RAIRO-Oper. Res. 52 (2018) 439–452
https://doi.org/10.1051/ro/2017041

RAIRO Operations Research
www.rairo-ro.org

SOCIAL OPTIMIZATION IN $M/M/1$ QUEUE WITH WORKING VACATION AND N-POLICY

Qing-qing Ma[1,3,*], Ji-hong Li[2] and Wei-qi Liu[2,4]

Abstract. This paper deals with the N-policy $M/M/1$ queueing system with working vacations. Once the system becomes empty, the server begins a working vacation and works at a lower service rate. The server resumes regular service when there are N or more customers in the system. By solving the balance equations, the stationary probability distribution and the mean queue length under observable and unobservable cases are obtained. Based on the reward-cost structure and the theory of Markov process, the social welfare function is constructed. Finally, the impact of several parameters and information levels on the mean queue length and social welfare is illustrated *via* numerical examples, comparison work shows that queues with working vacations(WV) and N-policy have advantage in controlling the queue length and improving the social welfare.

Mathematics Subject Classification. 90B22, 60K25, 68M20.

Received December 29, 2015. Accepted May 17, 2017.

1. Introduction

During the last decades, many researchers have devoted their efforts to the study of queueing systems with working vacations and(or) N-policy which are concerned with customer strategic behavior and social optimal welfare. A feasibly way for the study on customer strategic behavior is to quantize the waiting cost and the reward after service. The corresponding research can be found in Naor [20] and Edelson and Hildebrand [5] where a simple $M/M/1$ queue with linear reward-cost structure was considered, and the equilibrium strategy and optimal social strategy were obtained. From then on, most researchers study customer behavior with reward-cost structure, see [6–8, 17, 23, 30]. Decades later, Hassin and Haviv [9] summarized the main research results in this area and opened more questions for further research.

Customer strategic behaviour and social optimal benefit in queuing systems with working vacations have been studied extensively. Systems with working vacations have wide application in management of service systems. In these systems, the server keeps on working with a lower service rate rather than completely stopping working during a vacation period. Sevi and Finn [22] studied this working mechanism for the first time, they obtained

Keywords. Markov process, working vacations, N-policy, social optimization.

[1] School of Economics and Management, Shanxi University, Taiyuan 030006, China.
[*] Corresponding author: maqq0814@sina.com
[2] Institute of Management and Decision, Shanxi University, Taiyuan 030006, China.
[3] School of Mathematics and Statistics, Carleton University, Ottawa K1V5B6, Canada.
[4] Faculty of Finance and Banking, Shanxi University of Finance and Economics, Taiyuan 030006, China.

the stationary distribution of system and the mean sojourn time of customers. Liu, Xu and Tian [16] gave some results on the $M/M/1$ queue with working vacations. Doshi [3], Takagi [24] and Tian and Zhang [25] provided an amount of excellent models on vacation queueing systems. Liu, Ma and Li [17] studied customer equilibrium behavior in fully observable and almost observable $M/M/1$ and $Geo/Geo/1$ queueing systems with single vacation. Based on the reward-cost structure, the equilibrium threshold strategies were obtained and the stationary system behavior was analyzed. Zhang, Wang and Liu [30] investigated the customer strategic behavior in a single server queueing system with working vacations under four cases with respect to different information level. Along with the reward-cost structure, equilibrium threshold strategies and the corresponding Nash equilibria for each case were derived. Kasim and Gupur [10] proved that the $M/G/1$ queuing model with a single working vacation had a unique non-negative time-dependent solution and the time-dependent solution strongly converged to its steady-state solution. For more researches on working vacation, one can refer to [12–15, 19].

N-policy is another vacation policy which is commonly used in flexible manufacturing systems, made-to-order (MTO) systems and so on. With N-policy, the server begins a vacation when the system becomes empty and resumes regular work whenever there are $N(N > 1)$ or more customers in the systems. Due to the vacation time of the server always being dependent on the number of customers in the system, N-policy systems have an advantage in reducing the sojourn time of customers and increasing customer satisfaction. Queueing systems with N-policy was pioneered by Yadin and Naor [28], then Kella [11] applied it to working vacation queues, he or she laid the foundation of working vacation queue with N-policy. Guo and Hassin [6, 7] studied the customer equilibrium and optimal arrival rate in vacation queue under fully unobservable and almost observable cases, and obtained the optimal threshold N to maximize the social welfare. Guo and Li [8] extended Guo and Hassin [6] to the almost unobservable case, they also studied the customer equilibrium and optimal arrival rate, completing the model constructed by Guo and Hassin [6]. The study of [6–8] assumed that the server completely stopped working during vacation time, resumed to work as soon as there are N customers waiting in the queue. Sun, Li and Cheng [23] studied the customer equilibrium and socially optimal balking strategies in single-server Markovian queues with multiple vacations and N-policy under fully observable and fully unobservable cases. Different from [6–8], Sun, Li and Cheng [23] assumed that the server was reactivated only in case that there were at least N customers waiting in the queue after completing the vacation, otherwise, the server continued taking another vacation. They obtained both customers equilibrium and optimal balking strategies. Other research on $M/M/1$ queue with N-policy can be seen in [26, 29]. Wu, Tang and Yu [27] considered an M/G/1 repairable queueing system with N-policy and single vacations, in which the system had an unreliable service station and replaceable repair facility. They numerically determined the optimal threshold N for minimizing the cost function. Liu, Ma and Li [18] analyzed the customer social equilibrium strategy in an $M/M/1$ queue with working vacations and N-policy under the observable case where the arriving customers can observe the queue length upon arrival. After that, to the authors' knowledge, few researchers have discussed the case where the server served the customers with a lower service rate in working vacation queue with N-policy.

Previous studies have made great contributions to the Queueing Theory. However, to the best of our knowledge, there is no previous research dealing with the social optimization of queueing systems with working vacations and N-policy. Considering the difficulties in solving the $M/G/1$ working vacation queueing models, we consider an $M/M/1$ model first. The $M/G/1$ queue with working vacations and N-policy will be discussed in a separate paper. Our model is characterized as follows: In an $M/M/1$ queue, when the system becomes empty, the server begins a working vacation; once there are $N(N > 1)$ customers in the system, the server return to normal service period after completing current service. It's obvious that when the vacation service rate is 0, it becomes the model discussed in Guo and Hassin [23, 26, 29]; When $N = 1$, it becomes a classical queue. Queueing systems with such mechanisms could be extensively applied in management. An example using a machine assembling system can be illustrated as follows: when the facility finishes all the jobs in line and the line becomes empty, a working vacation begins. During the working vacation the facility keeps working but only at a fraction of its full capacity. Thus, when a new job arrives the facility handles the job with a lower service rate. Once there are N jobs in the system the facility comes back to work immediately and serves jobs with the normal service rate. For more application examples, one can refer to [5, 24, 25].

