International Journal of Pure and Applied Mathematics
Volume 28 No. 1 2006, 101-115

OPTIMAL PERFORMANCE ANALYSIS OF AN M/M/1/N QUEUE SYSTEM WITH BALKING, RENEGING AND SERVER VACATION

Dequan Yue[1], Yan Zhang[2], Wuyi Yue[3 §]
[1,2]Department of Statistics
College of Science
Yanshan University
Qinhuangdao, Hebei, 066004, P.R. CHINA
[1]e-mail: ydq@ysu.edu.cn
[2]e-mail: helen89@ysu.edu.cn
[3]Department of Information Science and Systems Engineering
Konan University
8-9-1, Okamoto, Higashinada-ku, Kobe, 658-8501, JAPAN
e-mail: yue@konan-u.ac.jp

Abstract: This paper presents an analysis for an M/M/1/N queueing system with balking, reneging and server vacations. Arriving customers balk with a probability and renege according to a negative exponential distribution. It is assumed that the server has multiple vacations. By using the Markov process method, we first develop the equations of the steady state probabilities. Then, we derive a matrix form solution of the steady-state probabilities. Next, we give some performance measures of the system such as the expected number of the waiting customers, the expected number of the customers in the system, the average rate of customer loss due to impatience and so on. And based on the performance analysis, we formulate a cost model to determine the optimal service rate. Finally, we present some numerical examples to demonstrate how the various parameters influence the behavior of the system.

AMS Subject Classification: 60K25
Key Words: vacation, balking, reneging, queueing system, optimal performance analysis

Received: April 16, 2006

[§]Correspondence author

1. Introduction

Many practical queueing systems, especially those with balking and reneging, have been widely applied to many real-life problems, such as the situations involving impatient telephone switchboard customers, hospital emergency rooms' handling of critical patients, and perishable goods storage inventory systems [9]. In this paper, we consider an M/M/1/N queueing system with balking and reneging. We also consider the server to have multiple vacations, i.e., the server leaves for a random length of time whenever the system becomes empty. At the end of a vacation, the server will take another vacation if the system is still empty.

Queueing systems with balking, reneging, or both have been studied by many researchers. Haight [6] first considered an M/M/1 queue with balking. An M/M/1 queue with customers reneging was also proposed by Haight [7]. The combined effects of balking and reneging in an M/M/1/N queue have been investigated by Ancker and Gafarian [2], [3]. Abou-EI-Ata and Hariri [1] considered a multiple servers queueing system M/M/c/N with balking and reneging. Wang and Chang [12] extended Abou-EI-Ata and Hariri's work to study an M/M/c/N queue with balking, reneging and server breakdowns.

On the other hand, queueing systems with server vacations have attracted a lot of attention from numerous researchers since the paper by Levy and Yechiali [8] was presented. Server vacations are useful for a system where the server wants to utilize its idle time for different purposes. An excellent survey of queueing systems with server vacations can be found in papers by Doshi [4] and Takagi [10].

However, most of the research works about queueing systems have not considered balking, reneging and server vacations together. There was only one paper [11] that we knew of that considered an $M^X/G/1$ queue with balking involving multiple vacations.

Queueing models with server vacations accommodate real-world situations more closely. Such model frequently occurs in areas of computer and communications, or manufacturing systems. For example, in an assembly line, a worker may have some idle time between subsequent jobs. To utilize the time effectively, managers can assign secondary jobs to the worker. However, it is important that the worker return to do his primary job when he completes the secondary jobs. This motivates us to study a queueing system with balking, reneging and server vacations.

The rest of this paper is organized as follows. In Section 2, we give a description of the queueing system model. In Section 3, we derive the steady-state equations by using the Markov process method. By writing the transition rate matrix as a block matrix, we get the matrix form solution of the steady-state probabilities. In Section 4, we give some performance measures of the system. Based on the performance analysis, we also formulate a cost model to determine the optimal service rate. In Section 5, we present some numerical examples to demonstrate how the various parameters influence the behavior of the system. Conclusions are given in Section 6.

