

Research Article

Analysis and Comparison of Queue with N -Policy and Unreliable Server

Qing Ma¹ and Xuelu Zhang² 

¹School of Data and Computer Science, Shandong Women's University, 250300 Jinan, China

²School of Mathematics and Statistics, Qilu University of Technology (Shandong Academy of Sciences), 250353 Jinan, China

Correspondence should be addressed to Xuelu Zhang; zhang_xuelu@163.com

Received 30 November 2019; Accepted 5 May 2020; Published 18 May 2020

Academic Editor: Xiaohua Ding

Copyright © 2020 Qing Ma and Xuelu Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This work considers a queueing system with N -policy and unreliable server, where potential customers arrive at the system according to Poisson process. If there is no customer waiting in the system, instead of shutting down, the server turns into dormant state and does not afford service until the number of customers is accumulated to a certain threshold. And in the working state, the server is apt to breakdown and affords service again only after it is repaired. According to whether the server state is observable or not, the numerical optimal arrival rates are computed to maximize the social welfare and throughput of the system. The results illustrate their tendency in two cases so that the manager has a strong ability to decide which is more crucial in making management decision.

1. Introduction

Queueing system models are widely used in other research areas, such as cognitive radio networks, signal transmission system, traffic system, and service-inventory system. The customer, the basic element of these systems, enters the system to seek service, and the service manager operates the system for the customer. From the point of the manager, he regards all the customers as a whole and expects to reach more social welfare as one of his ultimate goal. However, for an overcongested system, increasing the throughput and relieving the system congestion are more crucial. This paper considers a queueing model with N -policy vacation and unreliable server, analyzes the social welfare and throughput with respect to system parameters, and determines in which situation the social welfare and throughput have the same tendency.

Information level plays an important role in analyzing the customer behavior and manager decision. Naor [1] first investigated customer's behavior from the economic viewpoint in an $M/M/1$ queueing system, established the social welfare function, and obtained the optimal threshold

strategy when the queue length is observable to customers. Edelson and Hilderbrand [2] considered the same queueing model in unobservable cases and found the social optimal joining probability. Guo and Zipkin [3] analyzed the effects of different levels of information on customers and the overall system and focused on which delay information can improve performance or hurt the provider or the customers. Recently, many researchers focus on the influence of information level on kinds of queueing systems.

For an unreliable server, Economou and Kanta [4] obtained the optimal threshold of customer in fully and almost observable cases. For an N -policy vacation queue, Guo and Hassin [5] found the social optimal arrival rate and optimal threshold N , respectively, in observable and unobservable cases. Wang et al. [6] analyzed the strategic behavior of the primary user (PU) and secondary users (SUs) in the different information levels. Zhang and Wang [7] considered a make-to-stock production system where the system operates with different service rates. And the customer makes the decision on whether to stay for the product or leave without purchase on the basis of the system information. Sun et al. [8] considered a Markovian queue with

single vacation and abandonment opportunities in four information levels and analyzed the behavior of the customer along with the effect of abandonment on them.

Different from the strategic behavior of the customer, many papers also investigate the social optimal behavior from the social manager's preference. Sun et al. [9] considered site-clearing and non-site-clearing Markovian queues with server failures and unreliable repairer and compared the equilibrium and social optimal arrival rates along with the degree of deviation. In addition, Sun et al. [10] considered exhaustive and nonexhaustive $M/M/1/N$ queues with working vacation and threshold policy and compared the optimal arrival rate, optimal threshold, and corresponding social welfare. Wang and Wang [11] separated customers into two streams, so only part of them can observe the server state in retrial queue. According to the fraction of the observed customers, the authors determined the situation when the social welfare and throughput can reach maximum. Cui et al. [12] developed a queueing-game-theoretic model to analyze the interaction among customers, the line-sitting firm, and the service provider. And they also compared and examined the difference of line-sitting and selling priority on them. Therefore, the information level plays an important role to analyze the behavior of the customer and social manager.

In this paper, we consider a vacation queueing model with N -policy where the unreliable server has three different states. According to whether the server state is observed or not, we investigate the tendency of social welfare and throughput with information levels and system parameters. The results will offer managerial guidelines for the manager, and the main contributions can be summarized as follows: First, this paper considers an $M/M/1$ queueing system with an N -policy and unreliable server. Second, for observable and unobservable cases, we establish social welfare function to seek the optimal arrival rate along with the corresponding throughput. Third, by particle swarm optimization algorithm, numerical results illustrate that social welfare and throughput may not have the same tendency, and we get new insights into how to operate the queueing system for achieving more benefit and throughput.

This paper is organized as follows. The model description is given in Section 2. Section 3 presents the main performance measures in unobservable and observable cases. Finally, Sections 4 and 5 give the numerical experiments and conclusions, respectively.

2. Model Description

This paper considers an $M/M/1$ queueing system where all the customers arrive according to a Poisson process with rate Λ and the service discipline is first-come, first-served (FCFS). The service times are all assumed to be exponentially distributed with rate μ . Once the system is empty, the server turns into dormant state and the vacation time is exponentially distributed with rate α . In the dormant state, the server does not afford service until the number of customers in the system is accumulated to a

predetermined constant N . Besides this, the server is liable to break down and its lifetime is exponentially distributed with rate β . In the period of breakdown, customers always can join the system, but the server does not afford service similar to the dormant state. Assume that the repair times are also exponentially distributed with rate γ . Without loss of generality, the interarrival times, the service times, the lifetimes, and the repair times are mutually independent.

The system state is specified at time t by the pair $(I(t), J(t))$, where $I(t)$ denotes the number of customers in the system and $J(t)$ denotes the state of the server. Concretely, $J(t) = 0$ means the server is in the dormant state and does not afford service; $J(t) = 1$ means the server is in the working state with service rate μ ; and $J(t) = 2$ means the server is in the breakdown state and also does not afford service. It is clear that the process $\{(I(t), J(t)), t \geq 0\}$ is a continuous time Markov chain with state space $\Omega = \{(i, 0), 0 \leq i \leq N - 1; (i, j), i \geq 0, j = 1, 2\}$.

