ICIC Express Letters
Volume **10**, Number **11**, November **2016**

 ISSN 1881-803X
pp. **2735–2741**

PERFORMANCE ANALYSIS AND OPTIMIZATION OF THE (N, n)-PREEMPTIVE PRIORITY QUEUE WITH MULTIPLE WORKING VACATION

Zhanyou Ma, Xiaoming Zheng, Minjie Xu and Wenbo Wang

College of Science
Yanshan University
No. 438, Hebei Street, Qinhuangdao 066004, P. R. China
mzhy55@ysu.edu.cn

Received April 2016; accepted July 2016

Abstract. *In this paper, we consider a discrete time Geom/Geom/1 queue system with (N, n)-preemptive priority discipline and multiple synchronization working vacation. A discrete time three-dimension Markov chain (MC) of this queue system is given. By using the quasi birth and death chain and matrix-geometric solution theory, the average queue length of the two classes and the probability of a customer I being preempted are obtained. In the end, some numerical and optimization results are provided to illustrate the effect of the parameters on several performance characteristics.*
Keywords: Preemptive priority, Threshold, Quasi birth and death chain, Matrix-geometric solution, Optimization

1. **Introduction.** There are two fundamental priority disciplines in queueing system, non-preemptive (NP) and preemptive disciplines. The NP and preemptive disciplines are both extreme cases when a high-priority customer arrives during the provision of service to a low-priority customer. For instance, Guo and Liu researched a multi-class single-server queue model, which has a preemptive priority service discipline and K customer classes in [1]. Peköz analyzed a multi-server non-preemptive queue, and there was a decision maker, who decided when waiting customers could enter service in [2]. In order to balance this situation, many scholars had done a lot of trying. Kim first introduced the (N, n)-preemptive priority discipline in [3]; the preemption of the service of a low-class customer is determined by two thresholds N and n of the queue length of high-class customers. And the qualities for two types of customers can be controlled within a certain bound. This paper expands the strategy to a discrete time Geom/Geom/1 queue system.

When the server is idle, it may cause waste of resources to a certain extent, many researchers have done a lot to conserve the system resource. Servi and Finn first introduced working vacation policy in [4], in which the server worked with a lower service rate rather than stopping the service completely during a vacation period. Many scholars changed the arriving or serving distribution and added many other policies on this working vacation queue model, and improved the server utility in [5, 6, 7, 8]. In general, most of the articles are based on the two-dimension Markov process or two-dimension MC, considered the length of the customer and the state of the system. Our main work is to study a three-dimension MC composed by the average length of the two classes customers and the state of the server and combine (N, n)-preemptive priority discipline and working vacation policy on the basis of the Geom/Geom/1 queue model. And this research makes the theory closer to the reality.

The paper is organized as follows. In Section 2, the mathematical model of this queue system is given. In Section 3, the transition probability matrix and the existence condition of the steady-state distribution are analyzed. In Section 4, some performance measures

are considered. In Section 5, some numerical examples and optimization results are shown. In Section 6, conclusions are given.

2. **Mathematical Model.** In this model, customers are assumed to arrive in a queue system with a single server, where there is infinite buffer space for customer I and limited buffer space with the fixed number K for customer II. Both of the inter-arrival time, the service time and the vacation time are assumed to be mutually independent sequence. The queue model is referred as the late arrival system with delayed access based on the Entrance Protocols. In this model, let $\bar{x}$ be $1-x$, for $\forall x \in [0,1]$, and the specific description for this model is as follows.