2. System Model

In this paper, we consider an M/M/1/N queueing system with balking, reneging and server vacations together. The assumptions of the system model are as follows:

(a) Customers arrive at the system one by one according to a Poisson process with rate λ. On arrival a customer either decides to join the queue with probability b_n or balk with probability $1 - b_n$ when n customers are ahead of him $(n = 0, 1, ..., N - 1)$, where N is the maximum number of customers in the system, and

$$0 \leq b_{n+1} \leq b_n < 1, \quad 1 \leq n \leq N - 1, \qquad b_0 = 1 \quad \text{and} \quad b_n = 0, \quad n \geq N.$$

(b) After joining the queue each customer will wait a certain length of time T for service to begin. If his service has not begun by then, he will get impatient and leave the queue without getting service. This time T is a random variable according to an exponential distribution whose density function is given by

$$u(t) = \alpha e^{-\alpha t}, \quad t \geq 0, \ \alpha > 0\,,$$

where α is the occurring rate of the waiting time T, namely, $1/\alpha$ is the average length of the waiting time T. Since the arrival and the departure of the impatient customers without service are independent, the function of customer's average reneging rate can be given by

$$r(n) = (n - i)\alpha, \quad i \leq n \leq N, \ i = 0, 1, \qquad r(n) = 0, \quad n > N.$$

(c) The customers are served on a first-come, first served (FCFS) discipline. Once a customer's service commences, the service always proceeds to completion. The service times are assumed to be distributed according to an exponential distribution with density function as follows:

$$s(t) = \mu e^{-\mu t}, \quad t \geq 0, \ \mu > 0\,,$$

where μ is the service rate.

(d) Whenever the system is empty, the server goes on a sequence of vacations for a random period V of time. If the server returns from a vacation to find no customer waiting, it will begin another vacation immediately. It is assumed that V has an exponential distribution with the following density function:

$$v(t) = \eta e^{-\eta t}, \quad t \geq 0,\ \eta > 0\,,$$

where η is the vacation rate of the server.

3. Steady-State Probability

In this section, we derive the steady-state probabilities by using the Markov process method. Let $p_0(n)$ be the probability that there are n customers in the system when the server is on vacation. Let $p_1(n)$ be the probability that there are n customers in the system when the server is available.

By applying the Markov process theory, we can obtain the following set of steady-state equations.

$$\eta p_0(1) + (\mu + \alpha)p_1(2) = (\lambda b_1 + \mu)p_1(1), \tag{1}$$

$$\begin{aligned} &\lambda b_{n-1}p_1(n-1) + \eta p_0(n) + (\mu + n\alpha)p_1(n+1) \\ &\quad = [\lambda b_n + \mu + (n-1)\alpha]p_1(n), \quad n = 2, 3, ..., N-1, \end{aligned} \tag{2}$$

$$\lambda b_{N-1}p_0(N-1) + \eta p_0(N) = [\mu + (N-1)\alpha]p_0(N), \tag{3}$$

$$\mu p_1(1) + \alpha p_0(1) = \lambda p_0(0), \tag{4}$$

$$\begin{aligned} &\lambda b_{n-1}p_0(n-1) + (n+1)\alpha p_0(n+1) \\ &\quad = (n\alpha + \lambda b_n + \eta)p_0(n), \quad n = 1, 2, ..., N-1, \end{aligned} \tag{5}$$

$$\lambda b_{N-1}p_0(N-1) = (N\alpha + \eta)p_0(N). \tag{6}$$

The transition rate matrix $\boldsymbol{Q}$ of the Markov process has the following block form:

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{B}_0 & \boldsymbol{A} \\ \boldsymbol{C} & \boldsymbol{B}_1 \end{pmatrix} \tag{7}$$

where

$$\boldsymbol{C}=\begin{pmatrix}\mu & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 0\end{pmatrix}, \quad \boldsymbol{A}=\begin{pmatrix}0 & 0 & \cdots & 0\\ \eta & 0 & \cdots & 0\\ 0 & \eta & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & \eta\end{pmatrix},$$

$$\boldsymbol{B}_0=\begin{pmatrix}-\lambda & \lambda & 0 & \cdots & 0 & 0\\ \alpha & -c_1 & \lambda b_1 & \cdots & 0 & 0\\ 0 & 2\alpha & -c_2 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & -c_{N-1} & \lambda b_{N-1}\\ 0 & 0 & 0 & \cdots & N\alpha & -c_N\end{pmatrix},$$