Being served, the customer can obtain a reward of R units, and a waiting cost of c units per unit time is continuously accumulated from the time he enters the system to the time he completes the service. In this paper, the number of customers in the system is unobservable, and according to whether the state of the server is observable or not to customer upon arrival, the following two cases are investigated in this paper:

- (i) Unobservable case: customers cannot observe the state of the server
- (ii) Observable case: customers can observe the state of the server

3. Main Results

Since the number of customers is unobservable, this section considers the mixed joining strategy of the customer and presents the main performance measures of this system in two different situations. Based on these measures, it can help to establish the social welfare and throughput function to find the optimal arrival rate.

3.1. Unobservable Case. In this part, all the customers cannot observe the state of the server. Since all the customers are assumed to be homogeneous, this section considers the following mixed strategy: "customer joins the system with probability q ; balks with probability $1 - q$." And for simplicity, denote $\Lambda q = \lambda$. The corresponding transition rate diagram is shown in Figure 1.

To discuss this situation, the steady-state probabilities of the system are necessary and the preliminary results are summarized in the following proposition. It is readily proved that the system is stable if and only if $\mu\gamma - \lambda(\beta + \gamma) > 0$.

Proposition 1. *In the unobservable case, define the steady-state probabilities of the system state as $\{p_u(i, j) : (i, j) \in \Omega\}$. Then, the steady-state probabilities of the server in states 0, 1, and 2, respectively, are*

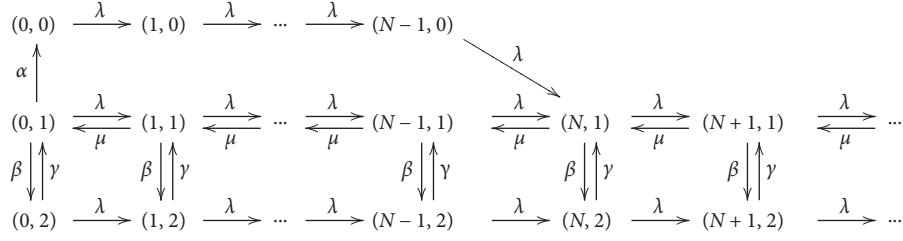


FIGURE 1: Transition rate diagram of the system in the unobservable case.

$$\begin{aligned}
 P_0 &= N p_u(0, 0), \\
 P_1 &= \frac{\lambda \gamma (\mu + \alpha N)}{\alpha [\mu \gamma - \lambda (\beta + \gamma)]} p_u(0, 0), \\
 P_2 &= \frac{\lambda \beta (\mu + \alpha N)}{\alpha [\mu \gamma - \lambda (\beta + \gamma)]} p_u(0, 0),
 \end{aligned} \tag{1}$$

where $p_u(0, 0) = \alpha [\mu \gamma - \lambda (\beta + \gamma)] / \mu [\lambda (\beta + \gamma) + \alpha \gamma N]$.

Proof. From the first line of Figure 1, the following steady-state equations

$$\begin{aligned}
 \lambda p_u(0, 0) &= \alpha p_u(0, 1), \\
 \lambda p_u(i, 0) &= \lambda p_u(i-1, 0), \quad i = 1, \dots, N-1,
 \end{aligned} \tag{2}$$

show that

$$\begin{aligned}
 p_u(0, 0) &= p_u(1, 0) = \dots = p_u(N-1, 0), \\
 p_u(0, 0) &= \frac{\alpha}{\lambda} p_u(0, 1).
 \end{aligned} \tag{3}$$

From the second and third lines of Figure 1, the following equations are at hand:

$$(\lambda + \alpha + \beta) p_u(0, 1) = \mu p_u(1, 1) + \gamma p_u(0, 2), \tag{4}$$

$$\begin{aligned}
 (\lambda + \mu + \beta) p_u(i, 1) &= \mu p_u(i+1, 1) + \lambda p_u(i-1, 1) \\
 &\quad + \gamma p_u(i, 2), \quad i = 1, \dots, N-1,
 \end{aligned} \tag{5}$$

$$\begin{aligned}
 (\lambda + \mu + \beta) p_u(N, 1) &= \mu p_u(N+1, 1) + \lambda p_u(N-1, 1) \\
 &\quad + \gamma p_u(N, 2) + \lambda p_u(N-1, 0),
 \end{aligned} \tag{6}$$

$$\begin{aligned}
 (\lambda + \mu + \beta) p_u(i, 1) &= \mu p_u(i+1, 1) + \lambda p_u(i-1, 1) \\
 &\quad + \gamma p_u(i, 2), \quad i = N+1, \dots,
 \end{aligned} \tag{7}$$

$$(\lambda + \gamma) p_u(0, 2) = \beta p_u(0, 1), \tag{8}$$

$$(\lambda + \gamma) p_u(i, 2) = \beta p_u(i, 1) + \lambda p_u(i-1, 2), \quad i = 1, \dots \tag{9}$$

Define the following partial generating functions as

$$\begin{aligned}
 P_0(z) &= \sum_{i=0}^{N-1} p_u(i, 0) z^i, \\
 P_1(z) &= \sum_{i=0}^{\infty} p_u(i, 1) z^i, \\
 P_2(z) &= \sum_{i=0}^{\infty} p_u(i, 2) z^i.
 \end{aligned} \tag{10}$$

From equation (3), it leads to

$$P_0(z) = \sum_{i=0}^{N-1} p_u(0, 0) z^i = \frac{1 - z^N}{1 - z} p_u(0, 0). \tag{11}$$

