The Sixth International Symposium on Operations Research and Its Applications (ISORA'06)
Xinjiang, China, August 8–12, 2006

Performance Analysis of an M/M/c/N Queueing System with Balking, Reneging and Synchronous Vacations of Partial Servers

Dequan Yue[1] Wuyi Yue[2] Yanping Sun[1]

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

Abstract In this paper, we present an analysis for an M/M/c/N queueing system with balking, reneging and synchronous vacations of partial servers together. It is assumed that arriving customers balk with a probability and renege according to an exponential distribution. The vacation policy prescribes that any d (0<d<c) servers take a vacation together when these d servers find that no customers are waiting in the line at a service completion instant and $c-d$ servers are always available, either serving the customers or remaining idle. At a vacation completion instant, if the number of customers in the system is not more than $c-d$, these d servers take another vacation together; otherwise, these d servers return to serving the queue. The service times and the vacation times are assumed to follow independent exponential distributions. By using the Markov process theory, we first develop the equations of the steady state probabilities and derive a matrix form solution of the steady-state probabilities. Then we give some performance measures of the system such as the expected number of waiting customers, the expected number of the customers in the system, the average rate of customer loss due to impatience, etc. Based on the performance analysis, we formulate a cost model to determine the optimal number of servers on vacation. Finally, we perform sensitivity analysis through numerical experiments.

Keywords synchronous vacation; balking; reneging; queueing system; steady-state probability; partial servers; cost model

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, the customers 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). Balking and reneging are not only a common phenomena in queues arising in daily activities, but also in various machine repair models. For related literature, interested readers may refer to [1], [2], and references therein. An interesting example of the occurrence of balking and reneging in air defence systems is given in Ancker and Gafarian [3].

Queueing systems with balking, reneging, or both have been studied by many researchers. Haight [4] first considered an M/M/1 queue with balking. An M/M/1 queue with customers reneging was also proposed by Haight [5]. The combined effects of balking and reneging in an M/M/1/N queue have been investigated by Ancker and Gafarian [6], [7]. Abou-EI-Ata and Hariri [8] considered the multiple servers queueing system M/M/c/N with balking and reneging. Wang and Chang [9] extended this work to study an M/M/c/N queue with balking, reneging and server breakdowns.

On the other hand, queueing models with vacations have been studied by many researchers during the past two decades and have been found to be applicable in analyzing numerous real world queueing situations such as flexible manufacturing systems, service systems, and telecommunication systems. Several excellent surveys on these vacation models have been done by Doshi [10], [11] and Takagi [12]. However, there are only a few works that take into consideration balking and reneging phenomena involving server vacations. The readers may refer to Zhang et al. [13] where an M/M/1/N queueing system with balking, reneging and server vacations was considered.

Multiple server vacation models are more flexible and applicable in practice than single server counterparts. In many practical multiple server systems, some servers perform secondary jobs (take vacations) when they become idle, while other servers are always available for serving arriving customers. These types of models are called “partial server vacation models". Recently, Zhang and Tian [14] studied M/M/c queueing system with synchronous vacations of partial servers. They obtained the stationary distribution of the queue length and proved several conditional stochastic decomposition results for the queue length and the customer waiting time. In this paper, we consider the balking and reneging phenomena in an M/M/c/N queueing system with the same partial server vacation policy as in [14].

The rest of this paper is organized as follows. In the next section, we give a description of the queueing model. In Section 3, we derive the steady-state equations by 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 formulate a cost model to determine the optimal number of servers on vacation. In Section 5, we perform sensitivity analysis through numerical experiments. Conclusions are given in Section 6.

2 System Model

In this paper, we consider an M/M/c/N queueing system with balking, reneging and a vacation policy for servers. The system capacity is finite N. The assumptions of the system model are as follows:

(a) Customers arrive at the system one by one according to a Poisson process with arrival rate λ. The service time for each server is assumed to be distributed according to an exponential distribution with service rate μ.

(b) A customer who on arrival finds n customers in the system, either decides to enter the queue with probability b_n or balk with probability $1-b_n$,

$$0 \leq b_{n+1} \leq b_n < 1, \quad R-d \leq n \leq N-1,$$
$$b_n = 0, \quad n \geq N.$$

(c) The vacation policy prescribes that any d $(0 < d < c)$ servers take a vacation together when these d servers find that no customers are waiting in the line at a service completion instant and $c-d$ servers are always available, either serving the customers or remaining idle. At a vacation completion instant, if the number of customers in the system is not more than $c-d$ (still no customers waiting in line), these d servers take another vacation together; otherwise, these d servers return to serving the queue. Because these d servers take vacations simultaneously, these vacations are synchronous vacations. The vacation time is assumed to be exponentially distributed with mean $1/\eta$.

(d) After joining the queue, in the case where all the servers available are occupied, each customer will wait a certain length of time T for service to begin before they get impatient and leave the queue without receiving service. This time T is assumed to be distributed according to an exponential distribution with mean $1/\alpha$. Let i denote the number of severs being busy and n represent the number of customers in the system. If n is less than or equal to i, the customers will get service instantly upon arrival to the server, and the phenomenon of reneging will not occur. If n is greater than i, then there are $n-i$ customers who have to wait in the queue. Since the arrival and the departure of the impatient customers without service are independent, the average reneging rate in this state is given by $(n-i)\alpha$.

(e) The service order is assumed on first-come first-served (FCFS) basis and the inter-arrival times, service times, and vacations are mutually independent.

3 Steady-State Probability

In this section, we derive the steady-state probabilities by the Markov process method.