In many studies, the explicit expression of the results always cannot be reached, so graphs is always used to analyze the results, see [17, 19, 30]. Besides graphs, PSO(Particle Swarm Optimization) algorithm proposed by [1, 2, 4, 21] is a way to solve optimization problems with no explicit expression. Yang and Wu [29] succeed in solving the optimization problem with PSO algorithm. In this paper, it is arduous task to develop the corresponding explicit expression of λ, N and R when the social welfare is optimized, due to the fact that the expressions of stationary probability and mean queue length are rather complex. If the explicit expressions of the optimized λ, N and R is needed, one can adopt PSO algorithm as well.

The rest of this paper is organized as follows. In Section 1, we describe the queueing model and the reward-cost structure; In Section 2, the mean queue length and social welfare for observable case is analyzed; In Section 3, the mean queue length and social welfare for unobservable case is analyzed; Finally, in Section 4, some conclusions are given.

2. Model description

Consider a single server $M/M/1$ queue with infinite waiting room, in which customers arrive according to a Poisson process of rate λ, the service order is FCFS(first-come first-served). The service times are independent and exponentially distributed with parameter $\mu_1(\mu_0)$. Different from previous N-policy, the server begins a working vacation once the system becomes empty. Customers who arrive during the vacation period are served with a lower service rate μ_0 which is also i.i.d and $\lambda < \mu_1$ holds.The server returns to normal service period after completing current service when there are N or more customers in the system. Unless otherwise stated, $N > 1$.

Let (n, s) be the system state, where n is the number of the customers and s is the server state. Define

$$s = \begin{cases} 0, \text{ If the service rate is } \mu_0 \text{ upon arrival} \\ 1, \text{ If the service rate is } \mu_1 \text{ upon arrival} \end{cases} \tag{2.1}$$

The process(n, s) is a two-dimensional embedded Markov chain with the state space:

$$\Omega = \left\{(0,0)\right\} \bigcup \left\{(k,j), k \geq 1, j = 0, 1\right\}.$$

To model customer behavior, assume that every customer receives a reward of R units for completing service, and pays a sojourn cost of C units per time unit when in the system (*i.e.*, the waiting time plus the service time). Let U be the expected customer net benefit, we have

$$U = R - CE[W]$$

Where $E[W]$ is the expected sojourn time. Customers choose to join the system if U is non-negative; balk otherwise. Customers are risk neutral and their decisions are irrevocable, meaning that reneging of entering customers are not allowed.

In the next sections the mean queue length and social welfare of customers will be studied based on the information they have upon arrival.

3. Customer strategies for observable case

In this section, we shall deal with the observable case where the arriving customers are informed of the queue length upon arrival. We use (n, s) to describe the system sate, the state space is:

$$\{(0,0), (1,0), \ldots, (N-1,0), (1,1), (2,1), \ldots, (N,1), (N+1,1), \ldots, (n_e,1)\}$$

Where n_e is the maximum number of customers in the system at state 1. Figure 1 illustrates the dynamics of queueing system in the observable case.

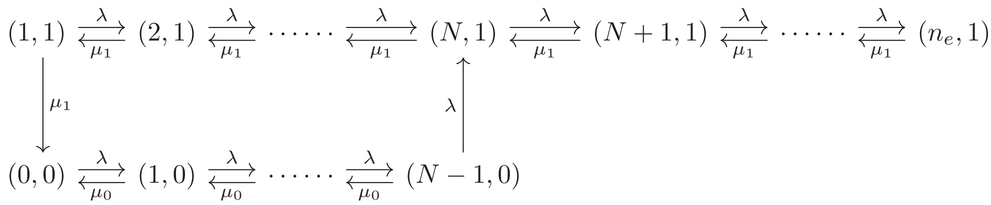

FIGURE 1. Transition rate diagram for the observable model.

Lemma 3.1. *The equilibrium threshold strategy of customer n_e is given by*

$$n_e = \left\lfloor \frac{R\mu_1}{C} \right\rfloor$$

To avoid the degenerate case(the state space consist of only one state (0,0)), we assume that $n_e \geq N$ always holds.

3.1. The stationary distribution and the mean queue length

The system stationary distribution and the derivation of the expected number of customers are detailed below.

Lemma 3.2. *In the observable M/M/1 queue with working vacations and N-policy where the arriving customers know the number of customers in the system, the stationary distribution is given as follows:*

$$P_{ob}(k,0) = \frac{\rho_0^k - \rho_0^N}{\rho_0^N(1-\rho_0)} \frac{\mu_1}{\mu_0} P_{ob}(1,1), k = 0,1,2,\ldots,N-1$$

$$P_{ob}(k,1) = \frac{1-\rho_1^k}{1-\rho_1} P_{ob}(1,1), k = 1,2,\ldots,N$$

$$P_{ob}(N+k,1) = \frac{\rho_1^k(1-\rho_1^N)}{1-\rho_1} P_{ob}(1,1), k = 0,1,2,\ldots,n_e-N$$

Where

$$P_{ob}(1,1) = \left[\frac{1-(N+1)\rho_0^N + N\rho_0^{N+1}}{\rho_0^N(1-\rho_0)^2} \frac{\mu_1}{\mu_0} + \frac{N - N\rho_1 - \rho_1^{n_e-N+1} + \rho_1^{n_e+1}}{(1-\rho_1)^2}\right]^{-1}$$

Proof. The system stationary distribution is obtained through solving the following balance equations:

$$\mu_1 P_{ob}(1,1) + \mu_0 P_{ob}(1,0) = \lambda P_{ob}(0,0) \tag{3.1}$$

$$(\lambda + \mu_1) P_{ob}(1,1) = \mu_1 P_{ob}(2,1) \tag{3.2}$$

$$(\mu_0 + \lambda) P_{ob}(k,0) = \lambda P_{ob}(k-1,0) + \mu_0 P_{ob}(k+1,0), k = 1,2,\ldots,N-2 \tag{3.3}$$

$$\lambda P_{ob}(N-2,0) = (\mu_0 + \lambda) P_{ob}(N-1,0) \tag{3.4}$$

$$\lambda P_{ob}(N-1,0) + \lambda P_{ob}(N-1,1) + \mu_1 P_{ob}(N+1,1) = (\lambda + \mu_1) P_{ob}(N,1) \tag{3.5}$$

$$\lambda P_{ob}(k-1,1)+\mu_1 P_{ob}(k+1,1)=(\lambda+\mu_1)P_{ob}(k,1), k=2,3,\ldots,N-1 \tag{3.6}$$

$$(\lambda+\mu_1)P_{ob}(N+k,1)=\lambda P_{ob}(N+k-1,1)+\mu_1 P_{ob}(N+k+1,1), k=1,2,\ldots,n_e-N \tag{3.7}$$

Donate $\rho_1=\frac{\lambda}{\mu_1}$,$\rho_0=\frac{\lambda}{\mu_0}$.

