \\ & & \cdots & \cdots & \cdots & & & \\ & & & \boldsymbol{A}_2 & \boldsymbol{B}_1 & \boldsymbol{C}_1 & & \\ & & & & \boldsymbol{A}_3 & \boldsymbol{B}_2 & \boldsymbol{C}_1 & \\ & & & & & \cdots & \cdots & \cdots \end{array}\right) \begin{array}{ll} \cdots & 0 \\ \cdots & 1 \\ \cdots & 2 \\ & \vdots \\ \cdots & r \\ \cdots & r+1 \\ & \vdots \end{array}$$

where

$$\boldsymbol{A}_1 = \begin{pmatrix} k\mu_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \boldsymbol{A}_2 = \begin{pmatrix} 0 & \cdots & 0 & k\mu_1 \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}, \quad \boldsymbol{A}_3 = \begin{pmatrix} 0 & \cdots & 0 & k\mu_2 \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix},$$

$$B_0 = -\lambda, \quad \boldsymbol{B}_1 = \begin{pmatrix} -\beta\lambda - k\mu_1 & & & \\ k\mu_1 & -\beta\lambda - k\mu_1 & & \\ & \cdots & \cdots & \\ & & k\mu_1 & -\beta\lambda - k\mu_1 \end{pmatrix},$$

$$\boldsymbol{B}_2 = \begin{pmatrix} -\beta\lambda - k\mu_2 & & & \\ k\mu_2 & -\beta\lambda - k\mu_2 & & \\ & \cdots & \cdots & \\ & & k\mu_2 & -\beta\lambda - k\mu_2 \end{pmatrix},$$

$$\boldsymbol{C}_0 = \begin{pmatrix} 0 & \cdots & 0 & \lambda \end{pmatrix}, \quad \boldsymbol{C}_1 = \begin{pmatrix} \beta\lambda & & & \\ & \beta\lambda & & \\ & & \cdots & \\ & & & \beta\lambda \end{pmatrix},$$

where $\boldsymbol{C}_0$ is a matrix of order $1 \times k$, $\boldsymbol{A}_1$ is a matrix of order $k \times 1$, and other matrixes are square matrixes of order k.

From the book written by Neuts [16], we know that process $\{N(t), J(t); t \geq 0\}$ is a QBD process.

2.3 Equilibrium condition of the system

In the following, we provide a necessary and sufficient condition to ensure the existence for the stationary probability distribution of the process $\{N(t), J(t); t \geq 0\}$.

Let $\boldsymbol{H} = \boldsymbol{A}_3 + \boldsymbol{B}_2 + \boldsymbol{C}_1$, then $\boldsymbol{H}$ is given by

$$\boldsymbol{H} = \begin{pmatrix} -k\mu_2 & 0 & \cdots & 0 & k\mu_2 \\ k\mu_2 & -k\mu_2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & k\mu_2 & -k\mu_2 \end{pmatrix}.$$

It is readily known that $\boldsymbol{H}$ is an irreducible generator. Let $\boldsymbol{\pi} = (\pi_1, \pi_2, ..., \pi_k)$ be the steady-state probability vector of $\boldsymbol{H}$. Then, $\boldsymbol{\pi}$ satisfies the linear equations $\boldsymbol{\pi H} = 0$ and $\boldsymbol{\pi e} = 1$, where $\boldsymbol{e}$ is a column vector whose elements are all equal to 1. Solving the above linear equations, we get that

$$\pi_i = \frac{1}{k}, \quad i = 1, 2, ..., k. \tag{1}$$

By Theorem 3.1.1 in Chapter 3 of Neuts [16], the equilibrium condition of the system is given by

$$\boldsymbol{\pi A}_3 \boldsymbol{e} > \boldsymbol{\pi C}_1 \boldsymbol{e}.$$

Making substitution for $\boldsymbol{\pi}$ by Eq. (1), then we have the equilibrium condition of the system given by

$$\frac{\beta\lambda}{\mu_2} < 1. \tag{2}$$

Remark 2.1 We observe from the above condition that the equilibrium condition of the system is dependent with the fast service rate μ_2 and independent with the slow service rate μ_1. This is in agreement with the equilibrium condition obtained by Rao [17], where Rao considered an M/G/1 queueing system in which customers balk with a constant probability $1-\beta$ and renege according to a negative exponential distribution. It has been shown that as long as reneging is permitted, the steady states always exist, but when no reneging is permitted, the steady states exist only when $\lambda\beta\eta < 1$, where λ is the arrival rate of customers, and η is the mean service time of a customer.