3.1 Steady-state equations

Let $L(t)$ be the number of customers in the system at time t and let

$$J(t) = \begin{cases} 0, & d \text{ vacation servers are on vacation at time } t \\ 1, & d \text{ vacation servers are not on vacation at time } t. \end{cases}$$

Then $\{L(t), J(t)\}$ is a Markov process with state space

$$\Omega = \{(k,0) : 0 \leq k \leq c-d\} \cup \{(k,j) : c-d+1 \leq k \leq N, j = 0,1\}.$$

Define the steady-state probabilities of the system as follows:

$$P_0(n) = \lim_{t\to\infty} P\{L(t) = n, J(t) = 0\}, \quad 0 \leq n \leq N,$$
$$P_1(n) = \lim_{t\to\infty} P\{L(t) = n, J(t) = 1\}, \quad c-d+1 \leq n \leq N.$$

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

$$\mu P_0(1) = \lambda P_0(0),$$

$$\lambda P_0(n-1) + (n+1)\mu P_0(n+1) = (\lambda + n\mu)P_0(n), \quad 1 \le n < c-d,$$

$$(c-d+1)\mu P_1(c-d+1) + \lambda P_0(c-d-1) + [(c-d)\mu + \alpha]P_0(c-d+1)$$
$$= [(c-d)\mu + \lambda b_{c-d}]P_0(c-d),$$

$$\lambda b_{N-1}P_0(n-1) + \{(c-d)\mu + (n+1+d-c)\alpha\}P_0(n+1)$$
$$= \{\eta + (c-d)\mu + (n+d-c)\alpha + \lambda b_n\}P_0(n), \quad c-d < n < N,$$

$$\lambda b_{N-1}P_0(N-1) = \{\eta + (c-d)\mu + (N+d-c)\alpha\}P_0(N),$$

$$\eta P_0(c-d+1) + (c-d+2)\mu P_1(c-d+2) = [(c-d+1)\mu + \lambda]P_1(c-d+1),$$

$$\eta P_0(n) + \lambda P_1(n-1) + (n-1)\mu P_1(n+1)$$
$$= (n\mu + \lambda)P_1(n), \quad c-d+2 \le n \le c-1,$$

$$\eta P_0(c) + \lambda P_1(c-1) + (c\mu + \alpha)P_1(c+1) = (c\mu + \lambda b_c)P_1(c),$$

$$\eta p_0(n) + \lambda b_{n-1}P_1(n-1) + [c\mu + (n+1-c)\alpha]P_1(n+1)$$
$$= [c\mu + \lambda b_n + (n-c)\alpha]P_1(n), \quad c < n < N,$$

$$\eta P_0(N) + \lambda b_{N-1}P_1(N-1) = [c\mu + (N-c)\alpha]P_1(N),$$

$$\sum_{n=0}^{N} P_0(n) + \sum_{n=c-d+1}^{N} P_1(n) = 1.$$

3.2 Matrix solution

In this subsection, we derive the steady-state probabilities by using the matrix analytical method. Let

$$\boldsymbol{P} = (P_0(0), P_0(1), ..., P_0(N), P_1(c-d+1), P_1(c-d+2), ..., P_1(N))$$

be the steady-state probability vector. Then, the steady-state probability equations above can be rewritten as the matrix form as follows:

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

where $\boldsymbol{e} = (1,1,...1)^T$ is a $(2N-c+d+1) \times 1$ vector, and the transition rate matrix $\boldsymbol{Q}$ of the Markov process has the following blocked matrix structure:

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \boldsymbol{0} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \boldsymbol{A}_{23} \\ \boldsymbol{A}_{31} & \boldsymbol{0} & \boldsymbol{A}_{33} \end{pmatrix}.$$

Each matrix $\boldsymbol{A}_{ij}$ $(i,j=1,2,3)$ is given as follows:

$$\boldsymbol{A}_{11}=\begin{pmatrix} s_0 & \lambda & 0 & \cdots & 0 & 0 & 0 \\ t_1 & s_1 & \lambda & \cdots & 0 & 0 & 0 \\ 0 & t_2 & s_2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & t_{c-d-1} & s_{c-d-1} & \lambda \\ 0 & 0 & 0 & \cdots & 0 & t_{c-d} & s_{c-d} \end{pmatrix},$$

$$\boldsymbol{A}_{12}=\begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \\ \lambda b_{c-d} & 0 & \cdots & 0 \end{pmatrix}, \quad \boldsymbol{A}_{21}=\begin{pmatrix} 0 & 0 & \cdots & t_{c-d+1} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

$$\boldsymbol{A}_{22}=\begin{pmatrix} s_{c-d+1} & \lambda b_{c-d+1} & 0 & \cdots & 0 & 0 & 0 \\ t_{c-d+2} & s_{c-d+2} & \lambda b_{c-d+2} & \cdots & 0 & 0 & 0 \\ 0 & t_{c-d+3} & s_{c-d+3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & t_{N-1} & s_{N-1} & \lambda b_{N-1} \\ 0 & 0 & 0 & \cdots & 0 & t_N & s_N \end{pmatrix},$$

$$\boldsymbol{A}_{23}=\eta \boldsymbol{I}, \quad \boldsymbol{A}_{31}=\begin{pmatrix} 0 & \cdots & 0 & v_1 \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix},$$

$$\boldsymbol{A}_{33}=\begin{pmatrix} u_1 & \lambda & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ v_2 & u_2 & \lambda & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & v_3 & u_3 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & u_{d-1} & \lambda & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & v_d & u_d & \lambda b_c & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & v_{d+1} & u_{d+1} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & u_{N+d-c-1} & \lambda b_{N-1} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & v_{N+d-c} & u_{N+d-c} \end{pmatrix}$$

where $\boldsymbol{A}_{11}$ is a $(c-d+1)\times(c-d+1)$ square matrix, $\boldsymbol{A}_{12}$ is a $(c-d+1)\times(N-c+d)$ matrix, $\boldsymbol{A}_{21}$ is an $(N-c+d)\times(c-d+1)$ matrix, $\boldsymbol{A}_{22}$, $\boldsymbol{A}_{23}$ and $\boldsymbol{A}_{33}$ are $(N-c+$