From (3.3), the probabilities $P_{ob}(k,0), k=1,2,\ldots,N-1$ are the solutions of the following homogeneous linear difference equation:

$$\mu_0 x_{n+1}-(\mu_0+\lambda)x_n+\lambda x_{n-1}=0, n=1,2,\ldots,N-2 \tag{3.8}$$

The corresponding characteristic equation of (3.8) is

$$\mu_0 x^2-(\mu_0+\lambda)x+\lambda=0$$

which has two roots: 1 and $\rho_0(\neq 1)$. So the general solution of (3.8), donated by x_n^{hom}, is $x_n^{hom}=A_1\rho_0^n+A_2$, where A_1 and A_2 are the coefficients to be determined. From (3.3), we know

$$P_{ob}(N-1,0)=A_1\rho_0^{N-1}+A_2 \tag{3.9}$$

and

$$P_{ob}(0,0)=A_1+A_2 \tag{3.10}$$

From (3.1), (3.4), (3.9) and (3.10), the following equations are obtained:

$$\begin{cases} \mu_1 P_{ob}(1,1)+\mu_0(A_1\rho_0+A_2)=\lambda(A_1+A_2) \\ \rho_0(A_1\rho_0^{N-2}+A_2)=(1+\rho_0)(A_1\rho_0^{N-1}+A_2), \end{cases} \tag{3.11}$$

which yields

$$\begin{cases} A_1=\dfrac{1}{\rho_0^N(1-\rho_0)}\frac{\mu_1}{\mu_0}P_{ob}(1,1), \\ A_2=-A_1\rho_0^N, \end{cases} \tag{3.12}$$

Thus

$$P_{ob}(k,0)=A_1\rho_0^k+A_2, k=0,1,\ldots,N-1 \tag{3.13}$$

where A_1 and A_2 are given in (3.12).

Then we consider the probabilities $P_{ob}(k,1), k=1,2,\ldots,N-1$ and $P_{ob}(N+k,1), k=1,2,\ldots,n_e-N$.

It follows (3.6) that

$$P_{ob}(k+1,1)-P_{ob}(k,1)=\rho_1^{k-1}(P_{ob}(2,1)-P_{ob}(1,1)), k=2,3,\ldots,N-1 \tag{3.14}$$

Taking (3.2) into consideration, we have:

$$P_{ob}(k+1,1)-P_{ob}(k,1)=\rho_1^k P_{ob}(1,1), k=2,3,\ldots,N-1 \tag{3.15}$$

Thus we have:

$$P_{ob}(k,1)=\frac{1-\rho_1^k}{1-\rho_1}P_{ob}(1,1), k=3,\ldots,N \tag{3.16}$$

Because $k=1,2$ also holds for (3.16), then:

$$P_{ob}(k,1)=\frac{1-\rho_1^k}{1-\rho_1}P_{ob}(1,1), k=1,2,\ldots,N \tag{3.17}$$

Solve the remaining equations in the same way, we have:

$$P_{ob}(N+k,1)=\frac{\rho_1^k(1-\rho_1^N)}{1-\rho_1}P_{ob}(1,1), k=1,2,\ldots,n_e-N \tag{3.18}$$

Besides, $k=0$ holds for (3.18).

Now, all the stationary state probabilities in terms of $P_{ob}(1,1)$ are obtained. The remaining probability $P_{ob}(1,1)$ can be found by the normalization condition

$$\sum_{k=0}^{N-1}P_{ob}(k,0)+\sum_{k=1}^{N-1}P_{ob}(k,1)+\sum_{k=0}^{n_e-N}P_{ob}(N+k,1)=1$$

$P_{ob}(1,1)$ is shown below:

$$P_{ob}(1,1)=\left[\frac{1-(N+1)\rho_0^N+N\rho_0^{N+1}}{\rho_0^N(1-\rho_0)^2}\frac{\mu_1}{\mu_0}+\frac{N-N\rho_1-\rho_1^{n_e-N+1}+\rho_1^{n_e+1}}{(1-\rho_1)^2}\right]^{-1} \tag{3.19}$$

Lemma 3.3. *In the observable M/M/1 queue with working vacations and N-policy where the arriving customers know the number of customers in the system, the mean sojourn time of customer* $E[W_{ob}]$ *and the mean queue length* L_{ob} *is given below*:

According to Little's Law, we have:

$$E[W_{ob}]=\frac{L_{ob}}{\lambda(1-P_{ob}(n_e,1))} \tag{3.20}$$

Where $P_{ob}(n_e,1)$ is the probability that the queue is at its maximum size, and $\lambda(1-P_{ob}(n_e,1))$ is the efficient arrival rate of customer, $P_{ob}(n_e,1)=\dfrac{\rho_1^{n_e-N}(1-\rho_1^N)}{1-\rho_1}P_{ob}(1,1)$.

$$\begin{aligned}
L_{ob}&=\sum_{k=1}^{N-1}k(P_{ob}(k,0)+P_{ob}(k,1))+\sum_{k=0}^{n_e-N}(N+k)P_{ob}(N+k,1)\\
&=\sum_{k=1}^{N-1}k\left(\frac{\rho_0^k-\rho_0^N}{\rho_0^N(1-\rho_0)}\frac{\mu_1}{\mu_0}P_{ob}(1,1)+\frac{1-\rho_1^k}{1-\rho_1}P_{ob}(1,1)\right)+\sum_{k=0}^{n_e-N}(N+k)\frac{\rho_1^k(1-\rho_1^N)}{1-\rho_1}P_{ob}(1,1)\\
&=\frac{1}{\rho_0^N(1-\rho_0)}\frac{\mu_1}{\mu_0}P_{ob}(1,1)\left(\rho_0\frac{1-N\rho_0^{N-1}+(N-1)\rho_0^N}{(1-\rho_0)^2}-\frac{N(N-1)}{2}\rho_0^N\right)\\
&\quad+\frac{1}{1-\rho_1}P_{ob}(1,1)\left(\frac{N(N-1)}{2}-\rho_1\frac{1-N\rho_1^{N-1}+(N-1)\rho_1^N}{(1-\rho_1)^2}\right)\\
&\quad+\frac{1-\rho_1^N}{1-\rho_1}P_{ob}(1,1)\left(N\frac{1-\rho_1^{n_e-N+1}}{1-\rho_1}+\rho_1\frac{1-(n_e-N+1)\rho_1^{n_e-N}+(n_e-N)\rho_1^{n_e-N+1}}{(1-\rho_1)^2}\right)
\end{aligned} \tag{3.21}$$

Lemma 3.4. *In the observable M/M/1 queue with working vacations and N-policy where the arriving customers know the number of customers in the system, the social welfare per time unit* SW_{ob} *is given below*:

$$SW_{ob}=(R-CE[W_{ob}])\lambda(1-P_{ob}(n_e,1)))=R\lambda(1-P_{ob}(n_e,1))-CL_{ob}$$