Multiplying $p_u(i, 1)$ by z^i , summing i from 0 to ∞ , and considering equations (3)–(7) yield

$$\begin{aligned}
 (\lambda + \mu + \beta) P_1(z) + (\alpha - \mu) p_u(0, 1) &= \frac{\mu}{z} (P_1(z) - p_u(0, 1)) \\
 &\quad + \lambda z P_1(z) + \gamma P_2(z) \\
 &\quad + \lambda p_u(0, 0) z^N.
 \end{aligned} \tag{12}$$

Considering equation (8) and multiplying equation (9) by z^i for all $i \geq 0$ yield

$$(\lambda + \gamma) P_2(z) = \beta P_1(z) + \lambda z P_2(z). \tag{13}$$

Solving $P_1(z)$ from equation (13) and putting it into (12), we obtain

$$P_2(z) = \frac{\lambda \beta [\mu (z-1) + \alpha z (z^N - 1)]}{\alpha (\lambda + \gamma - \lambda z) [(\lambda + \mu + \beta) z - \mu - \lambda z^2] - \alpha \beta \gamma z} p_u(0, 0), \tag{14}$$

$$\begin{aligned}
 P_1(z) &= \frac{1}{\beta} (\lambda + \gamma - \lambda z) P_2(z) \\
 &= \frac{\lambda (\lambda + \gamma - \lambda z) [\mu (z-1) + \alpha z (z^N - 1)]}{\alpha (\lambda + \gamma - \lambda z) [(\lambda + \mu + \beta) z - \mu - \lambda z^2] - \alpha \beta \gamma z} p_u(0, 0).
 \end{aligned} \tag{15}$$

From equations (11), (14), (15) and $\lim_{z \rightarrow 1} [P_0(z) + P_1(z) + P_2(z)] = 1$, the following probabilities are at hand:

$$\begin{aligned}
P_0 &= \lim_{z \rightarrow 1} P_0(z), \\
P_1 &= \lim_{z \rightarrow 1} P_1(z), \\
P_2 &= \lim_{z \rightarrow 1} P_2(z), \\
p_u(0,0) &= \frac{\alpha[\mu\gamma - \lambda(\beta + \gamma)]}{\mu[\lambda(\beta + \gamma) + \alpha\gamma N]}.
\end{aligned} \tag{16}$$

This completes the proof. ■

After getting $P_0(z)$, $P_1(z)$, $P_2(z)$, the expected number of customers in system $E[N_u]$ can be computed as follows:

$$\begin{aligned}
E[N_u] &= \left[\frac{dP_0(z)}{dz} + \frac{dP_1(z)}{dz} + \frac{dP_2(z)}{dz} \right]_{z=1} \\
&= p_u(0,0) \left\{ \frac{N(N-1)}{2} + \frac{\lambda \text{Num}_0}{2\alpha[\mu\gamma - \lambda(\beta + \gamma)]^2} \right\},
\end{aligned} \tag{17}$$

where $\text{Num}_0 = \alpha(\beta + \gamma)[\mu\gamma - \lambda(\beta + \gamma)]N^2 + \alpha[\lambda(\beta + \gamma)^2 + \mu(\gamma^2 + \beta\gamma + 2\lambda\beta)]N + 2\lambda\mu[(\beta + \gamma)^2 + \beta\mu]$.

According to Little's formula, the expected sojourn time of the customer in the system is $E[W_u] = E[N_u]/\lambda$. Moreover, the social welfare function is $S_u(\lambda) = \lambda(R - cE[W_u]) = \lambda R - cE[N_u]$, and the system throughput is $T_u(\lambda) = \lambda$.

3.2. Observable Case. In this section, all the customers can observe the state of the server. And according to the server states 0, 1, and 2, the customer chooses appropriate probabilities q_0 , q_1 , and q_2 to join the system. That is to say, "customer joins the system with probability q_i ($i = 0, 1, 2$); balks with probability $1 - q_i$." And for simplicity, denote $\Lambda q_i = \lambda_i$. The corresponding transition rate diagram is shown in Figure 2. Similar to Section 3.1, the system is stable if and only if $\mu\gamma - \lambda_1\gamma - \lambda_2\beta > 0$.

Proposition 2. *In the observable case, define the steady-state probabilities of the system state as $\{p_o(i, j) : (i, j) \in \Omega\}$. Then, the steady-state probabilities of the server in states 0, 1, and 2, respectively, are*

$$\begin{aligned}
Q_0 &= Np_o(0,0), \\
Q_1 &= \frac{\lambda_0\gamma(\mu + \alpha N)}{\alpha(\mu\gamma - \lambda_1\gamma - \lambda_2\beta)}p_o(0,0), \\
Q_2 &= \frac{\lambda_0\beta(\mu + \alpha N)}{\alpha(\mu\gamma - \lambda_1\gamma - \lambda_2\beta)}p_o(0,0),
\end{aligned} \tag{18}$$

where $p_o(0,0) = \alpha(\mu\gamma - \lambda_1\gamma - \lambda_2\beta)/\alpha N(\mu\gamma - \lambda_1\gamma - \lambda_2\beta) + \lambda_0(\mu + \alpha N)(\beta + \gamma)$.

Proof. From the first line of Figure 2, the following steady-state equations

$$\begin{aligned}
\lambda_0 p_o(0,0) &= \alpha p_o(0,1), \\
\lambda_0 p_o(i,0) &= \lambda_0 p_o(i-1,0), \quad i = 1, \dots, N-1,
\end{aligned} \tag{19}$$

show that

$$\begin{aligned}
p_o(0,0) &= p_o(1,0) = \dots = p_o(N-1,0), \\
p_o(0,1) &= \frac{\lambda_0}{\alpha} p_o(0,0).
\end{aligned} \tag{20}$$