and

$$\boldsymbol{B}_1=\begin{pmatrix}-d_1 & \lambda b_1 & 0 & \cdots & 0 & 0\\ e_1 & -d_2 & \lambda b_2 & \cdots & 0 & 0\\ 0 & e_2 & -d_3 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & -d_{N-1} & \lambda b_{N-1}\\ 0 & 0 & 0 & \cdots & e_{N-1} & -e_{N-1}\end{pmatrix}$$

where $c_i = i\alpha + \lambda b_i + \eta$, $d_i = \lambda b_i + \mu + (i-1)\alpha$, $e_i = \mu + i\alpha$, $i = 1, 2, ..., N-1$, $c_N = N\alpha + \eta$. $\boldsymbol{C}$ is an $N \times (N+1)$ matrix, $\boldsymbol{A}$ is an $(N+1) \times N$ matrix, $\boldsymbol{B}_0$ is an $N+1$ square matrix, $\boldsymbol{B}_1$ is an N square matrix.

Let $\boldsymbol{P} = (\boldsymbol{P}_0, \boldsymbol{P}_1)$ be the corresponding steady-state probability vector of $\boldsymbol{Q}$, where $\boldsymbol{P}_0 = (p_0(0), p_0(1), ..., p_0(N))$ and $\boldsymbol{P}_1 = (p_1(1), p_1(2), ..., p_1(N))$. The steady-state probability vector $\boldsymbol{P}$ must satisfy the following equations:

$$\begin{cases}\boldsymbol{PQ} = \boldsymbol{0}\,,\\ \boldsymbol{Pe} = 1\,,\end{cases} \tag{8}$$

where $\boldsymbol{e}$ is a $2N+1$ column vector that all of elements equal 1. From equations (7) and (8), we obtain by considering some routine substitutions that

$$\boldsymbol{P}_0\boldsymbol{B}_0 + \boldsymbol{P}_1\boldsymbol{C} = \boldsymbol{0}, \tag{9}$$

$$\boldsymbol{P}_0\boldsymbol{A} + \boldsymbol{P}_1\boldsymbol{B}_1 = \boldsymbol{0}, \tag{10}$$

$$\boldsymbol{P}_0\boldsymbol{e}_0 + \boldsymbol{P}_1\boldsymbol{e}_1 = 1 \tag{11}$$

where $\boldsymbol{e}_0$ is an $N+1$ column vector that all of elements equal 1, and $\boldsymbol{e}_1$ is an N column vector that all of elements equal 1.

Let $\boldsymbol{B}_0^{-1}$ and $\boldsymbol{B}_1^{-1}$ be the inverse matrix of $\boldsymbol{B}_0$ and $\boldsymbol{B}_1$, respectively. Solving equation (9), we get

$$\boldsymbol{P}_0 = -\boldsymbol{P}_1\boldsymbol{C}\boldsymbol{B}_0^{-1}. \tag{12}$$

Substituting equation (12) into equations (10) and (11), we obtain

$$\boldsymbol{P}_1(\boldsymbol{I} - \boldsymbol{C}\boldsymbol{B}_0^{-1}\boldsymbol{A}\boldsymbol{B}_1^{-1}) = \boldsymbol{0}, \tag{13}$$

$$\boldsymbol{P}_1(\boldsymbol{e}_1 - \boldsymbol{C}\boldsymbol{B}_0^{-1}\boldsymbol{e}_0) = 1. \tag{14}$$

Solving equations (12), (13) and (14), we can get the steady-state probabilities of the system given by the following theorem.

Let $\boldsymbol{a} = (a_{11}^{-1}, a_{12}^{-1}, ..., a_{1N+1}^{-1})$ be the first row vector of the matrix $\boldsymbol{B}_0^{-1}$, then $\widetilde{\boldsymbol{a}} = (a_{12}^{-1}, a_{13}^{-1}, ..., a_{1N+1}^{-1})$ is an N row vector. Let $\boldsymbol{\varepsilon}_i = (0, ..., 0, 1, 0, ..., 0)$ be an N column unit vector, $i = 1, 2, ..., N$.