(1) Suppose that a potential customer arrival occurs in the interval (t^-, t), $t = 0, 1, \ldots$. The probability of a customer arrival occurring in a slot is λ $(0 \leq \lambda \leq 1)$, the arriving customer is customer II with probability α $(0 \leq \alpha \leq 1)$, the arriving customer is customer I with probability $\bar{\alpha}$. The inter-arrival time T_1 and T_2 of customer I and customer II follow geometric distributions with parameters $\lambda\bar{\alpha}$ and $\lambda\alpha$ as follows

$$P\{T_1 = j\} = \lambda\bar{\alpha}(1-\lambda\bar{\alpha})^{j-1}, \quad P\{T_2 = j\} = \lambda\alpha(1-\lambda\alpha)^{j-1}, \quad j = 1, 2, \ldots$$

(2) The potential service occurs in the interval (t, t^+). The service time S_1, S_2 follow geometric distributions with parameters μ_1 and μ_2 $(0 < \mu_1,\ \mu_2 < 1)$ as follows

$$P\{S_1 = j\} = \mu_1\bar{\mu}_1^{j-1}, \quad P\{S_2 = j\} = \mu_2\bar{\mu}_2^{j-1}, \quad j = 1, 2, \ldots$$

(3) When the system is empty, the server would be in a working vacation period. If there is no customer when the working vacation period ends, the server will enter into the next working vacation period, otherwise, the vacation period is over and the new busy period is coming. The vacation time V follows geometric distribution with parameter θ $(0 < \theta < 1)$, the working vacation service time S_{1v}, S_{2v} follow geometric distributions with parameters μ_{1v} and μ_{2v} $(0 < \mu_{1v},\ \mu_{2v} < 1)$ as follows

$$P(V = j) = \theta\bar{\theta}^{j-1}, \quad P\{S_{1v} = j\} = \mu_{1v}\bar{\mu}_{1v}^{j-1}, \quad P\{S_{2v} = j\} = \mu_{2v}\bar{\mu}_{2v}^{j-1}, \quad j = 1, 2, \ldots$$

(4) In this system, when the server is serving for customer I, if the number of customer II (including the new customer II) in the system reaches the upper threshold N $(N \geq 1)$, the service for the customer I will be preempted, and this preempted service will be restarted as soon as the number of customer II in the system decreases to a certain lower threshold n $(0 \leq n \leq N-1)$. If the number of customer II in the system does not reach the upper threshold N, the customer II will not be served until the number of customer II in the system reaches the upper threshold N or the customer I has completely been served. We assumed that the service order is First-Come First-Served (FCFS) discipline.

3. **Analysis of the State Transition.** Let $L_1(t^+), L_2(t^+)$ represent the number of customer I and II in the system at time t^+, and

$$J_t = \begin{cases} 0, & \text{the instant of } t^+ \text{ is in working vacation period,} \\ 1, & \text{the instant of } t^+ \text{ is in busy period.} \end{cases}$$

Then we can obtain that $\{(L_1(t^+), L_2(t^+), J_t), t \geq 1\}$ is a discrete time three-dimension MC in this queue system and its state space is as follows

$$\Omega = \{(0,0,0)\} \cup \{(0,l,j),\ 1 \leq l \leq K,\ j = 0,1\} \cup \{(i,l,j),\ i \geq 1,\ 0 \leq l \leq K,\ j = 0,1\}.$$

All possible states: $(i,0,0)$, $(i,0,1)$, $(i,1,1)$, ..., $(i,K,0)$, $(i,K,1)$ are called level i, where $i \geq 1$. Specifically, level 0 has states: $(0,0,0)$, $(0,1,0)$, $(0,1,1)$, ..., $(0,K,0)$, $(0,K,1)$.

The transition probability matrix can be written as follows

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{A}_{00} & \boldsymbol{A}_{01} & & & \\ \boldsymbol{A}_{10} & \boldsymbol{A}_1 & \boldsymbol{A}_2 & & \\ & \boldsymbol{A}_0 & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \\ & & \ddots & \ddots & \ddots \end{bmatrix} \tag{1}$$

where $\boldsymbol{A}_{00}$, $\boldsymbol{A}_{01}$, $\boldsymbol{A}_{10}$, $\boldsymbol{A}_0$, $\boldsymbol{A}_1$, $\boldsymbol{A}_2$ represent state transition probability between levels. We define symbols δ_1, δ_{1v}, δ_2, δ_{2v}, δ_3, δ_{3v}, δ_4, δ_{4v}, as follows