From the second and third lines of Figure 2, the following equations are at hand:

$$(\lambda_1 + \alpha + \beta)p_o(0,1) = \mu p_o(1,1) + \gamma p_o(0,2), \tag{21}$$

$$\begin{aligned}
(\lambda_1 + \mu + \beta)p_o(i,1) &= \mu p_o(i+1,1) + \gamma p_o(i,2) \\
&\quad + \lambda_1 p_o(i-1,1), \quad i = 1, \dots, N-1,
\end{aligned} \tag{22}$$

$$\begin{aligned}
(\lambda_1 + \mu + \beta)p_o(N,1) &= \mu p_o(N+1,1) + \gamma p_o(N,2) \\
&\quad + \lambda_1 p_o(N-1,1) + \lambda_0 p_o(N-1,0),
\end{aligned} \tag{23}$$

$$\begin{aligned}
(\lambda_1 + \mu + \beta)p_o(i,1) &= \mu p_o(i+1,1) + \gamma p_o(i,2) \\
&\quad + \lambda_1 p_o(i-1,1), \quad i = N+1, \dots,
\end{aligned} \tag{24}$$

$$(\lambda_2 + \gamma)p_o(0,2) = \beta p_o(0,1), \tag{25}$$

$$(\lambda_2 + \gamma)p_o(i,2) = \beta p_o(i,1) + \lambda_2 p_o(i-1,2), \quad i = 1, \dots \tag{26}$$

Denote the corresponding partial generating functions as

$$\begin{aligned}
Q_0(z) &= \sum_{i=0}^{N-1} p_o(i,0)z^i, \\
Q_1(z) &= \sum_{i=0}^{\infty} p_o(i,1)z^i, \\
Q_2(z) &= \sum_{i=0}^{\infty} p_o(i,2)z^i.
\end{aligned} \tag{27}$$

Considering equation (20) yields

$$Q_0(z) = \frac{1 - z^N}{1 - z} p_o(0,0). \tag{28}$$

Considering equations (21)–(24) and multiplying $p_o(i,1)$ by z^i , $i = 0, 1, \dots$, yield

$$\begin{aligned}
(\lambda_1 + \mu + \beta)Q_1(z) + (\alpha - \mu)p_o(0,1) &= \frac{\mu}{z} [Q_1(z) - p_o(0,1)] \\
&\quad + \gamma Q_2(z) + \lambda_1 z Q_1(z) \\
&\quad + \lambda_0 p_o(0,0) z^N.
\end{aligned} \tag{29}$$

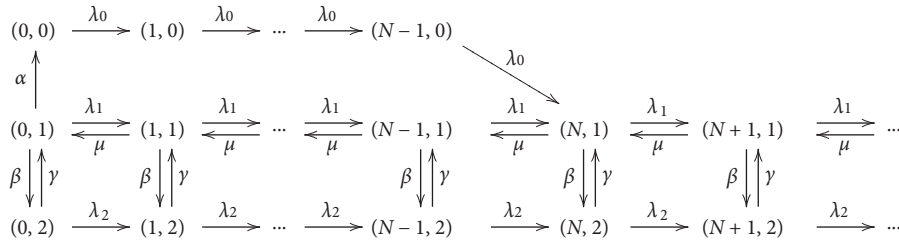


FIGURE 2: Transition rate diagram of the system in the observable case.

Similarly, considering equations (25)-(26) and multiplying $p_o(i, 2)$ by $z^i, i = 0, 1, \dots$, yield

$$(\lambda_2 + \gamma)Q_2(z) = \beta Q_1(z) + \lambda_2 z Q_2(z). \quad (30)$$

Solving $Q_1(z)$ from equation (30) and putting it into (29), we obtain

$$Q_2(z) = \frac{\lambda_0 \beta [(\mu - \alpha)z - \mu + \alpha z^{N+1}]}{\alpha \{[(\lambda_1 + \mu + \beta)z - \mu - \lambda_1 z^2][\lambda_2(1 - z) + \gamma] - \beta \gamma z\}} P_o(0, 0), \quad (31)$$

$$Q_1(z) = \frac{\lambda_0 [(\mu - \alpha)z - \mu + \alpha z^{N+1}][\lambda_2(1 - z) + \gamma]}{\alpha \{[(\lambda_1 + \mu + \beta)z - \mu - \lambda_1 z^2][\lambda_2(1 - z) + \gamma] - \beta \gamma z\}} P_o(0, 0). \quad (32)$$

Using equations (28), (31), (32) and $\lim_{z \rightarrow 1} [Q_0(z) + Q_1(z) + Q_2(z)] = 1$, the following probabilities show the results:

$$Q_0 = \lim_{z \rightarrow 1} Q_0(z),$$

$$Q_1 = \lim_{z \rightarrow 1} Q_1(z),$$

$$Q_2 = \lim_{z \rightarrow 1} Q_2(z),$$

$$p_o(0, 0) = \frac{\alpha(\mu\gamma - \lambda_1\gamma - \lambda_2\beta)}{\alpha N(\mu\gamma - \lambda_1\gamma - \lambda_2\beta) + \lambda_0(\mu + \alpha N)(\beta + \gamma)}. \quad (33)$$

This completes the proof.

And the expected number of customers in the system in states 0, 1, and 2 is

$$\begin{aligned} E[N_0] &= \left. \frac{dQ_0(z)}{dz} \right|_{z=1} = \frac{N(N-1)}{2} p_o(0, 0), \\ E[N_1] &= \left. \frac{dQ_1(z)}{dz} \right|_{z=1} = \frac{\text{Num}_1}{2\alpha(\mu\gamma - \lambda_1\gamma - \lambda_2\beta)^2} \lambda_0 p_o(0, 0), \\ E[N_2] &= \left. \frac{dQ_2(z)}{dz} \right|_{z=1} = \frac{\text{Num}_2}{2\alpha(\mu\gamma - \lambda_1\gamma - \lambda_2\beta)^2} \lambda_0 p_o(0, 0), \end{aligned} \quad (34)$$

where $\text{Num}_1 = \alpha\gamma(\mu\gamma - \lambda_1\gamma - \lambda_2\beta)N^2 + \alpha[2\lambda_2^2\beta + \lambda_2\beta\gamma + \gamma^2(\lambda_1 + \mu)]N + 2\mu(\lambda_2^2\beta + \lambda_2\beta\gamma + \lambda_1\gamma^2)$ and $\text{Num}_2 = \beta\{\alpha(\mu\gamma - \lambda_1\gamma - \lambda_2\beta)N^2 + \alpha[\gamma(\lambda_1 + \mu) + \lambda_2(\beta - 2\lambda_1 + 2\mu)]N + 2\mu[\lambda_1\gamma + \lambda_2(\mu - \lambda_1 + \beta)]\}$.