$d) \times (N-c+d)$ square matrixes, $\boldsymbol{A}_{31}$ is an $(N-c+d) \times (c-d+1)$ matrix, $\boldsymbol{I}$ is an $(N-c+d) \times (N-c+d)$ identity matrix, and

$$s_i = \begin{cases} -(\lambda+i\mu), & 0 \le i \le c-d-1 \\ -[\lambda b_{c-d}+(c-d)\mu], & i=c-d \\ -[\eta+\lambda b_i+(c-d)\mu+(i+d-c)\alpha], & c-d+1 \le i \le N-1 \\ -[\eta+(c-d)\mu+(N+d-c)\alpha], & i=N, \end{cases}$$

$$t_i = \begin{cases} i\mu, & 1 \le i \le c-d \\ (c-d)\mu+(i-c+d)\alpha, & c-d+1 \le i \le N, \end{cases}$$

$$u_i = \begin{cases} -[\lambda+(c-d+i)\mu], & 1 \le i \le d-1 \\ -[\lambda b_{i+c-d}+c\mu+(i-d)\alpha], & d \le i \le N+d-c-1 \\ -[c\mu+(N-c)\alpha], & i=N+d-c, \end{cases}$$

$$v_i = \begin{cases} (c-d+i\mu), & 1 \le i \le c \\ c\mu+(i-d)\alpha, & d+1 \le i \le N+d-c. \end{cases}$$

In the following, we derive the steady-state probabilities from Eq. (1). To accommodate the partitioned blocked structure of $\boldsymbol{Q}$, we partition the steady-state probability vector into the segments accordingly as

$$\boldsymbol{P} = (\boldsymbol{P}_{00}, \boldsymbol{P}_{01}, \boldsymbol{P}_1)$$

where

$$\boldsymbol{P}_{00} = (P_0(0), P_0(1), ..., P_0(c-d)),$$

$$\boldsymbol{P}_{01} = (P_0(c-d+1), P_0(c-d+2), ..., P_0(N)),$$

$$\boldsymbol{P}_1 = (P_1(c-d+1), P_1(c-d+2), ..., P_1(N)).$$

Based on these partitions of the vector $\boldsymbol{P}$, Eq. (1) can be rewritten as follows:

$$\boldsymbol{P}_{00}\boldsymbol{A}_{11} + \boldsymbol{P}_{01}\boldsymbol{A}_{21} + \boldsymbol{P}_1\boldsymbol{A}_{31} = 0, \tag{2}$$

$$\boldsymbol{P}_{00}\boldsymbol{A}_{12} + \boldsymbol{P}_{01}\boldsymbol{A}_{22} = 0, \tag{3}$$

$$\eta \boldsymbol{P}_{01} + \boldsymbol{P}_1\boldsymbol{A}_{33} = 0, \tag{4}$$

$$\boldsymbol{P}_{00}\boldsymbol{e}_0 + \boldsymbol{P}_{01}\boldsymbol{e}_1 + \boldsymbol{P}_1\boldsymbol{e}_1 = 1 \tag{5}$$

where the vector $\boldsymbol{e}_0 = (1,1,...,1)^T$ is a $(c-d+1) \times 1$ vector, and the vector $\boldsymbol{e}_1 = (1,1,...,1)$ is an $(N+d-c) \times 1$ vector.

In order to solve the equations above by using the blocked matrix method, we prove that the matrixes $\boldsymbol{A}_{11}$, $\boldsymbol{A}_{22}$, and $\boldsymbol{A}_{33}$ are invertible. We have the following lemmas.

Lemma 1. *The matrixes $\boldsymbol{A}_{11}$ and $\boldsymbol{A}_{33}$ are invertible.*

Proof. By applying the properties of the determinant of transformation, it is easy to show that the determinant of matrix

$$|\boldsymbol{A}_{11}| = (-\lambda)^{c-d+1} b_{c-d} \neq 0$$

and the determinant of matrix

$$|\boldsymbol{A}_{33}| = (-1)^{N+d-c} \prod_{i=c-d+1}^{c} i\mu \prod_{j=1}^{N-c} (c\mu + j\alpha) \neq 0.$$

Hence, the matrixes $\boldsymbol{A}_{11}$ and $\boldsymbol{A}_{33}$ are invertible. □

Lemma 2. *Let $\boldsymbol{A} = (a_{ij})$ be an $n \times n$ square matrix. If $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ for $i = 1, 2, ..., n$, then $|\boldsymbol{A}| \neq 0$.*