3 Performance Measures and Cost Model

In this section, we first derive the explicit expression for the steady-state probability vector. Then, we give some useful performance measures of the system. Based on these performance measures, we develop a cost model to determine the optimal critical value r to minimize the total expected cost per unit time.

3.1 Steady-state probability vector

Let $\boldsymbol{X} = (X_0, \boldsymbol{X}_1, ..., \boldsymbol{X}_r, \boldsymbol{X}_{r+1}, ...)$, where X_0 is a number, $\boldsymbol{X}_i$ $(i = 1, 2, ...)$ is a vector of order k. By applying the matrix geometric solution method [16], the stationary probability vector is given by

$$\boldsymbol{X}_i = \boldsymbol{X}_r \boldsymbol{R}^{i-r}, \qquad i = r,\ r+1,\ ... \tag{3}$$

where $\boldsymbol{R}$ is the minimal nonnegative solution to the equation $\boldsymbol{R}^2\boldsymbol{A}_3 + \boldsymbol{R}\boldsymbol{B}_2 + \boldsymbol{C}_1 = 0$, and $X_0, \boldsymbol{X}_1, ..., \boldsymbol{X}_r$ are given by solving the following equations:

$$\begin{aligned}
X_0 B_0 + \boldsymbol{X}_1\boldsymbol{A}_1 &= 0, \\
X_0\boldsymbol{C}_0 + \boldsymbol{X}_1\boldsymbol{B}_1 + \boldsymbol{X}_2\boldsymbol{A}_2 &= \boldsymbol{0}, \\
\boldsymbol{X}_i\boldsymbol{C}_1 + \boldsymbol{X}_{i+1}\boldsymbol{B}_1 + \boldsymbol{X}_{i+2}\boldsymbol{A}_2 &= \boldsymbol{0}, \qquad i = 1, 2, ..., r-2, \\
\boldsymbol{X}_{r-1}\boldsymbol{C}_1 + \boldsymbol{X}_r(\boldsymbol{B}_1 + \boldsymbol{R}\boldsymbol{A}_3) &= \boldsymbol{0}, \\
X_0 + \sum_{i=1}^{r-1} \boldsymbol{X}_i\boldsymbol{e} + \boldsymbol{X}_r(\boldsymbol{I}-\boldsymbol{R})^{-1}\boldsymbol{e} &= 1,
\end{aligned} \tag{4}$$

where $\boldsymbol{e}$ is a column vector of order k, and all its elements equal to 1.

In general, it is difficult to give an exact expression of $\boldsymbol{R}$ except for a few special cases. However, the matrix $\boldsymbol{R}$ can be approximately calculated by the following iterative procedure:

(a) $\boldsymbol{R}(0) = \boldsymbol{0}$,

(b) $\boldsymbol{R}(n+1) = -(\boldsymbol{C}_1 + \boldsymbol{R}^2(n)\boldsymbol{A}_3)\boldsymbol{B}_2^{-1}, \qquad n \geq 0.$

This iterative algorithm is convergent, i.e. $\boldsymbol{R} = \lim_{n\to\infty} \boldsymbol{R}(n)$ (Section 1.9 of Chapter 1 of [16]).

Remark 3.1 The inverse of the matrix $\boldsymbol{B}_2^{-1}$ in the above algorithm exists, and can be explicitly given by

$$\boldsymbol{B}_2^{-1} = \frac{1}{a^k}\begin{pmatrix} a^{k-1} & 0 & 0 & \cdots & 0 \\ a^{k-2}(-b) & a^{k-1} & 0 & \cdots & 0 \\ a^{k-3}(-b)^2 & a^{k-2}(-b) & a^{k-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ (-b)^{k-1} & a(-b)^{k-2} & a^2(-b)^{k-3} & \cdots & a^{k-1} \end{pmatrix},$$

where $a = (-\beta\lambda - k\mu_2)$, $b = k\mu_2$.

For a special case of $k = 2$ and $r = 2$, we can readily obtain an explicit expression for matrix $\boldsymbol{R}$ given in the following theorem.