Denote $T(i, 0), 1 \leq i \leq N - 1$, $T(i, 1), i \geq 1$, and $T(i, 2), i \geq 1$, respectively, as the expected sojourn time of a

tagged customer, given that he is at the i th position of the system and the state of the server is $j = 0, 1, 2$. The following conclusion gives their concrete description.

Proposition 3. *In the observable case, the expected sojourn time of a tagged customer, given that he is at the i th position of the system and the state of server is $j = 0, 1, 2$, is, respectively,*

$$\begin{aligned} T(i, 0) &= \frac{\beta + \gamma}{\mu\gamma} i + \frac{N - i}{\lambda_0}, \quad 1 \leq i \leq N - 1; \\ T(i, 1) &= \frac{\beta + \gamma}{\mu\gamma} i, \quad i \geq 1; \\ T(i, 2) &= \frac{1}{\gamma} + \frac{\beta + \gamma}{\mu\gamma} i, \quad i \geq 1. \end{aligned} \quad (35)$$

Proof. The transition rate diagram illustrated in Figure 2 shows that $T(i, j)$ satisfies the following linear system of equations:

$$T(i, 0) = \frac{N - i}{\lambda_0} + T(i, 1), \quad 1 \leq i \leq N - 1, \quad (36)$$

$$T(1, 1) = \frac{1}{\mu + \beta} + \frac{\beta}{\mu + \beta} T(1, 2), \quad (37)$$

$$T(i, 1) = \frac{1}{\mu + \beta} + \frac{\mu}{\mu + \beta} T(i - 1, 1) + \frac{\beta}{\mu + \beta} T(i, 2), \quad i \geq 2, \quad (38)$$

$$T(i, 2) = \frac{1}{\gamma} + T(i, 1), \quad i \geq 1. \quad (39)$$

Solving $T(i, 2)$ from equation (39) and substituting it into (38) yield

$$T(i, 1) - T(i - 1, 1) = \frac{\beta + \gamma}{\mu\gamma}, \quad i \geq 2. \quad (40)$$

Similarly, considering equations (37) and (39) yields $T(1, 1) = \beta + \gamma/\mu\gamma$, and moreover, combining the above equation yields

$$T(i, 1) = \frac{\beta + \gamma}{\mu\gamma} i, \quad i \geq 1. \quad (41)$$

Substituting $T(i, 1)$ into equations (36) and (39), the results are at hand. This completes the proof. ■

After this, conditioning on the number of customers in the system, the expected sojourn time of the customer, who can observe the server's state, can be computed.

Theorem 1. *In the observable case, given arrival rates $(\lambda_0, \lambda_1, \lambda_2)$, the expected sojourn time of a tagged customer seeing the server's states 0, 1, and 2 upon arrival is*

$$W_0(i) = \begin{cases} T(i+1, 0) = \frac{\beta + \gamma}{\mu\gamma} (i+1) + \frac{N-i-1}{\lambda_0}, & \text{if } 0 \leq i \leq N-2, \\ T(N, 1) = \frac{\beta + \gamma}{\mu\gamma} N, & \text{if } i = N-1. \end{cases} \quad (43)$$

Moreover, denote $p(i|0)$ as the probability that the number of customers in the system is i upon his arrival,

$$\begin{aligned} E[W_0] &= \left(\frac{\beta + \gamma}{\mu\gamma} - \frac{1}{\lambda_0} \right) \left(\frac{E[N_0]}{Q_0} + 1 \right) + \frac{N}{\lambda_0}, \\ E[W_1] &= \frac{\beta + \gamma}{\mu\gamma} \left(\frac{E[N_1]}{Q_1} + 1 \right), \\ E[W_2] &= \frac{\beta + \gamma}{\mu\gamma} \left(\frac{E[N_2]}{Q_2} + 1 \right) + \frac{1}{\gamma}. \end{aligned} \quad (42)$$

Proof. When a tagged customer observes the number of customers in system as i and the server is in state 0 upon arrival, denote $W_0(i)$ as his expected sojourn time. It is readily seen that his expected sojourn time is the expected sojourn time of the $(i+1)$ th customer in the system, i.e.,

conditioning on the server in state 0, i.e., $p(i|0) = p_o(i, 0)/Q_0$. Hence, the expected sojourn time upon seeing state 0 is

$$\begin{aligned} E[W_0] &= \sum_{i=0}^{N-1} W_0(i) p(i|0) = \frac{1}{Q_0} \sum_{i=0}^{N-1} \left[\frac{\beta + \gamma}{\mu\gamma} (i+1) + \frac{N-i-1}{\lambda_0} \right] p_o(i, 0) \\ &= \frac{\sum_{i=0}^{N-1} p_o(i, 0) [((\beta + \gamma)/(\mu\gamma)) - (1/\lambda_0)] i + ((\beta + \gamma)/(\mu\gamma)) + ((N-1)/\lambda_0)}{Q_0} \\ &= \frac{(((\beta + \gamma)/(\mu\gamma)) - (1/\lambda_0)) \sum_{i=0}^{N-1} i p_o(i, 0) + \frac{\beta + \gamma}{\mu\gamma} + \frac{N-1}{\lambda_0}}{Q_0} \\ &= \left(\frac{\beta + \gamma}{\mu\gamma} - \frac{1}{\lambda_0} \right) \frac{(dQ_0(z))/dz|_{z=1}}{Q_0} + \frac{\beta + \gamma}{\mu\gamma} + \frac{N-1}{\lambda_0} = \left(\frac{\beta + \gamma}{\mu\gamma} - \frac{1}{\lambda_0} \right) \frac{E[N_0]}{Q_0} + \frac{\beta + \gamma}{\mu\gamma} + \frac{N-1}{\lambda_0}. \end{aligned} \quad (44)$$