Proof. Suppose that $|\boldsymbol{A}| = 0$, then the matrix equation $\boldsymbol{Ax} = \boldsymbol{0}$ has a non-trivial solution, that is, a non-zero column vector $\boldsymbol{k} = (k_1, k_2, ..., k_n)^T$ that satisfies $\boldsymbol{Ak} = \boldsymbol{0}$. Let

$$|k_i| = \max\{|k_1|, |k_2|, ..., |k_n|\},$$

then $|k_i| \neq 0$ since $\boldsymbol{k}$ is non-zero vector. From $\boldsymbol{Ak} = \boldsymbol{0}$, we have

$$\sum_{j=1}^{n} a_{ij} k_j = 0$$

or equivalently

$$a_{ii} = -\sum_{j \neq i} a_{ij} \frac{k_j}{k_i}.$$

Hence,

$$|a_{ii}| = |\sum_{j \neq i} a_{ij} \frac{k_j}{k_i}| \leq \sum_{j \neq i} |a_{ij}| |\frac{k_j}{k_i}| \leq \sum_{j \neq i} |a_{ij}|.$$

This contradicts the condition that $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ for $i = 1, 2, ..., n$. This contradiction shows that $|\boldsymbol{A}| \neq 0$. Hence, we complete the proof. □

Lemma 3. *The matrix $\boldsymbol{A}_{22}$ is invertible.*

Proof. Let a_{ij} be the (i,j) elements of matrix $\boldsymbol{A}_{22}$. Then, it is easy to see that

$$\begin{aligned}
|a_{11}| &= \eta + \lambda b_{c-d+1} + (c-d)\mu + \alpha \\
&> \lambda b_{c-d+1} = \sum_{j\neq 1} a_{1j}, \\
|a_{ii}| &= \eta + \lambda b_{c-d+i} + (c-d)\mu + i\alpha \\
&> \lambda b_{c-d+i} + (c-d)\mu + i\alpha = \sum_{j\neq i} a_{ij}, \quad 2 \le i \le N+d-c-1, \\
|a_{N+d-c}| &= \eta + (c-d)\mu + i(N+d-c)\alpha \\
&> (c-d)\mu + (N+d-c)\alpha = \sum_{j\neq N+d-c} a_{N+d-cj}.
\end{aligned}$$

By Lemma 2, $|\boldsymbol{A}_{22}| \neq 0$. Hence, we prove Lemma 3. □

We are now ready to derive the steady-state probabilities from Eqs. (2)-(5). Let $\boldsymbol{\varepsilon} = (1,0,...,0)$ be a $1\times(c-d+1)$ unit vector and $\boldsymbol{x} = \boldsymbol{\varepsilon}\boldsymbol{A}_{22}^{-1}$. Then, $\boldsymbol{x}$ is a first row vector of the matrix $\boldsymbol{A}_{22}^{-1}$. Hence, $\boldsymbol{x}$ is a $1\times(N-c+d)$ vector. Let $\boldsymbol{y} = \boldsymbol{x}\boldsymbol{A}_{33}^{-1}$. Then, $\boldsymbol{y}$ is a $1\times(N-c+d)$ vector. Let $\boldsymbol{\delta} = (0,...,0,1)$ be a $1\times(c-d+1)$ unit vector and $\boldsymbol{z} = \boldsymbol{\delta}\boldsymbol{A}_{11}^{-1}$. Then, $\boldsymbol{z}$ is a last row vector of the matrix $\boldsymbol{A}_{11}^{-1}$. Hence, $\boldsymbol{z}$ is a $1\times(c-d+1)$vector. We have the following theorem.

Theorem 4. *The segments of steady-state probability vector are given by*

$$\boldsymbol{P}_{00} = -K(\eta v_1 \boldsymbol{y} - t_{c-d+1}\boldsymbol{x})\boldsymbol{\varepsilon}^T \boldsymbol{z}, \tag{6}$$

$$\boldsymbol{P}_{01} = -K\boldsymbol{x}, \tag{7}$$

$$\boldsymbol{P}_1 = \eta K \boldsymbol{y} \tag{8}$$

where $\boldsymbol{\varepsilon}^T$ is the transpose of the unit vector $\boldsymbol{\varepsilon}$, and the constant

$$K = \left[(\boldsymbol{y}-\boldsymbol{x})\boldsymbol{e}_1 - (\eta v_1 \boldsymbol{y} - t_{c-d+1}\boldsymbol{x})\boldsymbol{\varepsilon}^T \boldsymbol{z}\boldsymbol{e}_0\right]^{-1}. \tag{9}$$

Proof. From Eq. (3), based on Lemma 3, we have

$$\begin{aligned}
\boldsymbol{P}_{01} &= -\boldsymbol{P}_{00}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1} \\
&= -(\lambda b_{c-d} P_0(c-d), 0, ..., 0)\boldsymbol{A}_{22}^{-1} \\
&= -\lambda b_{c-d} P_0(c-d)\boldsymbol{\varepsilon}\boldsymbol{A}_{22}^{-1}.
\end{aligned} \tag{10}$$

Let $K = \lambda b_{c-d} P_0(c-d)$, then Eq. (10) can be rewritten as

$$\boldsymbol{P}_{01} = -K\boldsymbol{x}. \tag{11}$$

Based on Lemma 1, substituting Eq. (11) into Eq. (4) yields

$$\boldsymbol{P}_1 = \eta K \boldsymbol{y}. \tag{12}$$

Then substituting Eqs. (11) and (12) into Eq. (2), we get

$$\boldsymbol{P}_{00} = -K(\eta \boldsymbol{y}\boldsymbol{A}_{31} - \boldsymbol{x}\boldsymbol{A}_{21})\boldsymbol{A}_{11}^{-1}. \tag{13}$$