Theorem 3.1 *If $k = 2$ and $r = 2$, then the matrix $\boldsymbol{R}$ is explicitly given by*

$$\boldsymbol{R} = \frac{\beta\lambda}{4\mu_2^2}\begin{pmatrix} 2\mu_2 & \beta\lambda \\ 2\mu_2 & \beta\lambda + 2\mu_2 \end{pmatrix}. \tag{5}$$

Proof Let

$$\boldsymbol{R} = \begin{pmatrix} R_{00} & R_{01} \\ R_{10} & R_{11} \end{pmatrix},$$

then the equation $\boldsymbol{R}^2\boldsymbol{A}_3 + \boldsymbol{R}\boldsymbol{B}_2 + \boldsymbol{C}_1 = 0$ can be written as follows:

$$\begin{aligned} 2\mu_2(R_{00}^2 + R_{01}R_{10}) + (-\beta\lambda - 2\mu_2)R_{01} &= 0, \\ 2\mu_2(R_{00}R_{10} + R_{11}R_{10}) + (-\beta\lambda - 2\mu_2)R_{11} + \beta\lambda &= 0, \\ (-\beta\lambda - 2\mu_2)R_{00} + 2\mu_2 R_{01} + \beta\lambda &= 0, \\ (-\beta\lambda - 2\mu_2)R_{10} + 2\mu_2 R_{11} &= 0. \end{aligned} \tag{6}$$

Noting that $\boldsymbol{R}\boldsymbol{A}_3\boldsymbol{e} = \boldsymbol{C}_1\boldsymbol{e}$, (Eq. (3.1.6) in [16]), we obtain that

$$R_{00} = \frac{\beta\lambda}{2\mu_2}, \qquad R_{10} = \frac{\beta\lambda}{2\mu_2}. \tag{7}$$

Substituting Eq. (7) into Eq. (6), the other two elements R_{01} and R_{11} are readily obtained. Then, we obtained the matrix $\boldsymbol{R}$. □

In this special case, the vector $\boldsymbol{X}$ is given by

$$\boldsymbol{X}_i = \boldsymbol{X}_2\boldsymbol{R}^{i-2}, \qquad i = 2, 3, ..., \tag{8}$$

and the number X_0, the vectors $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ satisfy the following equations:

$$\begin{aligned} X_0 B_0 + \boldsymbol{X}_1\boldsymbol{A}_1 &= 0, \\ X_0\boldsymbol{C}_0 + \boldsymbol{X}_1\boldsymbol{B}_1 + \boldsymbol{X}_2\boldsymbol{A}_2 &= \boldsymbol{0}, \\ \boldsymbol{X}_1\boldsymbol{C}_1 + \boldsymbol{X}_2(\boldsymbol{B}_1 + \boldsymbol{R}\boldsymbol{A}_3) &= \boldsymbol{0}, \\ X_0 + \boldsymbol{X}_1\boldsymbol{e} + \boldsymbol{X}_2(\boldsymbol{I} - \boldsymbol{R})^{-1}\boldsymbol{e} &= 1, \end{aligned} \tag{9}$$

where

$$(\boldsymbol{I} - \boldsymbol{R})^{-1} = \frac{1}{4\mu_2\beta\lambda - 4\mu_2^2}\begin{pmatrix} \beta^2\lambda^2 + 2\mu_2\beta\lambda - 4\mu_2^2 & \beta^2\lambda^2 \\ (-2\mu_2)\beta\lambda & (-2\mu_2)(2\mu_2 - \beta\lambda) \end{pmatrix}. \tag{10}$$

Thus, X_0, $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are readily obtained by solving Eq. (9). However, it is not simple to present it explicitly since the expressions are tediously long. Thus, their expressions are omitted.

3.2 Performance measures

Using the steady-state probability vector presented above, we can obtain some performance measures of the system. Let

$$\boldsymbol{X}_i = (x_{i1}, x_{i2}, ..., x_{ik}), \qquad i = 1, 2, ...,$$

then we have the following theorem.