$$\delta_1 = \bar{\lambda}\bar{\mu}_2 + \lambda\alpha\mu_2, \quad \delta_{1v} = \bar{\lambda}\bar{\mu}_{2v} + \lambda\alpha\mu_{2v}, \quad \delta_2 = \bar{\mu}_2 - \lambda\bar{\alpha}\bar{\mu}_2 + \lambda\alpha\mu_2,$$
$$\delta_{2v} = \bar{\mu}_{2v} - \lambda\bar{\alpha}\bar{\mu}_{2v} + \lambda\alpha\mu_{2v}, \quad \delta_3 = \bar{\lambda}\bar{\mu}_1 + \mu_1\lambda\bar{\alpha}, \quad \delta_{3v} = \bar{\lambda}\bar{\mu}_{1v} + \mu_{1v}\lambda\bar{\alpha},$$
$$\delta_4 = \bar{\lambda}(1-\mu_1-\mu_2) + \lambda\alpha\mu_2 + \lambda\bar{\alpha}\mu_1, \quad \delta_{4v} = \bar{\lambda}(1-\mu_{1v}-\mu_{2v}) + \lambda\alpha\mu_{2v} + \lambda\bar{\alpha}\mu_{1v}.$$

The details are as follows

$$\boldsymbol{A}_{00} = \begin{bmatrix} \bar{\lambda} & \lambda\alpha\bar{\theta} & \lambda\alpha\theta & & & & & \\ \bar{\lambda}\mu_{2v} & \delta_{1v}\bar{\theta} & \delta_{1v}\theta & \lambda\alpha\bar{\mu}_{2v}\bar{\theta} & \lambda\alpha\bar{\mu}_{2v}\theta & & & \\ \bar{\lambda}\mu_2 & 0 & \delta_1 & 0 & \lambda\alpha\bar{\mu}_2 & & & \\ & \bar{\lambda}\mu_{2v}\bar{\theta} & \bar{\lambda}\mu_{2v}\theta & \delta_{1v}\bar{\theta} & \delta_{1v}\theta & \lambda\alpha\bar{\mu}_{2v}\bar{\theta} & \lambda\alpha\bar{\mu}_{2v}\theta & \\ & 0 & \bar{\lambda}\mu_2 & 0 & \delta_1 & 0 & \lambda\alpha\bar{\mu}_2 & \\ & & & & \ddots & \ddots & \ddots & \\ & & & ... & 0 & \delta_1 & 0 & \lambda\alpha\bar{\mu}_2 \\ & & & & \bar{\lambda}\mu_{2v}\bar{\theta} & \bar{\lambda}\mu_{2v}\theta & \delta_{2v}\bar{\theta} & \delta_{2v}\theta \\ & & & & 0 & \bar{\lambda}\mu_2 & 0 & \delta_2 \end{bmatrix},$$