Theorem 1. *The steady-state probabilities are given by*

$$p_1(1) = (1 - \mu\boldsymbol{a}\boldsymbol{e}_0 + \mu\eta\widetilde{\boldsymbol{a}}\sum_{i=2}^{N}\boldsymbol{B}_1^{-1}\boldsymbol{\varepsilon}_i)^{-1}, \tag{15}$$

$$p_1(i) = p_1(1)\mu\eta\widetilde{\boldsymbol{a}}\boldsymbol{B}_1^{-1}\boldsymbol{\varepsilon}_i, \quad i = 2, 3, ..., N, \tag{16}$$

$$p_0(i-1) = -p_1(1)\mu a_{1i}^{-1}, \quad i = 1, 2, ..., N+1. \tag{17}$$

Proof. $\boldsymbol{C}$ and $\boldsymbol{A}$ can be rewritten as the following block matrix:

$$\boldsymbol{C} = \begin{pmatrix} \mu & \boldsymbol{O}_1 \\ \boldsymbol{O}_2 & \boldsymbol{O}_3 \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} \boldsymbol{O}_1 \\ \eta\boldsymbol{I} \end{pmatrix},$$

where $\boldsymbol{O}_1$ is a $1 \times N$ matrix, $\boldsymbol{O}_2$ is an $(N-1) \times 1$ matrix, $\boldsymbol{O}_3$ is an $(N-1) \times N$ matrix, $\boldsymbol{I}$ is an N identity matrix, $\boldsymbol{C}$ is an $N \times (N+1)$ matrix and $\boldsymbol{A}$ is an $(N+1) \times N$ matrix.

Let a_{ij}^{-1}, $i = 1, 2, ..., N+1$, $j = 1, 2, ..., N+1$, be elements of the matrix $\boldsymbol{B}_0^{-1}$. Then, we have

$$\boldsymbol{C}\boldsymbol{B}_0^{-1} = \begin{pmatrix} \mu\boldsymbol{a} \\ \boldsymbol{O}_4 \end{pmatrix} \tag{18}$$

and

$$\boldsymbol{A}\boldsymbol{B}_1^{-1} = \begin{pmatrix} \boldsymbol{O}_1 \\ \eta\boldsymbol{B}_1{}^{-1} \end{pmatrix}, \tag{19}$$

where $\boldsymbol{O}_4$ is an $(N-1)\times(N+1)$ matrix. equation (18) is an $N\times(N+1)$ matrix, and equation (19) is an $(N+1)\times N$ matrix. $\boldsymbol{C}\boldsymbol{B}_0^{-1}$ can be rewritten as

$$\boldsymbol{C}\boldsymbol{B}_0^{-1}=\begin{pmatrix}\mu a_{11}^{-1} & \mu\widetilde{\boldsymbol{a}}\\ \boldsymbol{O}_2 & \boldsymbol{O}_5\end{pmatrix}, \tag{20}$$

where $\boldsymbol{O}_5$ is an $(N-1)\times N$ matrix. Thus, we have from equations (19) and (20) that

$$\boldsymbol{C}\boldsymbol{B}_0^{-1}\boldsymbol{A}\boldsymbol{B}_1^{-1}=\begin{pmatrix}\mu\eta\widetilde{\boldsymbol{a}}\boldsymbol{B}_1^{-1}\\ \boldsymbol{O}_5\end{pmatrix}. \tag{21}$$

From equations (13) and (21), we have

$$\boldsymbol{P}_1=\boldsymbol{P}_1\begin{pmatrix}\mu\eta\widetilde{\boldsymbol{a}}\boldsymbol{B}_1^{-1}\\ \boldsymbol{O}_5\end{pmatrix}=p_1(1)\mu\eta\widetilde{\boldsymbol{a}}\boldsymbol{B}_1^{-1}. \tag{22}$$

Then, we have

$$p_1(i)=p_1(1)\mu\eta\widetilde{\boldsymbol{a}}\boldsymbol{B}_1^{-1}\boldsymbol{\varepsilon}_i,\quad i=2,3,...,N. \tag{23}$$