Theorem 3.2 *(a) The expected number of customers in the system is given by*

$$E(N) = \sum_{n=1}^{r-1} n\boldsymbol{X}_n\boldsymbol{e} + \boldsymbol{X}_r\left[r(\boldsymbol{I}-\boldsymbol{R})^{-1} + \boldsymbol{R}(\boldsymbol{I}-\boldsymbol{R})^{-2}\right]\boldsymbol{e}. \tag{11}$$

(b) The expected number of customers in the queue is given by

$$E(N_q) = \sum_{n=1}^{r-2} n\boldsymbol{X}_{n+1}\boldsymbol{e} + \boldsymbol{X}_r\left[(r-1)(\boldsymbol{I}-\boldsymbol{R})^{-1} + \boldsymbol{R}(\boldsymbol{I}-\boldsymbol{R})^{-2}\right]\boldsymbol{e}. \tag{12}$$

(c) The mean balking rate of the system is given by

$$BR = (1-\beta)\lambda(1-X_0). \tag{13}$$

Proof The expected number of customers in the system is given by

$$E(N) = \sum_{i=1}^{k}\sum_{n=1}^{\infty} nx_{n,i} = \sum_{n=1}^{\infty} n\boldsymbol{X}_n\boldsymbol{e}.$$

From Eq. (3), we have

$$E(N) = \sum_{n=1}^{r-1} n\boldsymbol{X}_n\boldsymbol{e} + \sum_{n=r}^{\infty} n\boldsymbol{X}_r\boldsymbol{R}^{n-r}\boldsymbol{e}. \tag{14}$$

Hence, we obtain Eq. (11) by summation. Similarly, the expected number of customers in the queue is given by

$$E(N_q) = \sum_{i=1}^{k}\sum_{n=1}^{\infty} nx_{n+1,i} = \sum_{n=1}^{r-2} n\boldsymbol{X}_{n+1}\boldsymbol{e} + \sum_{n=r-1}^{\infty} n\boldsymbol{X}_r\boldsymbol{R}^{n-r+1}\boldsymbol{e}. \tag{15}$$

Hence, we obtain Eq. (12) by summation. Using the concept of Ancker and Gafarian [7], the mean balking rate of the system is given by

$$BR = \sum_{n=1}^{\infty} b_n\lambda\boldsymbol{X}_n\boldsymbol{e} = (1-\beta)\lambda(1-X_0). \tag{16}$$

Obviously, the probability that the server is busy is given by

$$P_B = 1 - X_0 \tag{17}$$

and the probability that the server is idle is given by

$$P_0 = X_0. \tag{18}$$

□

3.3 Cost model

In this subsection, we develop a steady-state expected cost function where the critical value r is a decision variable. Our objective is to determine the critical value r to minimize the total expected cost per unit time. Let

C_1=cost per unit time when customers are waiting for service,
C_2=cost per unit time when the server is busy,
C_3=loss cost per unit time when customers balk,
C_4=cost per unit time when the server is idle.

According to the definition of each cost of the parameters listed above, the total expected cost function per unit time is given by

$$F(r) = C_1 E(N_q) + C_2 P_B + C_3 BR + C_4 P_0, \tag{19}$$

where $E(N_q)$, BR, P_B, P_0 are given in Eqs. (12) and (13) and Eqs. (17) and (18). The first item of Eq. (19) is the cost incurred by the customer's waiting. The second and the last items of Eq. (19) are the costs incurred by the server. The third item of Eq. (19) is the cost incurred by the customer loss.

4 Sensitivity Analysis

In this section, we perform a sensitivity analysis on the optimal value r^* and its expected cost $F(r^*)$ based on changes in the values of the system parameters such as the arrival rate λ, the probability β, the slow service rate μ_1, the fast service rate μ_2 and cost parameters.

Let the service time follow a 2-stage Erlang distribution, and employ the cost parameters $C_1 = 100$, $C_2 = 150$, $C_3 = 300$ and $C_4 = 450$. The numerical results of the optimal critical value r^* and its expected minimum cost $F(r^*)$ are illustrated in Figures 4.1–4.4.

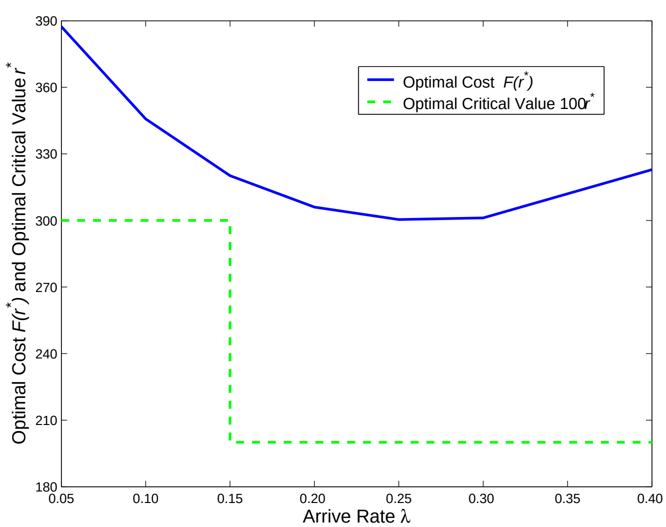

Figure 4.1: Optimal cost $F(r^*)$ and optimal critical value r^* versus arrive rate λ with $\mu_1 = 0.2$, $\mu_2 = 0.8$ and $\beta = 0.5$.