Similarly, denote $p(i|1)$ and $p(i|2)$ as the probabilities that the number of customers in the system is i upon arrival, conditioning on the server in states 1 and 2, and define $W_1(i)$ and $W_2(i)$ as his sojourn

time. Combining the definitions of $T(i+1, 1)$ and $T(i+1, 2)$, conditioning on the customer number i , the expected sojourn times upon seeing states 1 and 2 are

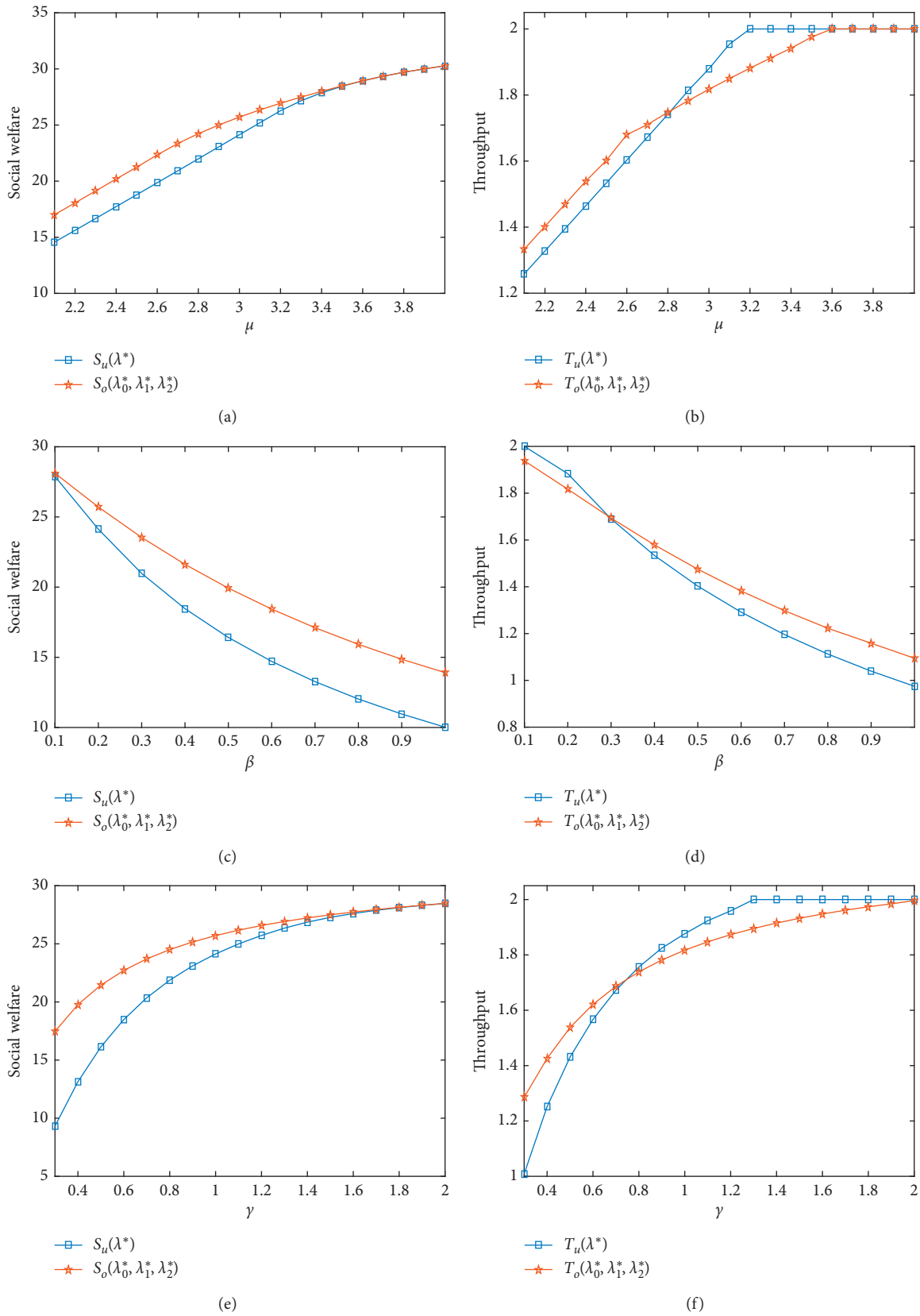


FIGURE 3: Social welfares $S_u(\lambda^*)$ and $S_o(\lambda_0^*, \lambda_1^*, \lambda_2^*)$ and the system throughputs $T_u(\lambda^*)$ and $T_o(\lambda_0^*, \lambda_1^*, \lambda_2^*)$ vs. μ , β , and γ .