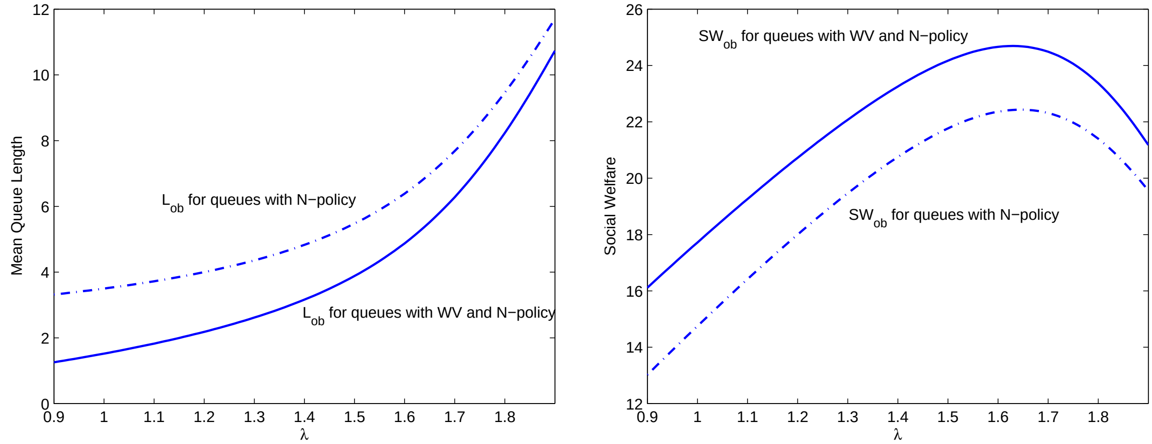

FIGURE 2. Mean queue length L_{ob} and social welfare SW_{ob} for observable queues with N-policy and queues with working vacations and N-policy ($\mu_0 = 1.5$, $\mu_1 = 2$, $C = 1.5$, $N = 6$, $R = 20$).

3.2. Numerical simulation for mean queue length and social welfare

In this section, the impact of λ, N and R on the mean queue length L_{ob} and social welfare SW_{ob} will be graphically presented. What's more, comparison work is carried out to show which mechanism, queues with N-policy or queues with working vacations and N-policy, is better in controlling the mean queue length and improving the social welfare.

Figure 2 shows that the mean queue length L_{ob} increases with λ, the social welfare SW_{ob} first increases and then decreases with λ in queues with working vacations and N-policy. With λ increasing, more customers join the system per time unit, the queue length become larger, the mean queue length increases. Once there are N customers in the system, the server resumes to normal work period, there will be more customers completing service per time unit, the social welfare increases. However, as λ keeps increasing, there will be congestion in the system, customers will suffer longer queue length and more waiting cost, which has a negative effect on the social welfare, social welfare decreases.

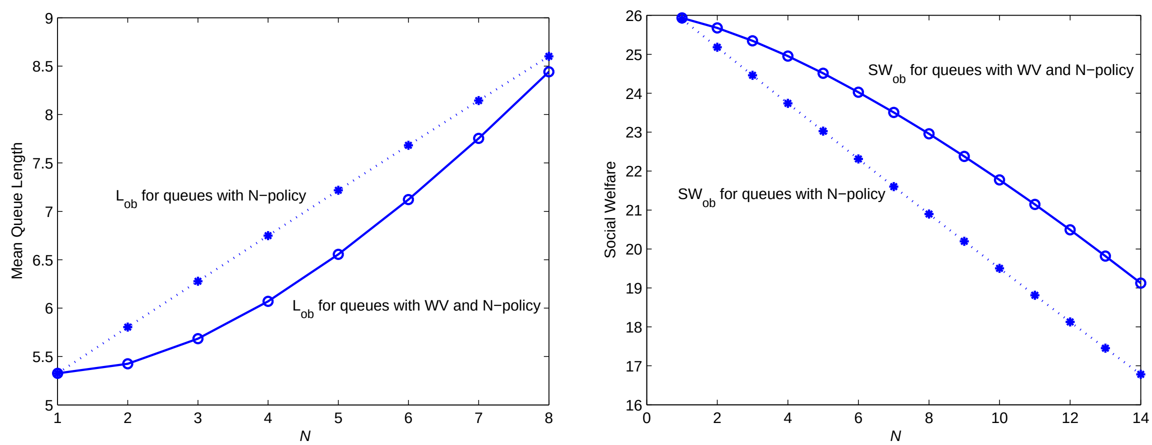

FIGURE 3. Mean queue length L_{ob} and social welfare SW_{ob} for observable queues with N-policy and queues with working vacations and N-policy ($\mu_0 = 1.5$, $\mu_1 = 2$, $\lambda = 1.7$, $C = 1.5$, $R = 20$).

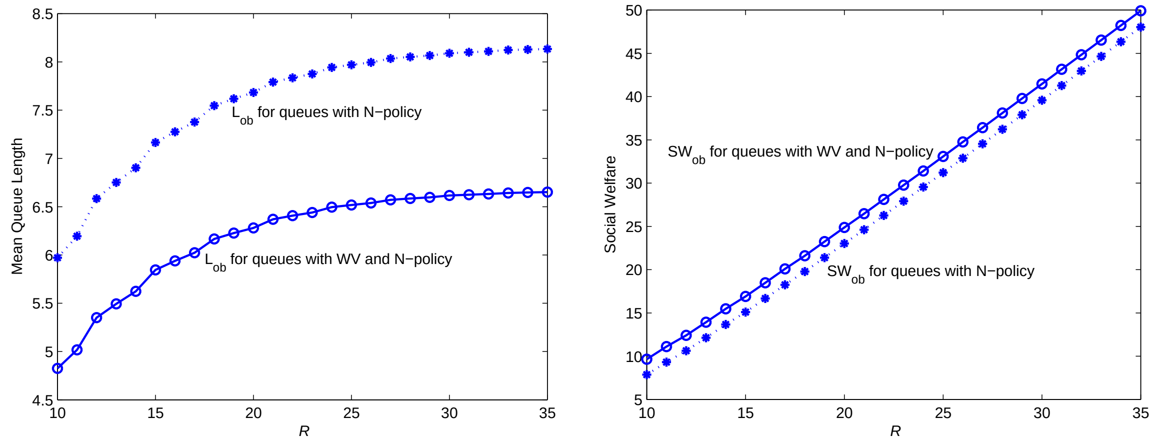

FIGURE 4. Mean queue length L_{ob} and social welfare SW_{ob} for observable queues with N-policy and queues with working vacations and N-policy ($\mu_0 = 1.5$, $\mu_1 = 2$, $\lambda = 1.7$, $C = 1.5$, $N = 6$.)