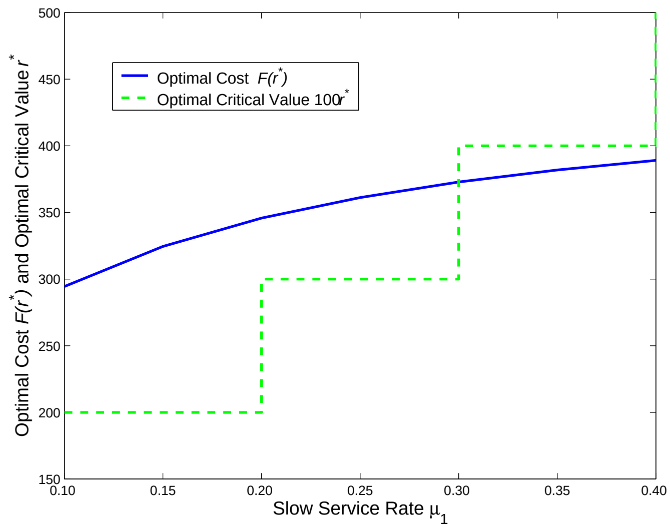

Figure 4.2: Optimal cost $F(r^*)$ and optimal critical value r^* versus slow service rate μ_1 with $\lambda = 0.1$, $\mu_2 = 0.8$ and $\beta = 0.5$.

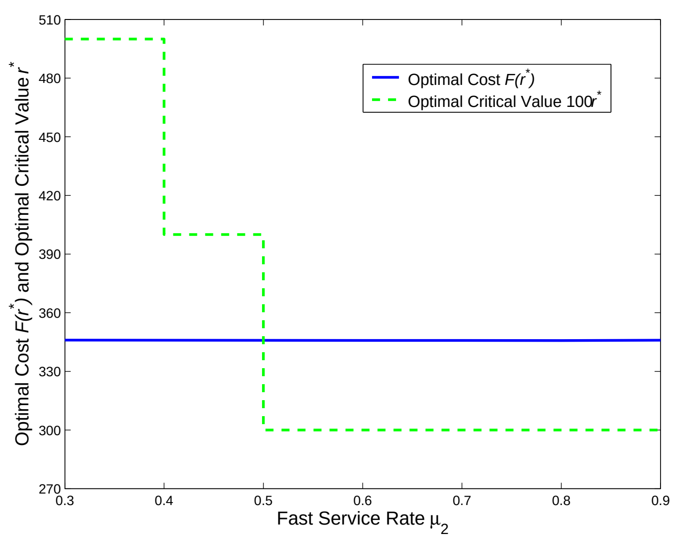

Figure 4.3: Optimal cost $F(r^*)$ and optimal critical value r^* versus fast service rate μ_2 with $\lambda = 0.1$, $\mu_1 = 0.2$ and $\beta = 0.5$.

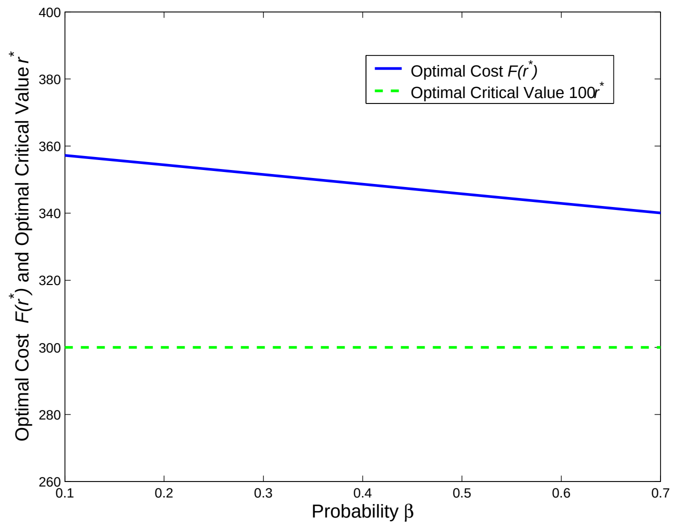

Figure 4.4: Optimal cost $F(r^*)$ and optimal critical value r^* versus probability β with $\lambda = 0.1$, $\mu_1 = 0.2$ and $\mu_2 = 0.8$.