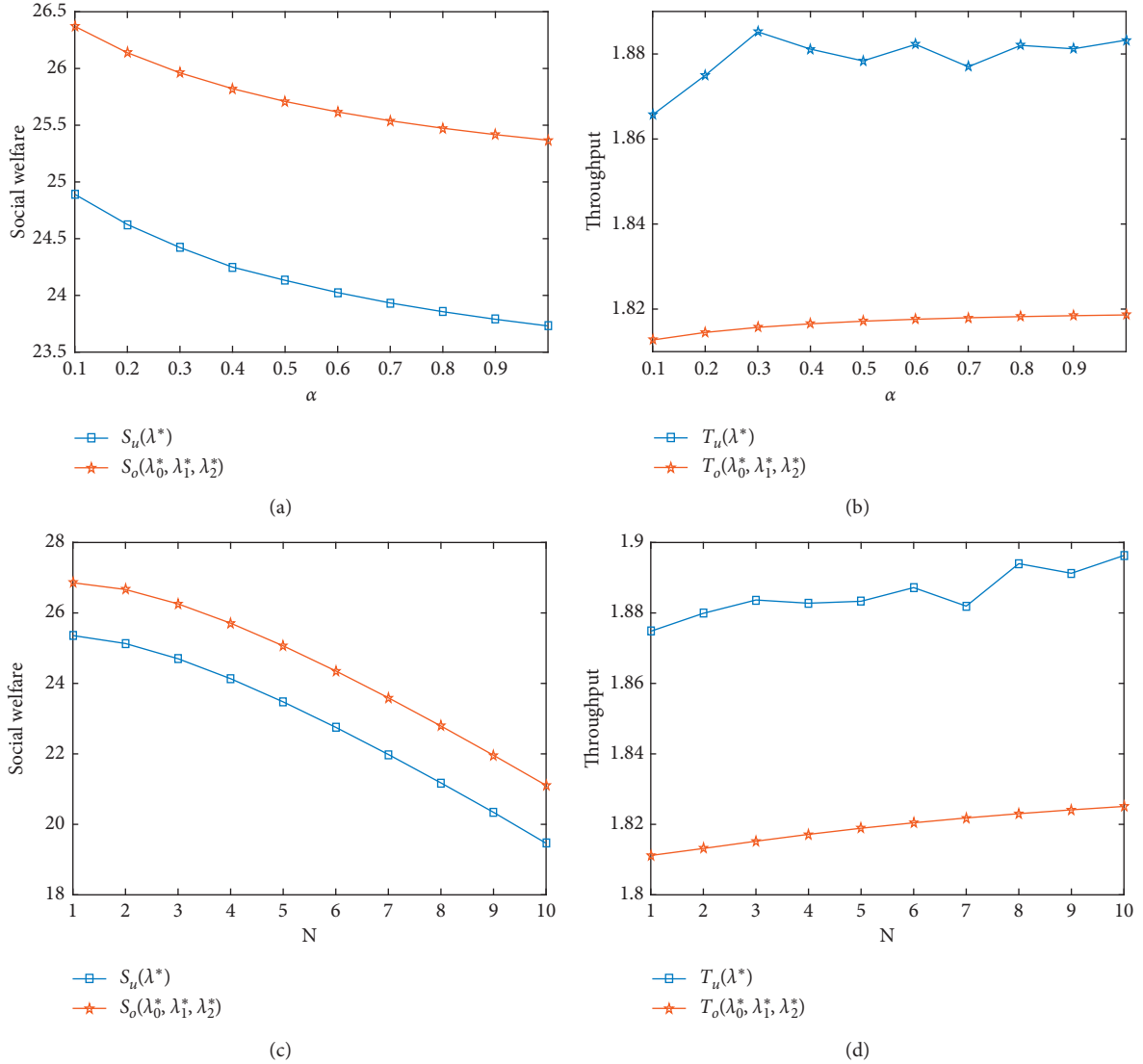


FIGURE 4: Social welfares $S_u(\lambda^*)$ and $S_o(\lambda_0^*, \lambda_1^*, \lambda_2^*)$ and the system throughput $T_u(\lambda^*)$ and $T_o(\lambda_0^*, \lambda_1^*, \lambda_2^*)$ vs. α and N .

$$\begin{aligned}
 E[W_1] &= \sum_{i=0}^{\infty} W_1(i) p(i|1) = \sum_{i=0}^{\infty} T(i+1, 1) \frac{p_o(i, 1)}{Q_1} = \frac{1}{Q_1} \sum_{i=0}^{\infty} \frac{\beta + \gamma}{\mu\gamma} (i+1) p_o(i, 1) \\
 &= \frac{(\beta + \gamma)/\mu\gamma}{Q_1} \sum_{i=0}^{\infty} (i+1) p_o(i, 1) = \frac{\beta + \gamma}{\mu\gamma} \left(\frac{E[N_1]}{Q_1} + 1 \right), \\
 E[W_2] &= \sum_{i=0}^{\infty} W_2(i) p(i|2) = \sum_{i=0}^{\infty} T(i+1, 2) \frac{p_o(i, 2)}{Q_2} = \frac{1}{Q_2} \sum_{i=0}^{\infty} \left[\frac{\beta + \gamma}{\mu\gamma} (i+1) + \frac{1}{\gamma} \right] p_o(i, 2) \\
 &= \frac{1}{Q_2} \left[\frac{\beta + \gamma}{\mu\gamma} \sum_{i=0}^{\infty} i p_o(i, 2) + \left(\frac{\beta + \gamma}{\mu\gamma} + \frac{1}{\gamma} \right) \sum_{i=0}^{\infty} p_o(i, 2) \right] = \frac{\beta + \gamma}{\mu\gamma} \frac{E[N_2]}{Q_2} + \frac{\beta + \mu + \gamma}{\mu\gamma}.
 \end{aligned} \tag{45}$$

This completes the proof. ■

After this, we can define the social welfare function $S_o(\lambda_0, \lambda_1, \lambda_2) = \sum_{i=0}^2 \lambda_i Q_i (R - cE[W_i])$ and the system throughput $T_o(\lambda_0, \lambda_1, \lambda_2) = \mu(Q_1 - p_o(0, 1))$.

4. Numerical Experiments

In this section, particle swarm optimization (PSO) algorithm presented by Carlisle and Dozier [13] is used to seek the optimal arrival rate λ^* or $(\lambda_0^*, \lambda_1^*, \lambda_2^*)$ of the social welfare

functions $S_u(\lambda)$ or $S_o(\lambda_0, \lambda_1, \lambda_2)$. And then, the tendency of the corresponding social welfares $S_u(\lambda^*)$ and $S_o(\lambda_0^*, \lambda_1^*, \lambda_2^*)$ and the system throughput $T_u(\lambda^*)$ and $T_o(\lambda_0^*, \lambda_1^*, \lambda_2^*)$ with respect to parameters μ , α , β , γ , and N is shown.