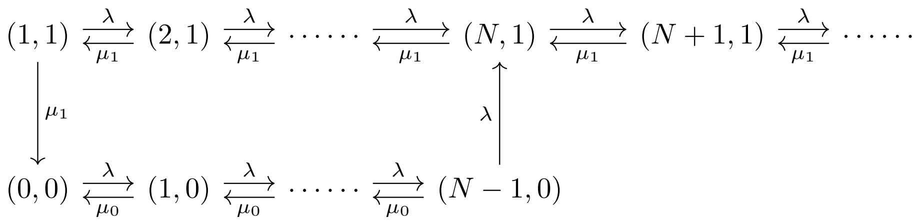

FIGURE 5. Transition rate diagram for the unobservable model.

Figure 3 shows that the mean queue length L_{ob} increases with N, the social welfare SW_{ob} decreases with N in queues with working vacations and N-policy. Because with a larger threshold N, it's difficult for the server returning to normal work period, there will be more customers served with the vacation service rate μ_0, customers who join the queue suffer longer waiting delay and the queue length grows larger, then the mean queue length becomes greater. Longer waiting delay causes more waiting cost of customers, thus the social welfare decreases with N.

Figure 4 shows that both the mean queue length L_{ob} and the social welfare SW_{ob} increase with R. With a larger R, customers receive higher reward after service. Social welfare SW_{ob}, as the sum of customers benefit, increases with R.

On the other hand, we can see that in Figures 2−4, the mean queue length for queues with N-policy is always larger than that for queues with working vacations and N-policy, the social welfare for queues with working vacations and N-policy is always greater than that for queues with N-policy. Because in queues with N-policy, the server doesn't serve customers in vacation period, customers suffer a longer waiting time and a greater waiting cost; while in queues with working vacations and N-policy, customers can be served with a lower service rate during vacation period.

4. Customer strategies for unobservable case

We now turn our interest to the unobservable case where the arriving customers have no information of the queue length upon arrival. The transition rate diagram is shown in Figure 5.

4.1. Analysis of the expected sojourn time and the mean queue length

Let $p_{un}(n,i),(i=0,1,n=i,i+1,i+2,\ldots)$ be the stationary distribution of the unobservable model. The following Lemmas hold.

Lemma 4.1. *In the unobservable M/M/1 queue with working vacations and N-policy where the arriving customers have no information of the queue length upon arrival, the system stationary distribution is given as follows*:

$$P_{un}(k,0)=\frac{\rho_0^k-\rho_0^N}{\rho_0^N(1-\rho_0)}\frac{\mu_1}{\mu_0}P_{un}(1,1), k=0,1,2,\ldots,N-1$$

$$P_{un}(k,1)=\frac{1-\rho_1^k}{1-\rho_1}P_{un}(1,1), k=1,2,\ldots,N$$

$$P_{un}(N+k,1)=\frac{\rho_1^k(1-\rho_1^N)}{1-\rho_1}P_{un}(1,1), k=0,1,2,\ldots$$

Where

$$P_{un}(1,1)=\left[\frac{1-(N+1)\rho_0^N+N\rho_0^{N+1}}{\rho_0^N(1-\rho_0)^2}\frac{\mu_1}{\mu_0}+\frac{N}{1-\rho_1}\right]^{-1}$$

Proof. The system stationary distribution is obtained through solving the following balance equations:

$$\mu_1 P_{un}(1,1)+\mu_0 P_{un}(1,0)=\lambda P_{un}(0,0) \tag{4.1}$$

$$(\lambda+\mu_1)P_{un}(1,1)=\mu_1 P_{un}(2,1) \tag{4.2}$$

$$(\mu_0+\lambda)P_{un}(k,0)=\lambda P_{un}(k-1,0)+\mu_0 P_{un}(k+1,0), k=1,2,\ldots,N-2 \tag{4.3}$$

$$\lambda P_{un}(N-2,0)=(\mu_0+\lambda)P_{un}(N-1,0) \tag{4.4}$$

$$\lambda P_{un}(N-1,0)+\lambda P_{un}(N-1,1)+\mu_1 P_{un}(N+1,1)=(\lambda+\mu_1)P_{un}(N,1) \tag{4.5}$$

$$\lambda P_{un}(k-1,1)+\mu_1 P_{un}(k+1,1)=(\lambda+\mu_1)P_{un}(k,1), k=2,3,\ldots,N-1 \tag{4.6}$$

$$(\lambda+\mu_1)P_{un}(N+k,1)=\lambda P_{un}(N+k-1,1)+\mu_1 P_{un}(N+k+1,1), k=1,2,\ldots \tag{4.7}$$

Define $\rho_1=\frac{\lambda}{\mu_1}$,$\rho_0=\frac{\lambda}{\mu_0}$.

Solve the equations in the same way as in the observable case, Lemma 4.1 is reached. Using the normalization condition

$$\sum_{k=0}^{N-1}P_{un}(k,0)+\sum_{k=1}^{N-1}P_{un}(k,1)+\sum_{k=0}^{+\infty}P_{un}(N+k,1)=1$$

$P_{un}(1,1)$ is given below:

$$P_{un}(1,1)=\left[\frac{1-(N+1)\rho_0^N+N\rho_0^{N+1}}{\rho_0^N(1-\rho_0)^2}\frac{\mu_1}{\mu_0}+\frac{N}{1-\rho_1}\right]^{-1} \tag{4.8}$$

□

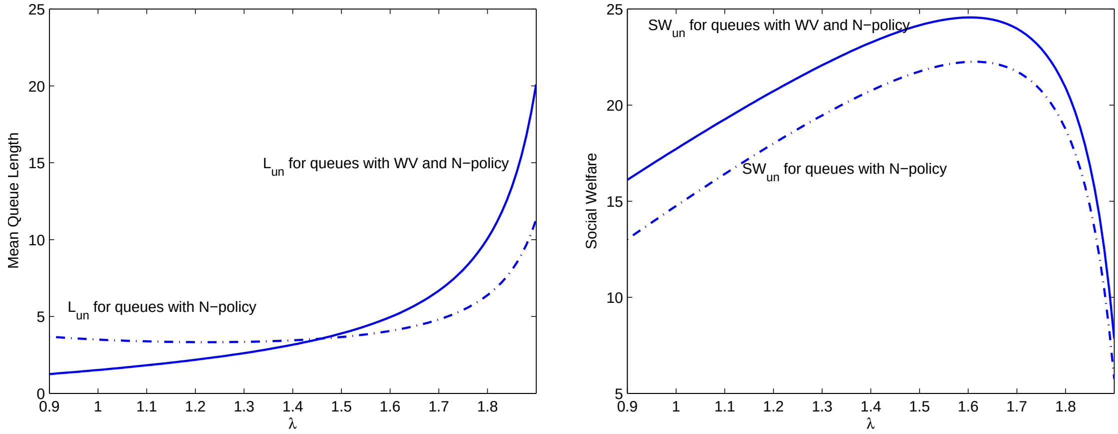

FIGURE 6. Mean queue length L_{un} and social welfare SW_{un} for unobservable queues with N-policy and queues with working vacations and N-policy($\mu_0 = 1.5$, $\mu_1 = 2$, $C = 1.5$, $R = 20$).