From equations (14) and (18), we have

$$\boldsymbol{P}_1\left[\boldsymbol{e}_1-\begin{pmatrix}\mu\boldsymbol{a}\\ \boldsymbol{O}_4\end{pmatrix}\boldsymbol{e}_0\right]=1, \tag{24}$$

or equivalently, we have

$$p_1(1)(1-\mu\boldsymbol{a}\boldsymbol{e}_0)+\sum_{i=2}^{N}p_1(i)=1. \tag{25}$$

Substituting equation (23) into equation (25), then we get equation (15). From equations (12) and (18), we obtain

$$\boldsymbol{P}_0=-\boldsymbol{P}_1\begin{pmatrix}\mu\boldsymbol{a}\\ \boldsymbol{O}_4\end{pmatrix}=-p_1(1)\mu\boldsymbol{a}. \tag{26}$$

Then, we have

$$p_0(i-1)=-p_1(1)\mu a_{1i}^{-1},\quad i=1,2,...,N+1.$$

This completes the proof. □

Remark 1. From Theorem 1, we need mainly to calculate the vectors $\boldsymbol{a}$ and $\boldsymbol{B}_1^{-1}\boldsymbol{\varepsilon}_i$, $i = 2, 3, ..., N$. Let $\boldsymbol{\varepsilon}_0 = (1, 0, ..., 0)$ be an $N+1$ unit row vector. Then, $\boldsymbol{a} = \boldsymbol{\varepsilon}_0 \boldsymbol{B}_1^{-1}$, or equivalently, $\boldsymbol{B}_0^T \boldsymbol{a}^T = \boldsymbol{\varepsilon}_0^T$. Thus, the transpose vector $\boldsymbol{a}^T$ of the vector $\boldsymbol{a}$ can be obtained by solving the following $N+1$ linear equations: $\boldsymbol{B}_0^T \boldsymbol{x} = \boldsymbol{\varepsilon}_0^T$. Similarly, let $\boldsymbol{B}_1^{-1}\boldsymbol{\varepsilon}_i = \boldsymbol{x}$, then $\boldsymbol{B}_1 \boldsymbol{x} = \boldsymbol{\varepsilon}_i$. Thus, $\boldsymbol{B}_1^{-1}\boldsymbol{\varepsilon}_i$ can be obtained by solving the following N linear equations: $\boldsymbol{B}_1 \boldsymbol{x} = \boldsymbol{\varepsilon}_i$.

4. Performance Measures and Cost Model

In this section, we give some performance measures of the system. Based on these performance measures, we develop a cost model to determine the optimal service rate of the system.

Using the steady-state probability presented in Section 3, we can obtain some performance measures of the system, such as the busy probability P_B of the server, the vacation probability P_V of the server, the expected number $E(N_q)$ of the waiting customers and the expected number $E(N)$ of the customers in the system as follows:

$$P_B = \sum_{i=1}^{N} p_1(i) = p_1(1)\left(1 + \mu\eta\widetilde{\boldsymbol{a}}\boldsymbol{\xi}_1\right), \tag{27}$$

$$P_V = \sum_{n=0}^{N} p_0(n) = 1 - P_B, \tag{28}$$

$$E(N_q) = \sum_{i=0}^{N} i p_0(i) + \sum_{i=1}^{N} (i-1) p_1(i) = p_1(1)\left(\mu\eta\widetilde{\boldsymbol{a}}(\boldsymbol{\xi}_2 - \boldsymbol{\xi}_1) - \mu \sum_{i=1}^{N} i a_{1i+1}^{-1}\right), \tag{29}$$

and

$$E(N) = \sum_{i=0}^{N} i p_0(i) + \sum_{i=1}^{N} i p_1(i) = p_1(1)\left(1 + \mu\eta\widetilde{\boldsymbol{a}}\boldsymbol{\xi}_2 - \mu \sum_{i=1}^{N} i a_{1i+1}^{-1}\right), \tag{30}$$

where $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are N column vectors, respectively given by

$$\boldsymbol{\xi}_1 = \boldsymbol{B}_1^{-1} \sum_{i=2}^{N} \boldsymbol{\varepsilon}_i, \boldsymbol{\xi}_2 = \boldsymbol{B}_1^{-1} \sum_{i=2}^{N} i\boldsymbol{\varepsilon}_i. \tag{31}$$