By rewriting the matrix $\boldsymbol{A}_{21}$ and the matrix $\boldsymbol{A}_{31}$ as the partitioned blocked matrixes, we have

$$\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} = t_{c-d+1}\begin{pmatrix} \boldsymbol{\delta} \\ 0 \\ \vdots \\ 0 \end{pmatrix}\boldsymbol{A}_{11}^{-1} = t_{c-d+1}\begin{pmatrix} \boldsymbol{z} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = t_{c-d+1}\boldsymbol{\varepsilon}^T\boldsymbol{z} \tag{14}$$

and

$$\boldsymbol{A}_{31}\boldsymbol{A}_{11}^{-1} = v_1\begin{pmatrix} \boldsymbol{\delta} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = v_1\begin{pmatrix} \boldsymbol{z} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = v_1\boldsymbol{\varepsilon}^T\boldsymbol{z}. \tag{15}$$

Substituting Eqs. (14) and (15) into Eq. (13), we get

$$\boldsymbol{P}_{00} = -K(\eta v_1 \boldsymbol{y} - t_{c-d+1}\boldsymbol{x})\boldsymbol{\varepsilon}^T\boldsymbol{z}. \tag{16}$$

The constant K can be obtained as the expression given in Eq. (9) by substituting Eqs. (11), (12) and (16) into Eq. (5). Hence, the proof is completed. □

Remark 1. Based on Theorem 4, to compute the steady-state probability vector, we need mainly to compute the vectors $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$. Since it is not easy to get the expressions of the matrix $\boldsymbol{A}_{11}^{-1}$, the matrix $\boldsymbol{A}_{22}^{-1}$ and the matrix $\boldsymbol{A}_{33}^{-1}$, the vectors $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ are not explicitly given. However, these vectors $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ can be obtained by solving the linear equation $\boldsymbol{x}\boldsymbol{A}_{22} = \boldsymbol{\varepsilon}$, $\boldsymbol{y}\boldsymbol{A}_{33} = \boldsymbol{x}$ and $\boldsymbol{z}\boldsymbol{A}_{11} = \boldsymbol{\delta}$, respectively.

Remark 2. Note first that the results for M/M/1/N queueing system with balking, reneging and multiple server vacations are obtained by setting $c = 1$ and $d = 1$ in Eqs. (6)-(9). Eqs. (6)-(9) then correspond to the existing results in [13]. Next, note that the results for M/M/c/N queueing systems with synchronous vacations of partial servers are obtained by setting $b_n = 1$ and $\alpha = 0$ in Eqs. (6)-(9).

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 number of servers being on vacation.

4.1 Performance measures

Let $\boldsymbol{\varepsilon}_i$ be a $(c-d+1)\times 1$ unit vector with the $(i+1)$th element equal to 1 and the other elements equal to 0, $i=0,1,...,c-d$. Let $\boldsymbol{\delta}_i$ be a $(N-c+d)\times 1$ unit vector with the ith element equal to 1 and the other elements equal to 0, $i=1,2,...,N-c+d$. It is easy to see that

$$P_0(n)=\begin{cases} \boldsymbol{P}_{00}\boldsymbol{\varepsilon}_n, & 0\le n\le c-d, \\ \boldsymbol{P}_{01}\boldsymbol{\delta}_{n-c+d}, & c-d+1\le n\le N \end{cases} \tag{17}$$

and

$$P_1(n)=\boldsymbol{P}_1\boldsymbol{\delta}_{n-c+d}, \quad c-d+1\le n\le N \tag{18}$$

where $\boldsymbol{P}_{00}$, $\boldsymbol{P}_{01}$ and $\boldsymbol{P}_1$ are given in Theorem 4.

Based on the expressions of the steady-state probabilities given by Eqs. (17) and (18), we can obtain some steady-state performance measures of the system. For notational convenience, we define the following vectors:

$$\boldsymbol{\xi}_1=\sum_{n=1}^{c-d} n\boldsymbol{\varepsilon}_n, \quad \boldsymbol{\xi}_2=\sum_{n=c-d+1}^{N} n\boldsymbol{\delta}_{n-c+d},$$

$$\boldsymbol{\xi}_3=\sum_{n=c+1}^{N}(n-c)\boldsymbol{\delta}_{n-c+d}, \quad \boldsymbol{\xi}_4=\sum_{n=c-d+1}^{c} n\boldsymbol{\delta}_{n-c+d}+\sum_{n=c+1}^{N} c\boldsymbol{\delta}_{n-c+d},$$

$$\boldsymbol{\xi}_5=\sum_{n=c-d+1}^{N}(1-b_n)\boldsymbol{\delta}_{n-c+d}, \quad \boldsymbol{\xi}_6=\sum_{n=c}^{N}(1-b_n)\boldsymbol{\delta}_{n-c+d}$$

where $\boldsymbol{\xi}_1$ is a $(c-d+1)\times 1$ vector, $\boldsymbol{\xi}_2$, $\boldsymbol{\xi}_3$, $\boldsymbol{\xi}_4$, $\boldsymbol{\xi}_5$ and $\boldsymbol{\xi}_6$ are $(N-c+d)\times 1$ vectors.