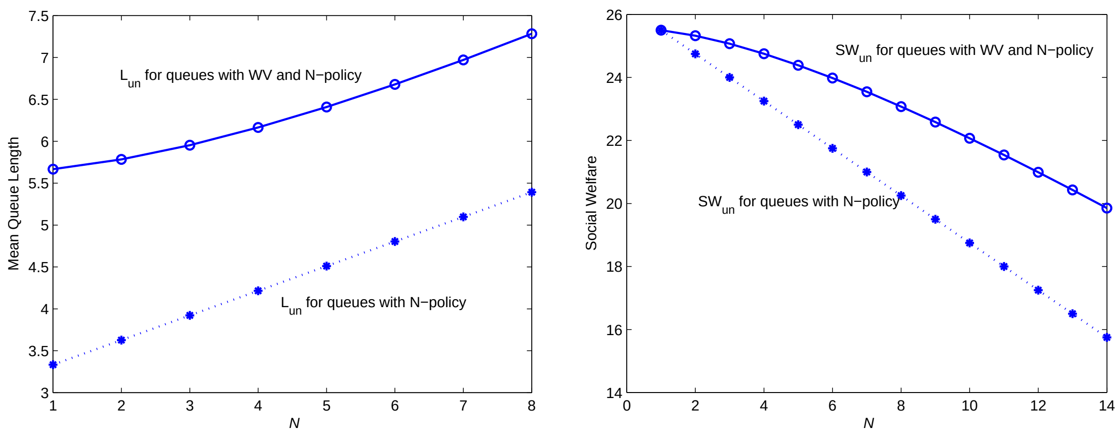

FIGURE 7. Mean queue length L_{un} and social welfare SW_{un} for unobservable queues with N-policy and queues with working vacations and N-policy($\mu_0 = 1.5$, $\mu_1 = 2$, $\lambda = 1.7$, $C = 1.5$, $R = 20$).

Lemma 4.2. *In the unobservable M/M/1 queue with working vacations and N-policy where the arriving customers have no information of the queue length upon arrival, the mean sojourn time of customer $E[W_{un}]$ and the mean queue length L_{un} is given below:*

According to Little's Law,

$$E[W_{un}] = \frac{L_{un}}{\lambda} \tag{4.9}$$

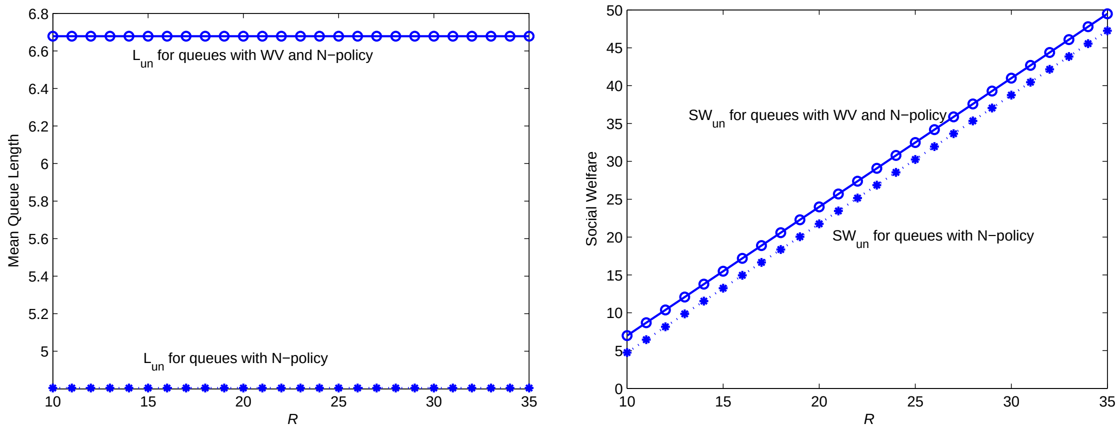

FIGURE 8. Mean queue length L_{un} and social welfare SW_{un} for unobservable queues with N-policy and queues with working vacations and N-policy($\mu_0 = 1.5$, $\mu_1 = 2$, $\lambda = 1.7$, $C = 1.5$, $N = 6$).

The mean queue length is:

$$L_{un} = \sum_{k=1}^{N-1} kP_{un}(k,0) + \sum_{k=1}^{N} kP_{un}(k,1) + \sum_{k=1}^{+\infty}(N+k)P_{un}(N+k,1)$$

$$= \frac{1}{\rho_0^N(1-\rho_0)}\frac{\mu_1}{\mu_0}P_{un}(1,1)\left[\rho_0\frac{1-N\rho_0^{N-1}+N\rho_0^N-\rho_0^N}{(1-\rho_0)^2} - \frac{N(N-1)}{2}\rho_0^N\right]$$

$$+\frac{1}{1-\rho_1}P_{un}(1,1)\left[\frac{N(N-1)}{2} - \rho_1\frac{1-N\rho_1^{N-1}+N\rho_1^N-\rho_1^N}{(1-\rho_1)^2}\right] + \frac{1-\rho_1^N}{1-\rho_1}P_{un}(1,1)\left(\frac{N}{1-\rho_1}+\frac{\rho_1}{(1-\rho_1)^2}\right)$$

$$= \frac{1}{\rho_0^N(1-\rho_0)}\frac{\mu_1}{\mu_0}\left[\rho_0\frac{1-N\rho_0^{N-1}+N\rho_0^N-\rho_0^N}{(1-\rho_0)^2} - \frac{N(N-1)}{2}\rho_0^N\right]P_{un}(1,1)$$

$$+\frac{1}{1-\rho_1}(\frac{N(N+1)}{2}+\frac{N\rho_1}{1-\rho_1})P_{un}(1,1) \tag{4.10}$$

Lemma 4.3. *In the unobservable M/M/1 queue with working vacations and N-policy where the arriving customers have no information of the queue length upon arrival, the social welfare per time unit SW_{un} is given below*:

$$SW_{un} = (R - CE[W_{un}])\lambda = R\lambda - CL_{un}$$

4.2. Numerical simulation for mean queue length and social welfare

In this section, the effect of N, R and λ on the mean queue length L_{un} and social welfare SW_{un} will be graphically presented. What's more, comparison work is carried out to show which mechanism, N-policy or working vacations and N-policy, is better in controlling the mean queue length and improving the social welfare.