Using the concept proposed by Ancker and Gafarian [2], we can obtain the average balking rate B and the average reneging rate R as follows:

$$B = \sum_{i=0}^{N} \lambda(1-b_i)p_0(i) + \sum_{i=1}^{N} \lambda(1-b_i)p_1(i)$$
$$= p_1(1)\left(\mu\eta\widetilde{\boldsymbol{a}}\boldsymbol{\xi}_3 - \lambda\mu\sum_{i=0}^{N}(1-b_i)a_{1i+1}^{-1}\right) \quad (32)$$

and

$$R = \sum_{i=0}^{N} \alpha i p_0(i) + \sum_{i=1}^{N} \alpha(i-1)p_1(i) = \alpha E(N_q)\,, \quad (33)$$

where $\lambda(1-b_i)$ is the instantaneous balking rate, $i\alpha$ is the instantaneous reneging rate when there are i customers in the system, and $\boldsymbol{\xi}_3$ is an N column vector given by

$$\boldsymbol{\xi}_3 = \boldsymbol{B}_1^{-1}\sum_{i=1}^{N}(1-b_i)\boldsymbol{\varepsilon}_i.$$

The average rate L of customer loss due to impatience is given as follows:

$$L = B + R. \quad (34)$$

Remark 2. The vectors $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$ and $\boldsymbol{\xi}_3$ are solution vectors of the following N linear equations, respectively. Therefore, these vectors can be easily obtained by solving the following equations:

$$\boldsymbol{B}_1\boldsymbol{x} = \sum_{i=2}^{N}\boldsymbol{\varepsilon}_i, \quad \boldsymbol{B}_1\boldsymbol{x} = \sum_{i=2}^{N} i\boldsymbol{\varepsilon}_i, \quad \boldsymbol{B}_1\boldsymbol{x} = \sum_{i=1}^{N}(1-b_i)\boldsymbol{\varepsilon}_i.$$

In the following, we develop an expected cost model, in which service rate μ is a control variable. Our objective is to control this service rate to minimize the system's total expected cost per unit. Define the cost elements as follows:

C_1: cost per unit time when the server is busy;
C_2: cost per unit time when the server is on vacation;
C_3: cost per unit time when a customer joins the queue and waits for service;
C_4: cost per unit time when a customer balks or reneges.

Using the definitions of cost elements listed above, the total expected cost function per unit time can be given by

$$F(\mu) = C_1 P_B + C_2 P_V + C_3 E(N_q) + C_4 L\,, \tag{35}$$

where P_B, P_V, $E(N_q)$, L are given in equations (27)-(29) and (33).

In equation (34), the first two items are the cost incurred by the server, the third item $C_3 E(N_q)$ is the cost incurred by the customer's waiting, and the last item $C_4 L$ is the cost incurred by the customer loss. In next section, we numerically consider the optimal service rate and the optimal expected cost.

5. Numerical Results

We define an optimal service rate μ^* to be a service rate that minimize the total expected cost function, and an optimal expected cost $F(\mu^*)$ to be a corresponding minimum expected cost. Although the expected cost function is too complicated to derive an explicit expression of the optimal service rate, the performance measures and the optimal service rate can be numerically calculated. That is if the maximum number N of customers in the system is fixed, then the vectors $\boldsymbol{a}$, $\boldsymbol{B}_1^{-1}\boldsymbol{\varepsilon}_i$, $i = 2, 3, ..., N$, $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$ and $\boldsymbol{\xi}_3$ mentioned in Remark 1 and Remark 2 can be obtained by using the symbol operation function of *Matlab*. Thus, the performance measures presented in equation (34) and the expected cost function $F(\mu)$ can be obtained by the formulas in Section 4. Therefore, the optimal service rate μ^* and the optimal expected cost $F(\mu^*)$ can be numerically calculated by using *Matlab*.