PSO algorithm has been shown that to occupy preferred properties in convergence and so in this paper, the algorithm presented in [13] is used to search the optimal arrival rate λ^* or $(\lambda_0^*, \lambda_1^*, \lambda_2^*)$. And the same parameters given in [13] are used except that the number of particles is $S = 100$ and the number of iterations is $M = 2000$. And in each situation, the related parameter setting is $\Lambda = 2$, $\mu = 3$, $\alpha = 0.5$, $\beta = 0.2$, $\gamma = 1$, $R = 18$, $c = 2$, and $N = 4$, if necessary. Any example in this work is repeated 10 times, and the almost identical result is achieved each time. This fact shows PSO algorithm is robust and appropriate to this model.

From Figure 3, the social welfare and the throughput have the same tendency with respect to parameters μ , β , and γ . As the service rate μ increases, the sojourn time of the customer in the system could be shortened and his benefit is also improved. Therefore, more customers prefer to join the system and the throughput is increasing, which also incurs the increasing of social welfare.

Second, in observable cases, the social welfare is greater than that in unobservable case. Hence, disclosing the server state to the customer is advantageous to the system manager. However, this adjustment is not fit for the throughput. For greater μ , incomplete information is liable to reach more throughput. Once μ is large enough, there is no obvious difference between the observable and unobservable cases. The same tendency can be seen in Figure 3(f).

Third, for the smaller failure rate β , customers without any information have more motivations to join the system. And as β increases, they cannot determine the state of server and so their motivations decrease. This tendency can be seen in Figure 3(d). And it is readily understood that the social welfare and throughput are all decreasing in Figure 3(c).

From the point of system manager, enhancing the system throughput can relieve the system congestion and strengthen the service capacity. However, in some situations, this adjustment is adverse to social welfare. With the increase in vacation rate α , the system has the greater probability in state 0 and hence more customers cannot be served timely which leads to decrease in social welfare. On the other hand, if the reward R is large enough, all the customers would like to join the system and the throughput is equal to the maximum of arrival rate Λ . If not, customers also have the motivation to join the system only if their benefit is positive. Therefore, the throughput is increasing no matter what α is. But the social welfare is decreasing even if the throughput is increasing, because customer has a higher probability in state 0 and cannot obtain the reward without being served. The same tendency for parameter N can be seen in Figures 4(c) and 4(d).

Besides, information disclosing level also influences the social welfare and throughput. For example, information level has the same effect on these two performance measures in the case of smaller μ , greater β , or smaller γ . Hence, which is more crucial to the system manager is the basis of making decision.

5. Conclusions

This work considers a queueing model with N -policy vacation and unreliable server. According to whether the server state is observable or not, the main two performance measures, social welfare and throughput, are derived and maximized by particle swarm optimization algorithm. Numerical results illustrate that these two measures may not have the same tendency on the information level and system parameters. That is to say, increased social welfare may lead to decreased throughput. From the opinion of the manager, how to achieve more benefit and greater throughput, how to adjust the system parameter, and whether or not reveal system information to customers are the key factors that the manager should consider.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 11601251) and the Development Plan of Young Innovation Team in Colleges and Universities of Shandong Province (Grant no. 2019KJN011).

References

- [1] P. Naor, "The regulation of queue size by levying tolls," *Econometrica*, vol. 37, no. 1, pp. 15–24, 1969.
- [2] N. M. Edelson and D. K. Hilderbrand, "Congestion tolls for poisson queueing processes," *Econometrica*, vol. 43, no. 1, pp. 81–92, 1975.
- [3] P. Guo and P. Zipkin, "Analysis and comparison of queues with different levels of delay information," *Management Science*, vol. 53, no. 6, pp. 962–970, 2007.
- [4] A. Economou and S. Kanta, "Equilibrium balking strategies in the observable single-server queue with breakdowns and repairs," *Operations Research Letters*, vol. 36, no. 6, pp. 696–699, 2008.
- [5] P. Guo and R. Hassin, "Strategic behavior and social optimization in Markovian vacation queues," *Operations Research*, vol. 59, no. 4, pp. 986–997, 2011.
- [6] J. T. Wang, S. Zhu, and W. Li, "Strategic behavior of cognitive radio networks with different information," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4810–4823, 2019.
- [7] X. Zhang and J. Wang, "Optimal inventory threshold for a dynamic service make-to-stock system with strategic customers," *Applied Stochastic Models in Business and Industry*, vol. 35, no. 4, pp. 1103–1123, 2019.
- [8] W. Sun, S. Y. Li, and N. S. Tian, "Comparisons of customer behavior in Markovian queues with vacation policies and geometric abandonments," *RAIRO-Operations Research*, vol. 54, no. 3, pp. 616–636, 2020.

- [9] W. Sun, S. Li, W. Liu, and N. Tian, "Comparisons of site-clearing and non-site-clearing Markovian queues with server failures and unreliable repairer," *Quality Technology & Quantitative Management*, vol. 16, no. 2, pp. 211–230, 2019.
- [10] W. Sun, S. Li, Y. Wang, and N. Tian, "Comparisons of exhaustive and nonexhaustive M/M/1/N queues with working vacation and threshold policy," *Journal of Systems Science and Systems Engineering*, vol. 28, no. 2, pp. 154–167, 2019.
- [11] Z. B. Wang and J. T. Wang, "Information heterogeneity in a retrieval queue: throughput and social welfare maximization," *Queueing Systems*, vol. 92, no. 1-2, pp. 131–172, 2019.
- [12] S. Cui, Z. Wang, and L. Yang, "The economics of line-sitting," *Management Science*, vol. 66, no. 1, pp. 227–242, 2020.
- [13] A. Carlisle and G. Dozier, *An Off-The-Shelf PSO. Proceeding of the Workshop on Particle Swarm Optimization*, Purdue School of Engineering and Technology, Indianapolis, Ind., USA, 2001.