Theorem 5. *The expected number of customers in the system is given by*

$$E(N)=\boldsymbol{P}_{00}\boldsymbol{\xi}_1+(\boldsymbol{P}_{01}+\boldsymbol{P}_1)\boldsymbol{\xi}_2. \tag{19}$$

The expected number of the waiting customers in the queue is given by

$$E(N_q)=\boldsymbol{P}_{01}[\boldsymbol{\xi}_2-(c-d)\boldsymbol{e}_1]+\boldsymbol{P}_1\boldsymbol{\xi}_3. \tag{20}$$

The expected number of servers on vacation is given by

$$E(V)=d[\boldsymbol{P}_{00}\boldsymbol{e}_0+\boldsymbol{P}_{01}\boldsymbol{e}_1]. \tag{21}$$

The expected number of busy servers in the system is given by

$$E(B)=\boldsymbol{P}_{00}\boldsymbol{\xi}_1+(c-d)\boldsymbol{P}_{01}\boldsymbol{e}_1+\boldsymbol{P}_1\boldsymbol{\xi}_4. \tag{22}$$

The expected number of idle servers in the system is given by

$$E(I)=\boldsymbol{P}_{00}[(c-d)\boldsymbol{e}_0-\boldsymbol{\xi}_1]+\boldsymbol{P}_1(c\boldsymbol{e}_1-\boldsymbol{\xi}_4). \tag{23}$$

Proof. Note that

$$E(N) = \sum_{n=0}^{N} nP_0(n) + \sum_{n=c-d+1}^{N} nP_1(n),$$

$$E(N_q) = \sum_{n=c-d+1}^{N} (n-c+d)P_0(n) + \sum_{n=c+1}^{N} (n-c)P_1(n),$$

$$E(V) = \sum_{n=0}^{N} dP_0(n).$$

Substituting $P_0(n)$ of Eq. (17) and $P_1(n)$ of Eq. (18) into the equations above, we obtain the results given by Eqs. (19), (20) and (21). It is easy to see that

$$E(B) = E(N) - E(N_q). \tag{24}$$

Substituting Eqs. (19) and (20) into the equation above yields Eq. (22). Note that

$$E(I) = c - E(B) - E(V), \tag{25}$$

the result for $E(I)$ given by Eq. (23) is obtained by substituting Eqs. (21) and (24) into the equations above. □

Theorem 6. *The average balking rate is given by*

$$B.R. = \lambda\{(1-b_{c-d})\boldsymbol{P}_{00}\boldsymbol{\varepsilon}_{c-d} + \boldsymbol{P}_{01}\boldsymbol{\xi}_5 + \boldsymbol{P}_1\boldsymbol{\xi}_6\}. \tag{26}$$

The average reneging rate is given by

$$R.R. = \alpha\{\boldsymbol{P}_{01}[\boldsymbol{\xi}_2 - (c-d)\boldsymbol{e}_0] + \boldsymbol{P}_1\boldsymbol{\xi}_3\}. \tag{27}$$

The average rate of customer loss is given by

$$\begin{aligned} L.L. = \lambda\{(1-b_{c-d})\boldsymbol{P}_{00}\boldsymbol{\varepsilon}_{c-d} + \boldsymbol{P}_{01}\boldsymbol{\xi}_5 + \boldsymbol{P}_1\boldsymbol{\xi}_6\} \\ +\alpha\{\boldsymbol{P}_{01}[\boldsymbol{\xi}_2 - (c-d)\boldsymbol{e}_0] + \boldsymbol{P}_1\boldsymbol{\xi}_3\}. \end{aligned} \tag{28}$$

Proof. Since the probability that a customer balks in the system is $1-b_n$ when the customer on arrival finds n customers in the system, then the instantaneous balking rate is $\lambda(1-b_n)$. Following the model of Ancker and Gafarian [6], the average balking rate $B.R.$ is given by

$$B.R. = \sum_{n=c-d}^{N} \lambda(1-b_n)P_0(n) + \sum_{n=c}^{N} \lambda(1-b_n)P_1(n).$$

Substituting $P_0(n)$ of Eq. (17) and $P_1(n)$ of Eq. (18) into the equation above yields Eq. (26).

If there are n customers in the system and i servers available, then there are $n-i$ waiting customers in the queue. Since any one of the $n-i$ customers in the queue may renege, the instantaneous reneging rate is $(n-i)\alpha$. Again, following the model of Ancker and Gafarian [6], the average reneging rate $R.R.$ is given by

$$R.R. = \sum_{n=c-d+1}^{N} (n-c+d)\alpha P_0(n) + \sum_{n=c+1}^{N} (n-c)\alpha P_1(n) = \alpha E(N_q).$$

Substituting $E(N_q)$ of Eq. (20) yields Eq. (27).

The average rate of customer loss, $L.R.$, is simply the sum of the average balking rate and the average reneging rate. Thus, we have

$$L.R. = B.R. + R.R.$$

Then, Eq. (28) is obtained by substituting Eqs. (26) and (27) into the equation above. □

4.2 Cost model

In this subsection, we develop an expected cost function where the number of servers on vacation d is a decision variable. Our objective is to determine the optimum number of servers on vacation d^* to minimize this cost function.

Let

$C_1 \equiv$ cost per unit time when one server is on vacation,

$C_2 \equiv$ cost per unit time when one server is idle,

$C_3 \equiv$ cost per unit time when one server is busy,

$C_4 \equiv$ cost per unit time when one customer is waiting for service,

$C_5 \equiv$ cost per unit time when one customer joins the system and is served,

$C_6 \equiv$ cost per unit time when a customer balks or reneges.

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

$$F(d) = C_1E(V) + C_2E(I) + C_3E(B) + C_4E(N_q) + C_5(E(N) - E(N_q)) + C_6L.L.$$

or equivalently

$$F(d) = C_1E(V) + C_2E(I) + (C_3 + C_5)E(B) + C_4E(N_q) + C_6L.L.$$

where $E(V)$, $E(I)$, $E(B)$ and $E(N_q)$ are given in Theorem 5, and $L.L.$ is given in Theorem 6.

5 Numerical Results

In this subsection, we perform a sensitivity analysis on the optimal number of servers on vacation d^* and its expected cost $F(d^*)$ based on changes in the values of the system parameters such as the arrival rate λ, the service rate μ, the reneging rate α and the vacation rate η.