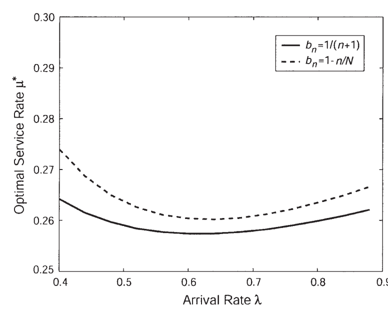

Figure 1: Optimal service rate μ^* versus arrival rate λ

In this section, we present some numerical examples to demonstrate how the various system parameters influence the optimal service rate μ^* and the optimal expected cost $F(\mu^*)$.

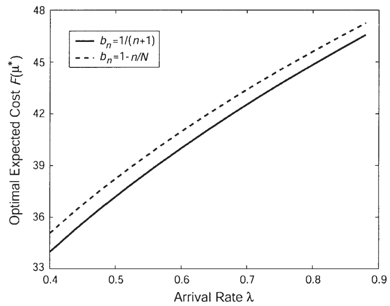

Figure 2: Optimal expected cost $F(\mu^*)$ versus arrival rate λ

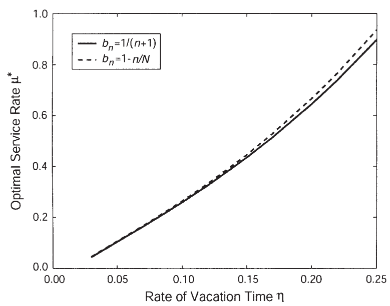

Figure 3: Optimal service rate μ^* versus rate of vacation time η

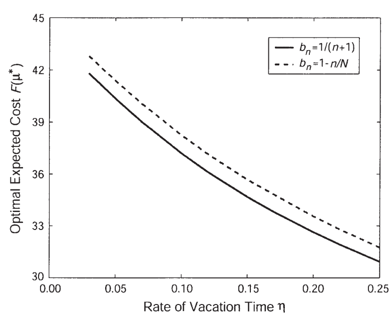

Figure 4: Optimal expected cost $F(\mu^*)$ versus rate of vacation time η

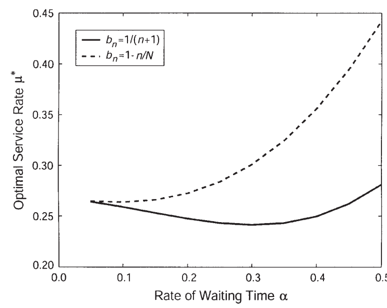

Figure 5: Optimal service rate μ^* versus rate of waiting time α

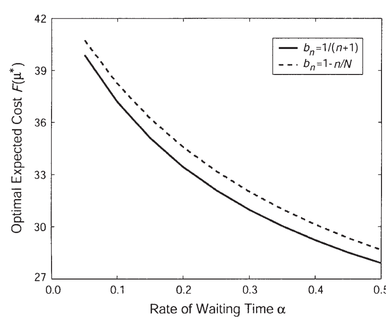

Figure 6: Optimal expected cost $F(\mu^*)$ versus rate of waiting time α

In order to reduce the computing time, we fix the maximum number $N = 3$ of customers in the system. We select the cost elements: $C_1 = 15$, $C_2 = 12$, $C_3 = 18$ and $C_4 = 12$ by referencing [12]. We consider the following two cases for the probability b_n that customers decide to join the queue by referencing [2] and [3]: $b_n = 1 - n/N$ and $b_n = 1/(n + 1)$.

In the first case, the probability $b_n = 1 - n/N$ is a linear monotonically decreasing function of system size n. In this case, the model in which only balking occurs, i.e., no reneging and no vacation are permitted, can be interpreted as a "Machine Interference" problem. We refer to [2] to see the equivalence of these two problems. In the "Machine Interference" problem, the total number of machines is N. The time required for servicing a machine is a random variable with an exponential distribution having service rate μ, and the time to failure of a given machine is exponential distributed with failure rate λ/N. Thus, the arrival rate of machines requiring repair when n of them are in need of repair is $(1 - n/N)\lambda$. Hence, this "Machine Interference" problem is equivalent to the model presented in this paper in which only balking occurs with the probability that customers decide to join the queue $b_n = 1 - n/N$. In the second case, the probability $b_n = 1/(n + 1)$ is not a linear monotonically decreasing function of system size n. It is fractional. That is that the customers become discouraged to join the queue as the queue becomes longer [5].