$$\boldsymbol{A}_{10} = \begin{bmatrix} \bar{\lambda}\mu_{1v} & \lambda\alpha\mu_{1v}\bar{\theta} & \lambda\alpha\mu_{1v}\theta & & & & & & & \\ \bar{\lambda}\mu_1 & 0 & \lambda\alpha\mu_1 & & & & & & & \\ & \bar{\lambda}\mu_{1v}\bar{\theta} & \bar{\lambda}\mu_{1v}\theta & \lambda\alpha\mu_{1v}\bar{\theta} & \lambda\alpha\mu_{1v}\theta & & & & & \\ & 0 & \bar{\lambda}\mu_1 & 0 & \lambda\alpha\mu_1 & & & & & \\ & & & \ddots & \ddots & \ddots & \ddots & & & \\ & & & \bar{\lambda}\mu_{1v}\bar{\theta} & \bar{\lambda}\mu_{1v}\theta & \lambda\alpha\mu_{1v}\bar{\theta} & \lambda\alpha\mu_{1v}\theta & & & \\ & & & 0 & \bar{\lambda}\mu_1 & 0 & \lambda\alpha\mu_1 & & & \\ & & & & 0 & 0 & 0 & 0 & & \\ & & & & & \ddots & \ddots & \ddots & \ddots & \\ & & & & & & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\boldsymbol{A}_0 = \begin{bmatrix} \bar{\lambda}\mu_{1v}\bar{\theta} & \bar{\lambda}\mu_{1v}\theta & \lambda\alpha\mu_{1v}\bar{\theta} & \lambda\alpha\mu_{1v}\theta & & & & & \\ 0 & \bar{\lambda}\mu_1 & 0 & \lambda\alpha\mu_1 & & & & & \\ & \ddots & \ddots & \ddots & \ddots & & & & \\ & & \bar{\lambda}\mu_{1v}\bar{\theta} & \bar{\lambda}\mu_{1v}\theta & \lambda\alpha\mu_{1v}\bar{\theta} & \lambda\alpha\mu_{1v}\theta & & & \\ & & 0 & \bar{\lambda}\mu_1 & 0 & \lambda\alpha\mu_1 & & & \\ & & & 0 & 0 & 0 & 0 & & \\ & & & & \ddots & \ddots & \ddots & \ddots & \\ & & & & & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\boldsymbol{A}_1 = \begin{bmatrix}
\delta_{3v}\bar{\theta} & \delta_{3v}\theta & \lambda\alpha\bar{\mu}_{1v}\bar{\theta} & \lambda\alpha\bar{\mu}_{1v}\theta & & & & & & & & & & \\
0 & \delta_3 & 0 & \lambda\alpha\bar{\mu}_1 & & & & & & & & & & \\
 & & \delta_{3v}\bar{\theta} & \delta_{3v}\theta & \lambda\alpha\bar{\mu}_{1v}\bar{\theta} & \lambda\alpha\bar{\mu}_{1v}\theta & & & & & & & & \\
 & & 0 & \delta_3 & 0 & \lambda\alpha\bar{\mu}_1 & & & & & & & & \\
 & & \ddots & \ddots & \ddots & \ddots & & & & & & & & \\
 & & & \delta_{3v}\bar{\theta} & \delta_{3v}\bar{\theta} & \lambda\alpha\bar{\mu}_{1v}\bar{\theta} & \lambda\alpha\bar{\mu}_{1v}\theta & & & & & & & \\
 & & & 0 & \delta_3 & 0 & \lambda\alpha\bar{\mu}_1 & & & & & & & \\
 & & & \bar{\lambda}\mu_{2v}\bar{\theta} & \bar{\lambda}\mu_{2v}\theta & \delta_{4v}\bar{\theta} & \delta_{4v}\theta & \lambda\alpha\bar{\mu}_v\bar{\theta} & \lambda\alpha\bar{\mu}_v\theta & & & & & \\
 & & & 0 & \bar{\lambda}\mu_2 & 0 & \delta_4 & \ddots & \lambda\alpha\bar{\mu} & \ddots & & & & \\
 & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & & & \\
 & & & & \bar{\lambda}\mu_{2v}\bar{\theta} & \bar{\lambda}\mu_{2v}\theta & \delta_{4v}\bar{\theta} & \delta_{4v}\theta & \lambda\alpha\bar{\mu}_v\bar{\theta} & \lambda\alpha\bar{\mu}_v\theta & & & & \\
 & & & & 0 & \bar{\lambda}\mu_2 & 0 & \delta_4 & 0 & \lambda\alpha\bar{\mu} & & & & \\
 & & & & & & \bar{\lambda}\mu_{2v}\bar{\theta} & \bar{\lambda}\mu_{2v}\theta & \delta_{1v}\bar{\theta} & \delta_{1v}\theta & \lambda\alpha\bar{\mu}_{2v}\bar{\theta} & \lambda\alpha\bar{\mu}_{2v}\theta & & \\
 & & & & & & 0 & \bar{\lambda}\mu_2 & 0 & \delta_1 & 0 & \lambda\alpha\bar{\mu}_2 & & \\
 & & & & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\
 & & & & & & & 0 & \bar{\lambda}\mu_2 & 0 & \delta_1 & 0 & \lambda\alpha\bar{\mu}_2 & \\
 & & & & & & & & & \bar{\lambda}\mu_{2v}\bar{\theta} & \bar{\lambda}\mu_{2v}\theta & \delta_{2v}\bar{\theta} & \delta_{2v}\theta & \\
 & & & & & & & & & 0 & \bar{\lambda}\mu_2 & 0 & \delta_2 &
\end{bmatrix},$$