Since the decision variable d is an integer value and the expected cost function is non-linear and complex, it would be an arduous task to derive analytic results for the optimal value d^*. Thus, we use a heuristic approach to obtain the optimal value d^* which is determined by satisfying the following inequality:

$$F(d^*-1) > F(d^*) < F(d^*+1).$$

We set the system capacity $N = 10$ and the number of servers $c = 4$. We select the probability $b_0 = 0.95$ and $b_n = 1 - n/N$ for $1 \leq n \leq N$ by referencing [6], and employ the following cost elements by referencing [9]: $C_1 = 100$, $C_2 = 110$, $C_3 = 120$, $C_4 = 150$, $C_5 = 130$ and $C_6 = 140$. The numerical results for the optimal value d^* and the optimal cost $F(d^*)$ are illustrated in Figs. 1-4.

In Fig. 1, we fix $\mu = 1.0$, $\eta = 0.3$ and $\alpha = 0.2$, and display the optimal critical value d^* and the optimal cost $F(d^*)$ by varying the arrival rate λ. Fig. 1 shows that the optimal value d^* decreases as λ increases, while the optimal cost $F(d^*)$ increases significantly as λ increases.

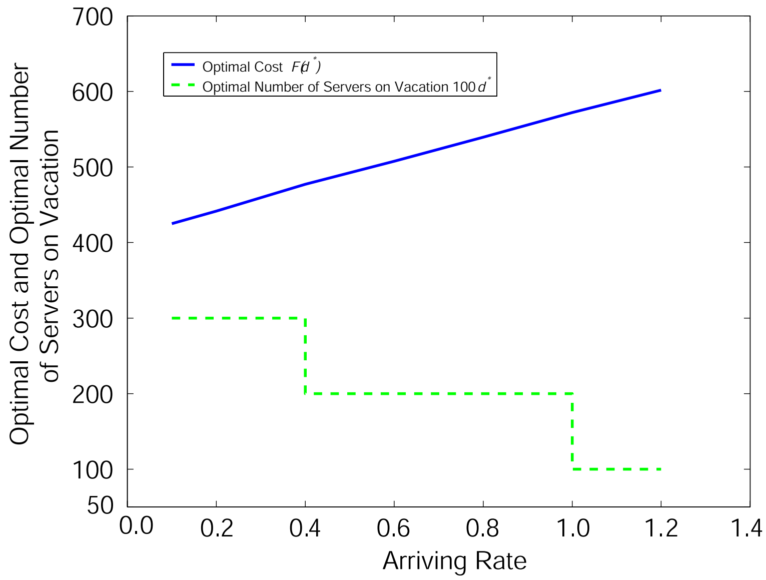

Figure 1: Optimal cost $F(d^*)$ and optimal value d^* versus arrive rate λ with $\mu_1 = 1.0$, $\eta = 0.3$ and $\alpha = 0.2$.

In Fig. 2, we fix $\lambda = 0.5$, $\eta = 0.3$ and $\alpha = 0.2$, and display the optimal critical value d^* and the optimal cost $F(d^*)$ by varying the service rate μ. Fig. 2 shows that the optimal value d^* increases as μ increases, while the optimal cost $F(d^*)$ decreases significantly as μ increases.

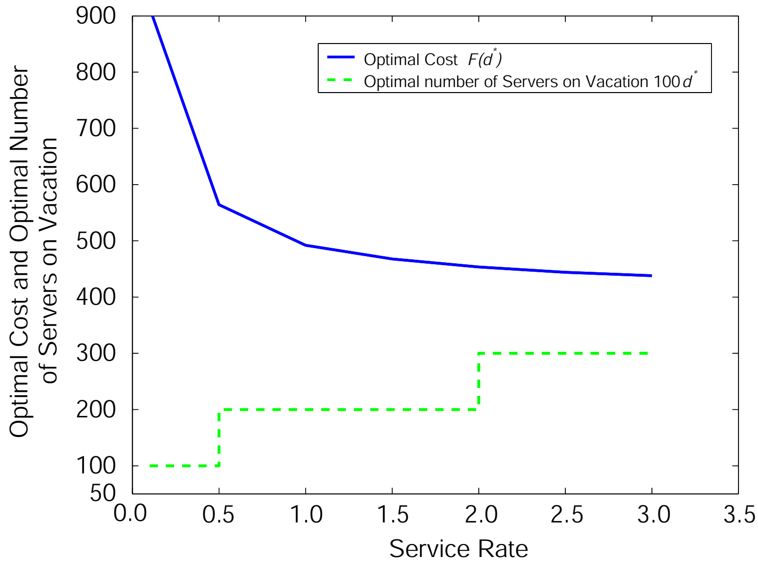

Figure 2: Optimal cost $F(d^*)$ and optimal value d^* versus service rate μ with $\lambda = 0.5$, $\eta = 0.3$ and $\alpha = 0.2$.

In Fig. 3, we fix $\lambda o.5$, $\mu = 1$ and $\eta = 0.3$, and display the optimal critical value d^* and the optimal cost $F(d^*)$ by varying the reneging rate α. Fig. 3 shows that the optimal value d^* increases as α increases, while the optimal cost $F(d^*)$ decreases slightly as α increases.

In Fig. 4, we fix $\lambda = 0.5$, $\mu = 1.0$ and $\alpha = 0.2$, and display the optimal critical value d^* and the optimal cost $F(d^*)$ by varying the vacation rate η. Fig. 4 shows that the optimal value d^* increases as η increases, while the optimal cost $F(d^*)$ decreases slightly as η increases.