The numerical results of the optimal service rate μ^* and the optimal expected cost $F(\mu^*)$ changed with the various system parameters are illustrated in Figures 1-6.

In Figure 1 and Figure 2, we select the rate of the waiting time $\alpha = 0.1$, the rate of the vacation time $\eta = 0.1$ as an example, and change the values of the arrival rates of customers λ. From Figure 1 and Figure 2, we observe that: (i) the optimal service rate μ^* first decreases and then increases slightly with the increasing of λ, and (ii) the optimal expected cost $F(\mu^*)$ increases greatly with the increasing of λ. This is because the number of the customers in the system increases with the increasing of λ. Thus, the busy probability P_B, the expected number of waiting customers $E(N_q)$ and the average rate of customers loss L all increase which results in the increasing of the optimal expected cost.

In Figure 3 and Figure 4, we select $\alpha = 0.1$, $\lambda = 0.5$ as an example, and change the values of η. We can see from Figure 3 and Figure 4 that: (i) the optimal service rate μ^* increases greatly with the slightly increasing of η, and (ii) the optimal expected cost $F(\mu^*)$ decreases with the increasing of η. This is because the mean vacation time of the server $1/\eta$ decreases with the increasing of η. Thus, the busy probability P_B, the expected number of waiting customers $E(N_q)$ and the average rate of customers loss L all decrease which results in the decreasing of the optimal expected cost.

In Figure 5 and Figure 6, we select $\lambda = 0.5$, $\eta = 0.1$ as an example, and change the values of α. Figure 5 shows that: the optimal service rate μ^* for the case with the probability $b_n = 1/(n+1)$ first decreases and then increases slightly with the increasing of α, while the optimal service rate μ^* for the case with the probability $b_n = 1-n/N$ first decreases and then increases significantly with the increasing of α. Figure 6 shows that the optimal expected cost $F(\mu^*)$ for both cases decreases with the increasing of α. This is because the mean waiting time of impatient customers decreases with the increasing of α. Thus, the average rate of customers loss L increases, while the expected number of waiting customers $E(N_q)$ and the busy probability P_B decreases which results in the decreasing of the optimal expected cost.

From the six figures, we observe that the optimal service rate μ^* and the optimal expected cost $F(\mu^*)$ with the probability $b_n = 1-n/N$ are always larger than that with the probability $b_n = 1/(n+1)$. This is because the probability $b_n = 1 - n/N$ is larger than the probability $b_n = 1/(n+1)$ when $1 < n < N$.

6. Conclusions

In this paper, we analyzed an M/M/1/N queueing system with balking, reneging and server vacations. We developed the equations of the steady state probabilities and derived the matrix form solution of the steady-state probabilities. We also gave some performance measures of the system, and formulated a cost model to determine the optimal service rate. Although the expected cost function is too complicated to derive an explicit expression of the optimal service rate, the optimal service rate and the optimal expected cost can be numerically calculated by the formulas in Section 4. Some numerical examples were presented to demonstrate how the various parameters of the model influence the behavior of the system.

Acknowledgments

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] M.O. Abou El-Ata, A.M.A. Hariri, The M/M/C/N queue with balking and reneging, *Computers and Operations Research*, **19** (1992), 713-716.

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

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

[4] B.T. Doshi, Single server queues with vacation: a survey, *Queueing System*, **1** (1986), 29-66.

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

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

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

[8] Y. Levy, U. Yechiali, Utilization of idle time in an M/G/1 queueing system, *Management Science*, **22** (1975), 202-211.

[9] E. Robert, Reneging phenomenon of single channel queues, *Mathematics of Operations Research*, **4** (1979), 162-178.

[10] H. Takagi, *Queueing Analysis, A Foundation of Performance Evaluation, Volume 1: Vacation and Priority Systems*, Elsevier, Amsterdam (1991).

[11] L.A. Thomo, Multiple vacation model $M^X/G/1$ with balking, *Nonlinear Analysis, Theory, Methods and Applications*, **30** (1997), 2025-2030.

[12] K-H. Wang, Y-C. Chang, Cost analysis of a finite M/M/R queueing system with balking, reneging and server breakdowns, *Mathematical Methods of Operations Research*, **56** (2002), 169-180.