$$\boldsymbol{A}_2 = \begin{bmatrix}
\lambda\bar{\alpha}\bar{\mu}_{1v}\bar{\theta} & \lambda\bar{\alpha}\bar{\mu}_{1v}\theta & & & & & & \\
0 & \lambda\bar{\alpha}\bar{\mu}_1 & & & & & & \\
\ddots & \ddots & & & & & & \\
 & \lambda\bar{\alpha}\bar{\mu}_{1v}\bar{\theta} & \lambda\bar{\alpha}\bar{\mu}_{1v}\theta & & & & & \\
 & 0 & \lambda\bar{\alpha}\bar{\mu}_1 & & & & & \\
 & \lambda\bar{\alpha}\mu_{2v}\bar{\theta} & \lambda\bar{\alpha}\mu_{2v}\theta & \lambda\bar{\alpha}\bar{\mu}_v\bar{\theta} & \lambda\bar{\alpha}\bar{\mu}_v\theta & & & \\
 & 0 & \lambda\bar{\alpha}\mu_2 & 0 & \lambda\bar{\alpha}\bar{\mu} & & & \\
 & \ddots & \ddots & \ddots & \ddots & & & \\
 & & \lambda\bar{\alpha}\mu_{2v}\bar{\theta} & \lambda\bar{\alpha}\mu_{2v}\theta & \lambda\bar{\alpha}\bar{\mu}_v\bar{\theta} & \lambda\bar{\alpha}\bar{\mu}_v\theta & & \\
 & & 0 & \lambda\bar{\alpha}\mu_2 & 0 & \lambda\bar{\alpha}\bar{\mu} & & \\
 & & \lambda\bar{\alpha}\mu_{2v}\bar{\theta} & \lambda\bar{\alpha}\mu_{2v}\theta & \lambda\bar{\alpha}\bar{\mu}_{2v}\bar{\theta} & \lambda\bar{\alpha}\bar{\mu}_{2v}\theta & & \\
 & & 0 & \lambda\bar{\alpha}\mu_2 & 0 & \lambda\bar{\alpha}\bar{\mu}_2 & & \\
 & & \ddots & \ddots & \ddots & \ddots & & \\
 & & & \lambda\bar{\alpha}\mu_{2v}\bar{\theta} & \lambda\bar{\alpha}\mu_{2v}\theta & \lambda\bar{\alpha}\bar{\mu}_{2v}\bar{\theta} & \lambda\bar{\alpha}\bar{\mu}_{2v}\theta & \\
 & & & 0 & \lambda\bar{\alpha}\mu_2 & 0 & \lambda\bar{\alpha}\bar{\mu}_2 &
\end{bmatrix}.$$

According to the block structure of transition probability matrix, we can obtain that $\{(L_1(t^+), L_2(t^+), J_t), t \geq 1\}$ is a quasi birth and death chain. When the MC is positive recurrent, the steady-state distribution is indicated as

$$\pi_{i,l,j} = \lim_{t\to\infty} P\{L_1(t^+) = i, L_2(t^+) = l, J_t = j\},\ (i,l,j) \in \Omega,\ \boldsymbol{\Pi} = (\boldsymbol{\pi_0}, \boldsymbol{\pi_1}, \boldsymbol{\pi_2}, \ldots)$$

where

$$\boldsymbol{\pi}_0 = (\pi_{0,0,0}, \pi_{0,1,0}, \pi_{0,1,1}, \ldots, \pi_{0,K,0}, \pi_{0,K,1}),$$
$$\boldsymbol{\pi}_i = (\pi_{i,0,0}, \pi_{i,0,1}, \pi_{i,1,0}, \pi_{i,1,1}, \ldots, \pi_{i,K,0}, \pi_{i,K,1}),\ i \geq 1.$$