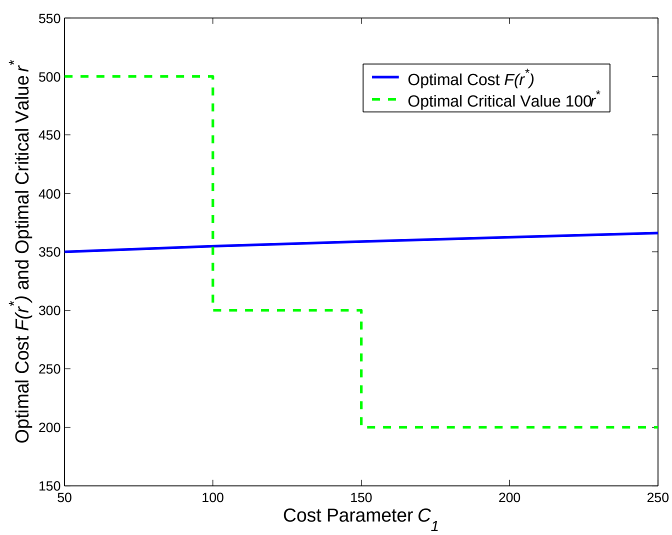

Figure 4.5: Optimal cost $F(r^*)$ and optimal critical value r^* versus cost parameter C_1 with $C_2 = 150$, $C_3 = 300$ and $C_4 = 450$.

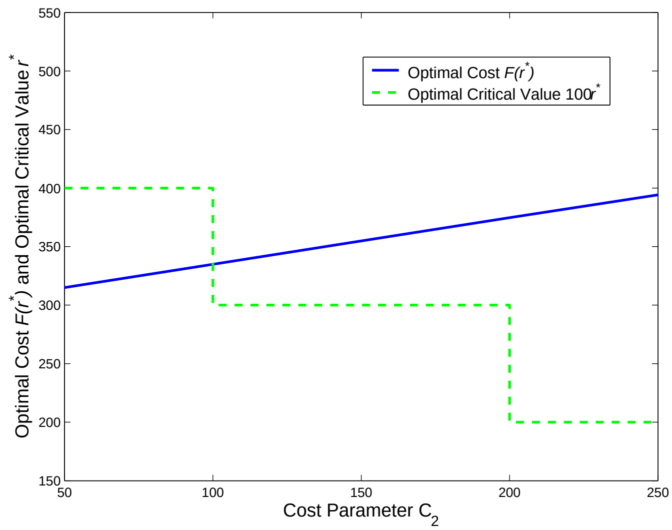

Figure 4.6: Optimal cost $F(r^*)$ and optimal critical value r^* versus cost parameter C_2 with $C_1 = 100$, $C_3 = 300$ and $C_4 = 450$.

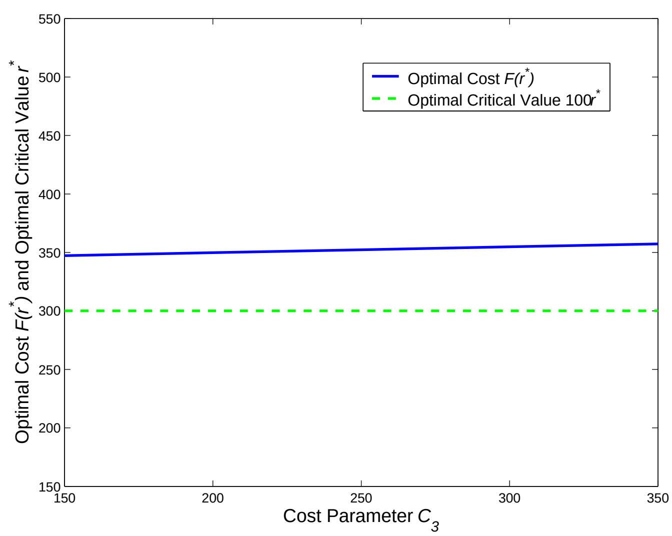

Figure 4.7: Optimal cost $F(r^*)$ and optimal critical value r^* versus cost parameter C_3 with $C_1 = 100$, $C_2 = 150$ and $C_4 = 450$.