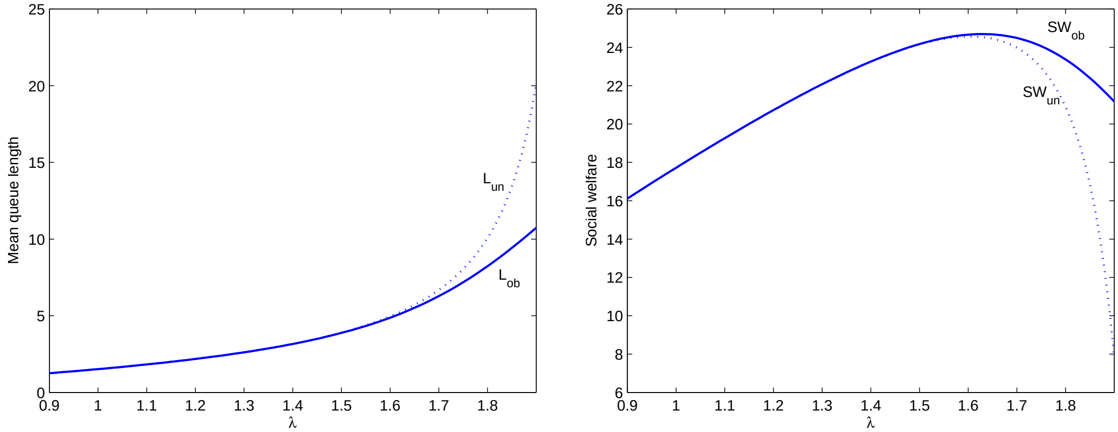

FIGURE 9. Mean queue length and social welfare for observable and unobservable queues with working vacations and N-policy ($\mu_0 = 1.5$, $\mu_1 = 2$, $C = 1.5$, $N = 6$, $R = 20$).

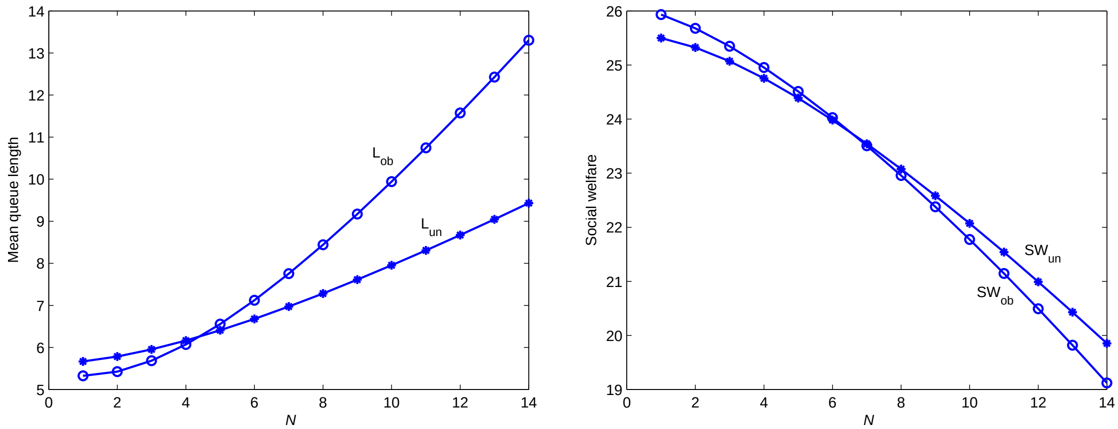

FIGURE 10. Mean queue length and social welfare for observable and unobservable queues with working vacations and N-policy ($\mu_0 = 1.5$, $\mu_1 = 2$, $\lambda = 1.7$, $C = 1.5$, $R = 20$).

Figure 6 shows that the mean queue length L_{un} increases with λ, the social welfare SW_{un} first increases then decreases with λ in queues with working vacations and N-policy. With λ increasing, more customers join the system per time unit, the queue length become larger, the mean queue length increases. Once there are N customers in the system, the server resumes to normal work period, there will be more customers completing service per time unit, the social welfare increases. However, as λ keeps increasing, there will be congestion in the system, customers will suffer longer queue length and more waiting cost, which has a negative effect on the social welfare, social welfare decreases.

Figure 7 shows that the mean queue length L_{un} increases with N, the social welfare SW_{un} decreases with N in queues with working vacations and N-policy. Because with a lager N, it's more difficult for the server returning to normal service period, there will be more customers served with the vacation service rate μ_0, customers who join the queue suffer longer waiting time and longer queue length, the mean queue length becomes larger. Longer waiting delay causes more waiting cost of customers, thus the social welfare decreases with N.

Figure 8 shows that the mean queue length L_{un} remains fixed with R and equation (4.10) demonstrates this, the social welfare SW_{un} increases with R in queues with working vacations and N-policy. Because

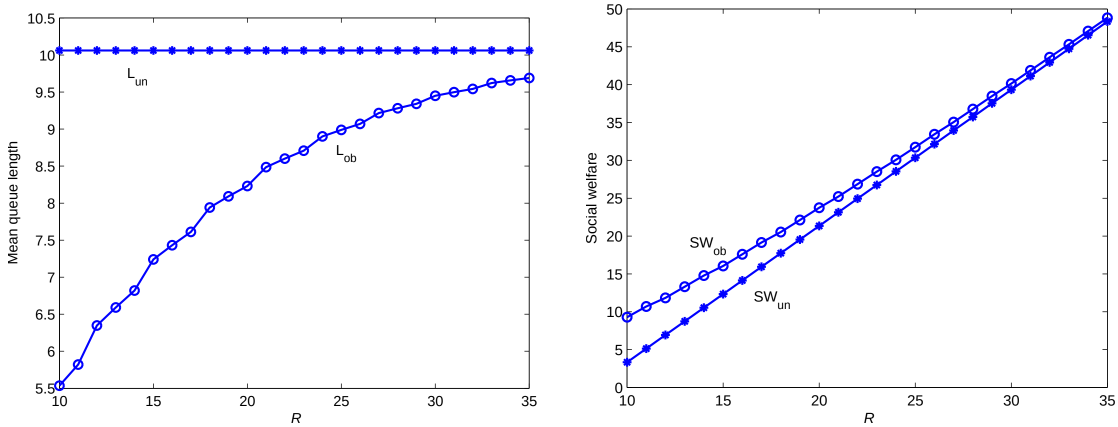

FIGURE 11. Mean queue length and social welfare for observable and unobservable queues with working vacations and N-policy ($\mu_0 = 1.5$, $\mu_1 = 2$, $\lambda = 1.7$, $C = 1.5$, $N = 6$).

with a larger R, customers receive higher reward after service. Social welfare SW_{un}, as the sum of customer benefit, increases with R necessarily.

On the other hand, Figures 6–8 show that the mean queue length for queues with N-policy is always larger than that for queues with working vacations and N-policy except in the left of Figure 6, the social welfare for queues with working vacations and N-policy is always greater than that for queues with N-policy. Because in queues with N-policy, the server doesn't serve customers in vacation period, customers suffer a longer waiting time and a greater waiting cost; while in queues with working vacations and N-policy, customers can be served with a lower service rate during vacation period.