The necessary and sufficient conditions that $\{(L_1(t^+), L_2(t^+), J_t),\ t \geq 1\}$ is positive recurrent are the matrix quadratic equation $\boldsymbol{R}^2\boldsymbol{A}_0 + \boldsymbol{R}\boldsymbol{A}_1 + \boldsymbol{A}_2 = \boldsymbol{R}$ has a minimal non-negative solution $\boldsymbol{R}$ and the spectral radius $\mathrm{SP}(\boldsymbol{R}) < 1$, and the $(4K+3)$ dimensional matrix

$$B[\boldsymbol{R}] = \begin{bmatrix} \boldsymbol{A}_{00} & \boldsymbol{A}_{01} \\ \boldsymbol{A}_{10} & \boldsymbol{R}\boldsymbol{A}_0 + \boldsymbol{A}_1 \end{bmatrix} \tag{2}$$

has left-invariant vector. When MC is positive recurrent, its stationary distribution is

$$\begin{cases} (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1) B[\boldsymbol{R}] = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1), \\ \boldsymbol{\pi}_0\boldsymbol{e} + \boldsymbol{\pi}_1(\boldsymbol{I} - \boldsymbol{R})^{-1}\boldsymbol{e} = 1, \\ \boldsymbol{\pi}_i = \boldsymbol{\pi}_1\boldsymbol{R}^{i-1},\ i \geq 1 \end{cases} \tag{3}$$

where $\boldsymbol{e}$ is an appropriate dimensional column vector with all element being equal to one. The proof of Equation (3) can be obtained by using equilibrium equation $\boldsymbol{\Pi P} = \boldsymbol{\Pi}$ and matrix-geometric solution method presented in [9].

4. **Performance Measures.** According to the results from Section 3, we can obtain the average queue length of customer I and customer II, the probability of a customer I being preempted and so on.

(1) The average queue length of customer I is given by

$$E[L_1] = \sum_{i=0}^{\infty} iP(L_1 = i) = \sum_{i=1}^{\infty} i\boldsymbol{\pi}_i\boldsymbol{e} = \sum_{i=1}^{\infty} i \sum_{l=0}^{K} \sum_{j=0}^{1} \pi_{i,l,j}.$$

(2) The average queue length of customer II is given by

$$E[L_2] = \sum_{l=0}^{K} lP(L_2 = l) = \sum_{l=1}^{K} l \sum_{i=0}^{\infty} \sum_{j=0}^{1} \pi_{i,l,j}.$$

(3) The probability that a customer I is preempted by customer II is given by

$$P_1 = \sum_{j=0}^{1} \sum_{i=0}^{\infty} \pi_{i,N,j}.$$

(4) The probability that the server is released from customer II is given by

$$P_2 = \sum_{j=0}^{1} \sum_{i=1}^{\infty} \pi_{i,n,j}.$$

5. **Numerical and Optimization Results.** Firstly, we provide some numerical results to describe the effect of parameters on performance measures. Take $\alpha = 0.4$, $\mu_1 = 0.3$, $\mu_2 = 0.6$, $\mu_{1v} = 0.05$, $\mu_{2v} = 0.1$, $\theta = 0.3$, $K = 8$.