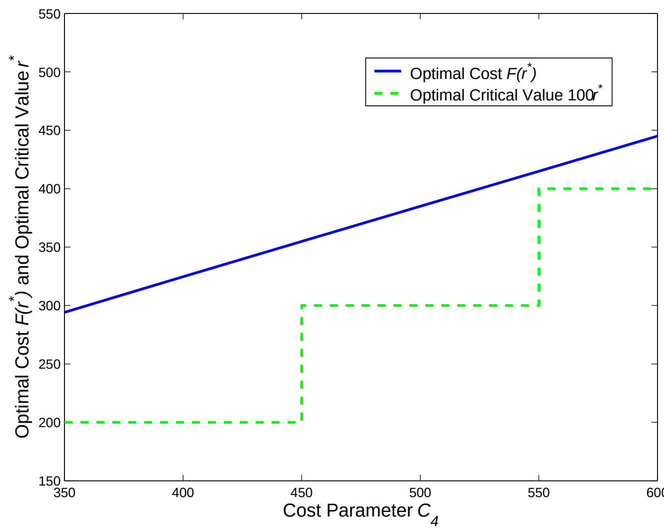

Figure 4.8: Optimal cost $F(r^*)$ and optimal critical value r^* versus cost parameter C_4 with $C_1 = 100$, $C_2 = 150$ and $C_3 = 300$.

In Figure 4.1, we fix $\mu_1 = 0.2$, $\mu_2 = 0.8$ and $\beta = 0.5$, and display the optimal critical value r^* and its expected minimum cost $F(r^*)$ by varying the arrival rate λ. Figure 4.1 shows that: (i) The optimal critical value r^* decreases as λ increases from 0.05 to 0.15, while it does not change at all when λ varies from 0.15 to 0.4. (ii) The minimum expected cost $F(r^*)$ first decreases and then increases as λ increases. Intuitively, the optimal critical value r^* is insensitive to changes with λ.

In Figure 4.2, we fix $\lambda = 0.1$, $\mu_2 = 0.8$ and $\beta = 0.5$, and display the optimal critical value r^* and its expected minimum cost $F(r^*)$ by varying the slow service rate μ_1. Figure 4.2 shows that the optimal critical value r^* and its minimum expected cost $F(r^*)$ increase as μ_1 increases.

In Figure 4.3, we fix $\lambda = 0.1$, $\mu_1 = 0.2$ and $\beta = 0.5$, and display the optimal critical value r^* and its expected minimum cost $F(r^*)$ by varying the fast service rate μ_2. Figure 4.3 shows that: (i) The optimal critical value r^* decreases as μ_2 increases from 0.3 to 0.5, while it does not change at all when μ_2 varies from 0.5 to 0.9. (ii) The minimum expected cost $F(r^*)$ rarely changes when μ_2 varies from 0.3 to 0.9. Intuitively, the optimal critical value r^* and its minimum expected cost may be too insensitive to changes with μ_2.

In Figure 4.4, we fix $\lambda = 0.1$, $\mu_1 = 0.2$ and $\mu_2 = 0.8$, and display the optimal critical value r^* and its expected minimum cost $F(r^*)$ by varying the probability β. Figure 4.4 shows that: (i) The optimal critical value r^* does not change at all when β varies from 0.1 to 0.7. (ii) The minimum expected cost $F(r^*)$ decreases slightly as β increases. Intuitively, the optimal critical value r^* and its expected minimum cost $F(r^*)$ are insensitive to changes with β.

It appears from Figures 4.1–4.4 that: (i) β does not affect r^*, but slightly affects $F(r^*)$. (ii) λ affects $F(r^*)$ and slightly affects r^*. (iii) μ_2 rarely affects r^* and $F(r^*)$. And (iv) μ_1 affects r^* and $F(r^*)$ significantly.

Furthermore, we perform a sensitivity analysis on the optimal value r^* and its expected cost $F(r^*)$ based on changes in values of the cost parameters C_1, C_2, C_3 and C_4.

Let the service time follow a 2-stage Erlang distribution, and employ the system parameters $\lambda = 0.1$, $\mu_1 = 0.2$, $\mu_2 = 0.8$ and $\beta = 0.5$. The numerical results of the optimal critical value r^* and its expected minimum cost $F(r^*)$ are illustrated in Figures 4.5–4.8.