4.3. Numerical simulation for observable case and unobservable case

Figures 9–11 focus on comparing the mean queue length and the social welfare for observable case and unobservable case in queues with working vacations and N-policy. Figure 9 shows that when λ is small, the mean queue length and the social welfare are the same in observable case and unobservable case; when λ keeps increasing, L_{ob} is smaller than L_{un} and SW_{ob} is greater than SW_{un}, which means that the observable case is better in decreasing the mean queue length and improving the social welfare. Figure 10 show that when N is small, L_{ob} is smaller than L_{un} and SW_{ob} is greater than SW_{un}; As N keeps increasing, L_{ob} is greater than L_{un} and SW_{ob} is smaller than SW_{un}. Figure 11 shows that L_{ob} is smaller than L_{un} and SW_{ob} is greater than SW_{un}. We can make the summary that when N is small, the observable case has an advantage in avoiding congestion and improving the social welfare; otherwise, the unobservable case is better.

5. CONCLUSION

In this paper, the customer social welfare in a single server queue with working vacations and N-policy under observable and unobservable cases is analyzed. Comparison work shows that queues with working vacations and N-policy lead to smaller mean queue length and larger social welfare than queues with N-policy. The main reason is that the arriving customers who decide to join a queue with working vacations and N-policy can be served during vacation period. A risk neutral customer prefers to join a queue with working vacations and N-policy. However, from the perspective of agent managers, running a server during vacation period will lead to extra cost in maintaining it. The future work could take both customer and manager's strategies into consideration. What's more, we give some advice from a managerial point of view on how to choose the information level through the numerical results.

Acknowledgements. The authors thank the editor and the anonymous reviewers for their valuable comments and suggestions to improve the paper. The first author thanks China Scholarship Council for supporting her visit to Carleton University(Ottawa, CA). This research was supported by the National Natural Science Foundation of China (No. 71301091), and the Program for the Outstanding Innovation Teams of Higher Learning Institute of Shanxi (OIT).

References

[1] A. Banks, J. Vincent and C. Anyakoha, A review of particle swarm optimization. Part I: background and development. *Natural Comput.* **6** (2007) 467–484.

[2] M. Clerc, Particle Swarm Optimization.Vol. 93. John Wiley and Sons (2010).

[3] B.T. Doshi, Queueing systems with vacationsa survey. *Queueing Syst.* **1** (1986) 29–66.

[4] R.C. Eberhart and J. Kennedy, A new optimizer using particle swarm theory. In *Proc. 6th Inter. Symp. Micro Mach. Human Sci.* **1** (1995) 39–43.

[5] N.M. Edelson and D.K. Hilderbrand, Congestion tolls for Poisson queuing processes. *Econ.: J. Econ. Soc.* (1975) 81–92.

[6] P. Guo and R. Hassin, Strategic behavior and social optimization in Markovian vacation queues. *Oper. Res.* **59** (2011) 986–997.

[7] P. Guo and R. Hassin, Strategic behavior and social optimization in Markovian vacation queues: the case of heterogeneous customers. *Eur. J. Oper. Res.* **222** (2012) 278–286.

[8] P. Guo and Q. Li, Strategic behavior and social optimization in partially-observable Markovian vacation queues. *Oper. Res. Lett.* **41** (2013) 277–284.

[9] R. Hassin and M. Haviv, To queue or not to queue: Equilibrium behavior in queueing systems. Vol. 59. Springer Science and Business Media (2003).

[10] E. Kasim and G. Gupur, Dynamic Analysis of the $M/G/1$ Queueing Model with Single Working Vacation. *Inter. J. Appl. Comput. Math.* (2016) 1–31.

[11] O. Kella, The threshold policy in the $M/G/1$ queue with server vacations. *Naval Res. Logistics* (NRL) **36** (1989) 111–123.

[12] J. Li and N. Tian, The $M/M/1$ queue with working vacations and vacation interruptions. *J. Syst. Sci. Syst. Eng.* **16** (2007) 121–127.

[13] J. Li and N. Tian, Performance analysis of a $GI/M/1$ queue with single working vacation. *Appl. Math. Comput.* **217** (2011) 4960–4971.

[14] J.H. Li, N.S. Tian and Z.Y. Ma, Performance analysis of $GI/M/1$ queue with working vacations and vacation interruption. *Appl. Math. Model.* **32** (2008) 2715–2730.

[15] J.H. Li, N.S. Tian, Z.G. Zhang and H.P. Luh, Analysis of the $M/G/1$ queue with exponentially working vacationsa matrix analytic approach. *Queueing Syst.* **61** (2009) 139–166.

[16] W.Y. Liu, X.L. Xu and N.S. Tian, Stochastic decompositions in the $M/M/1$ queue with working vacations. *Oper. Res. Lett.* **35** (2007) 595–600.

[17] W. Liu, Y. Ma and J. Li, Equilibrium threshold strategies in observable queueing systems under single vacation policy. *Appl. Math. Model.* **36** (2012) 6186–6202.

[18] W. Liu, Q. Ma and J. Li, Analysis of customer behaviors in $M/M/1$ queueing systems with N-policy and working vacation. *Syst. Engineer–Theory Practice* **36** (2016) 1848–1856.

[19] Y. Ma, W.Q. Liu and J.H. Li, Equilibrium balking behavior in the $Geo/Geo/1$ queueing system with multiple vacations. *Appl. Math. Model.* **37** (2013) 3861–3878.

[20] P. Naor, The regulation of queue size by levying tolls. *Econ.: J. Econ. Soc.* (1969) 15–24.

[21] R. Poli, J. Kennedy and T. Blackwell, Particle swarm optimization. *Swarm Intelligence* **1** (2007) 33–57.

[22] L.D. Servi and S.G. Finn, $M/M/1$ queues with working vacations $(m/m/1/wv)$. *Performance Evaluation* **50** (2002) 41–52.

[23] W. Sun, S. Li and E. Cheng−Guo, Equilibrium and optimal balking strategies of customers in Markovian queues with multiple vacations and N-policy. *Appl. Math. Model.* **40** (2016) 284–301.

[24] H. Tagaki, Queueing Analysis–a Foundation of Performance Evaluation, Vol. 3. Discrete-Time Systems (1993).

[25] N. Tian and Z.G. Zhang, Vacation queueing models: Theory and Applications. Vol. 93. Springer Science and Business Media (2006).

[26] J. Wang, X. Zhang and P. Huang, Strategic behavior and social optimization in a constant retrial queue with the N-policy. *Eur. J. Oper. Res.* **256** (2017) 841–849.

[27] W. Wu, Y. Tang and M. Yu, Analysis of an $M/G/1$ queue with N-policy, single vacation, unreliable service station and replaceable repair facility. *Opsearch* **52** (2015) 670–691.

[28] M. Yadin and P. Naor, Queueing systems with a removable service station. *J. Oper. Res. Soc.* **14** (1963) 393-405.

[29] D.Y. Yang and C.H. Wu, Cost-minimization analysis of a working vacation queue with N-policy and server breakdowns. *Comput. Industrial Eng.* **82** (2015) 151–158.

[30] F. Zhang, J. Wang and B. Liu, Equilibrium balking strategies in Markovian queues with working vacations. *Appl. Math. Model.* **37** (2013) 8264–8282.