It appears from Figs. 1-4 that: (i) The arrival rate λ and the service rate μ affect the optimal value d^* and the optimal cost $F(d^*)$ significantly. (ii) The reneging rate α and the vacation rate η affect the optimal value d^* and the optimal cost $F(d^*)$ slightly.

6 Conclusions

In this paper, we have considered an M/M/c/N queueing system with balking, reneging and synchronous vacations of partial servers. We have developed the equations of the steady state probabilities and derived the matrix form solution for the steady-state probability vector. Based on the steady-state probability vector, we have obtained some performance measures of the system and formulated a cost model to determine the optimal number of servers on vacation. Furthermore, we have performed sensitivity analysis through numerical experiments.

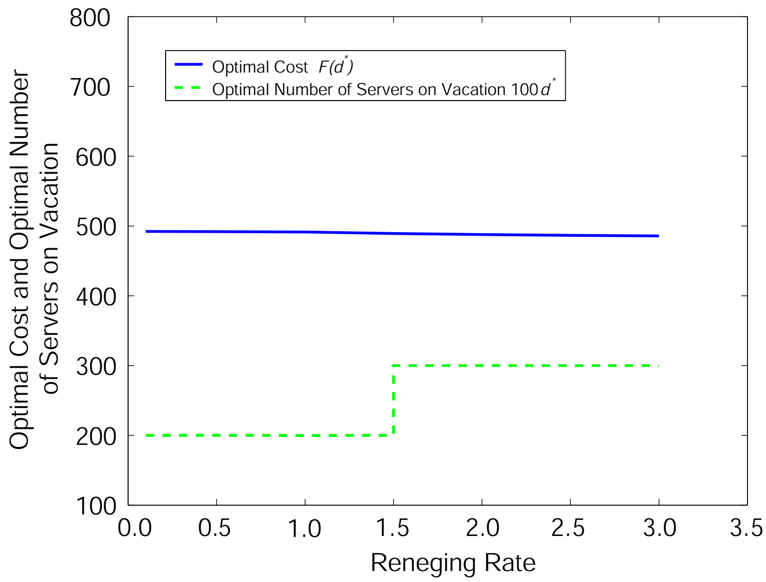

Figure 3: Optimal cost $F(d^*)$ and optimal value d^* versus reneging rate α with $\lambda = 0.5$, $\mu = 1.0$ and $\eta = 0.3$.

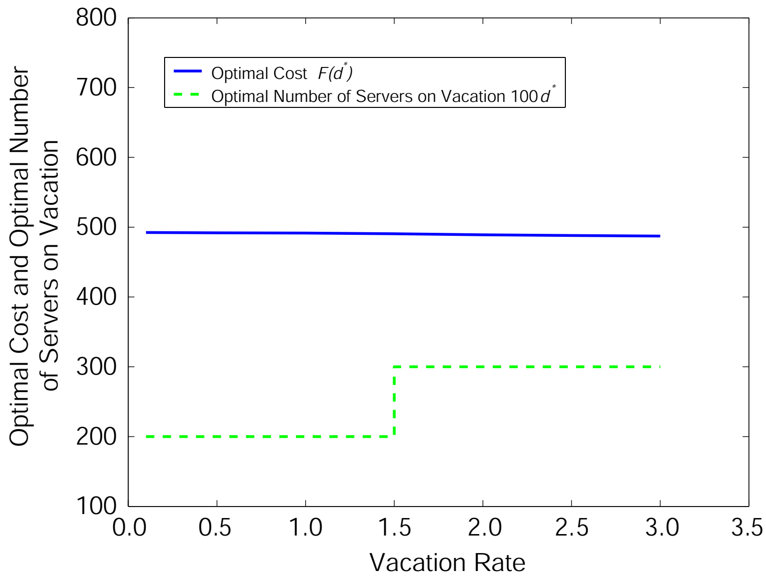

Figure 4: Optimal cost $F(d^*)$ and optimal value d^* versus vacation rate η with $\lambda = 0.5$, $\mu = 1.0$ and $\alpha = 0.2$.

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] J-C. Ke and K-H. Wang. The reliability analysis of balking and reneging in a repairable system with warm standbys. *Quality and reliability Engineering International*, 18, 211–218, 2002.

[2] A. Shawky. The machine interference model: M/M/C/K/N with balking, reneging and spares. *Opsearch*, 37, 25–35, 2000.

[3] Jr., C. Ancker and A. Gafarian. Queueing with impatient customers who leave at random. *Journal of Industry Engineering*, XIII, 84–90, 1962.

[4] F. Haight. Queueing with balking. *Biometrika*, 44, 360–369, 1957.

[5] F. Haight. Queueing with reneging. *Metrika*, 2, 186–197, 1959.

[6] Jr., C. Ancker and A. Gafarian, Some queueing problems with balking and reneging: I. *Operations Research*, 11, 88–100, 1963.

[7] Jr., C. Ancker and A. Gafarian. Some queueing problems with balking and reneging: II. *Operations Research*, 11, 928–937, 1963.

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

[9] K-H. Wang and 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(2), 169–180, 2002.

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

[11] B. Doshi. Single server queues with vacations. In, H. Takagi (Ed.), *Stochastic Analysis of Computer and Communication Systems*, North-Holland, 217–265, 1990.

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

[13] Y. Zhang, D. Yue and W. Yue. Analysis of an M/M/1/N queue with balking, reneging and server vacations. *Proc. of the Fifth International Symposium*, 37–47, 2005.

[14] Z. Zhang and N. Tian. Analysis on queueing systems with synchronous vacations of partial servers. *Performance Evaluation*, 52, 269–282, 2003.