It can be easily seen from Figures 4.5–4.8 that: (i) The cost C_1 slightly affects r^* and $F(r^*)$ when C_1 is larger than 150. (ii) The cost C_2 and C_4 affect r^* and $F(r^*)$ significantly. And (iii) the cost C_3 rarely affects r^* and $F(r^*)$.

5 Conclusions

We have considered an M/E$_k$/1 queuing system with balking and state-dependent service rate. By using the matrix geometric solution, we have obtained the matrix solution of the steady-state probability distribution and the explicit expressions of some performance measures of the system. Based on these performance measures, we have developed a cost model to determine the optimal critical value r^* to minimize the total expected cost per unit time. Furthermore, we have performed sensitivity analysis for the optimal critical value r^* and its expected minimum cost $F(r^*)$ with various parameters. Our numerical investigations indicate that: (i) The slow service rate μ_1, the cost parameters C_2 and C_4 affect the optimal critical value r^* and its expected minimum cost $F(r^*)$ significantly. (ii) The other parameters such as λ, μ_2, β and the cost parameters C_1 and C_3 rarely affect the optimal critical value r^* and its expected minimum cost $F(r^*)$.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (No. 10271102) and the Natural Science Foundation of Hebei Province (No. A2004000185), China, and was supported in part by GRANT-IN-AID FOR SCIENTIFIC RESEARCH (No. 16560350) and MEXT.ORC (2004-2008), Japan.

References

[1] Gaver, D. P. and Lehoczky, J. P. Gaussian approximations to service problems: a communication system example. *Journal of Applied Probability* **13** (1976) 768–780.

[2] Sharma, R. R. and Rai, R. C. and Mishra, A. Optimal bus services on express basis in the case of balking and reneging. *European Journal of Operational Research* **66** (1993) 113–123.

[3] Ackere, V. and Ninios, P. Simulation and queueing theory applied to a single-server queue with advertising and balking. *Journal of the Operational Research Society* **44** (1993) 407–414.

[4] Shawky, A. I. The single-server machine interference model with balking, reneging and an additional server for longer queues. *Microelectronic and Reliability* **37** (1997) 355–357.

[5] Haight, F. A. Queueing with balking. *Biometrika* **44** (1957) 360–369.

[6] Haight, F. A. Queueing with reneging. *Metrika* **2** (1959) 186–197.

[7] Ancker, Jr. C. J. and Gafarian, A. V. Some queueing problems with balking and reneging: I. *Operation Research* **11** (1963) 88–100.

[8] Ancker, Jr. C. J. and Gafarian, A. V. Some queueing problems with balking and reneging: II. *Operation Research* **11** (1963) 928–937.

[9] Abou-EI-Ata, M. O. The state-dependent queue: M/M/1/N with reneging and general balk functions. *Microelectronic and Reliability* **31** (1991) 1001–1007.

[10] Abou-EI-Ata, M. O. and Kotb, K. A. M. Linearly dependent service rate for the queue: M/M/1/N with general balk function, reflecting barrier, reneging and additional service for longer queue. *Microelectronic and Reliability* **32** (1992) 1693–1698.

[11] Abou-EI-Ata, M. O., Alseedy, R. O. and Kotb, K. A. M. Transient solution of the state-dependent queue: M/M/1/N with balking and reflecting barrier. *Microelectronic and Reliability*, **33** (1993) 681–688.

[12] Abou-EI-Ata, M. O. and Shawky, A. I. Single-server Markovian overflow queue with balking, reneging and an additional server for longer queues. *Microelectronic and Reliability* **32** (1992) 1389–1394.

[13] Drekic, S. and Woolford, D. G. A preemptive priority queue with balking. *European Journal of Operational Research* **164** (2005) 387–401.

[14] Al-Seedy, R. O. and Kotb, K. A. M. Transient solution of the state-dependent queue: M/M/1 with balking. *Advances in Modelling* & *Simulation* **35** (1991) 55–64.

[15] Al-Seedy, R. O. Analytical solution of the state-dependent Erlangian queue: M/E_j/1/N with balking. *Microelectronic and Reliability* **36** (1996) 203–206.

[16] Neuts, M. F. *Matrix-Geometric Solutions in Stochastic Models.* Johns Hopkins University Press, Baltimore, 1981.

[17] Rao, S. S. Queuing with balking and reneging in M/G/1 systems. *Metrika* **12** (1968) 173–188.