In Figure 1, taking $n = 3$, we can find that $E[L_1]$ increases with the increase of λ. When λ is unchanged, $E[L_1]$ decreases with the increase of the value N. That is mainly because N increases, customer II has less chances to preempt the service from customer I, therefore, customer I has more chances to get service, and the number of customer I decreases. In Figure 2, taking $n = 3$, we can find that P_1 increases with the increase of λ. When λ is unchanged, P_1 increases with the increase of the value n. That is mainly because with the increase of n, the frequency of server transformation between two types

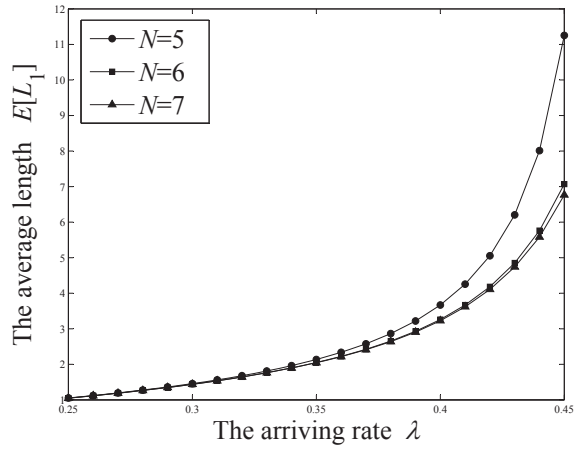

FIGURE 1. The relation of $E[L_1]$ with λ and N

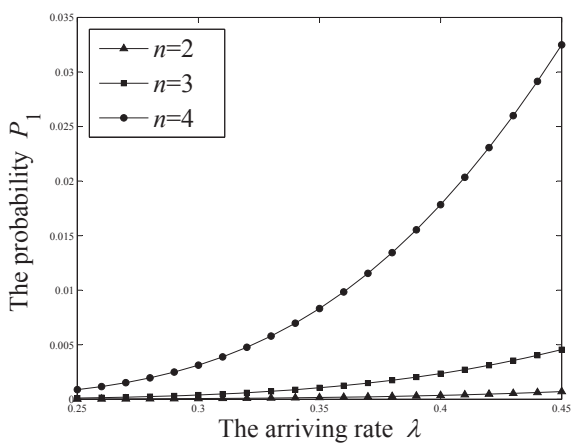

FIGURE 2. The relation of P_1 with λ and n

TABLE 1. The relation of P_2 with μ_2 and N

N	The probability P_2				
	$\mu_2 = 0.4$	$\mu_2 = 0.45$	$\mu_2 = 0.5$	$\mu_2 = 0.55$	$\mu_2 = 0.6$
5	0.4941	0.4995	0.5055	0.5111	0.5144
6	0.4669	0.4725	0.4744	0.4748	0.4745
7	0.4581	0.4663	0.4703	0.4720	0.4728

of customers increases; hence, the probability of customer I being preempted increases. With the increase of n, the effect of λ for P_1 is more obvious.

In Table 1, taking $n = 3$, P_2 increases with the increase of μ_2. When μ_2 is unchanged, P_2 decreases with the increase of N. It is mainly because the increase of N makes more customer II get service and have less chances to release server. Therefore, the probability of the server being released from customer II decreases.

Then, we discuss the individual optimality for each customer. We assumed that R_1, R_2 represent the reward which per customer I and customer II could obtain when they got service; C_1, C_2 represent the cost which the customer can produce per waiting time in the system. $U_{\mathrm{I}1}$, $U_{\mathrm{I}2}$ represent the individual benefit of customer I and customer II. The individual benefit $U_{\mathrm{I}i}$ is given as follows

$$U_{\mathrm{I}i} = (\mu_i/\lambda_i)R_i - w_i C_i, \quad i = 1, 2$$

where $\lambda_1 = \lambda\bar{\alpha}$, $\lambda_2 = \lambda\alpha$, $w_i = E(L_i)/\lambda_i$.

Finally, we discuss the social optimality of the system. And we assume that R_0 represents the average reward when the customer had been served; C represents the average cost which customers can produce per waiting time; C_p represents the cost when a customer I had been preempted. The social benefit U_s of the system is given as follows

$$U_s = \lambda_s[(\mu_s/\lambda_s)R_0 - w_s C] - P_1 C_p$$

where $R_0 = (R_1 + R_2)/2$, $C = (C_1 + C_2)/2$, $w_s = (w_1 + w_2)/2$, $\mu_s = (\mu_1 + \mu_2)/2$.

This social optimality arrival rate λ_s^* with the maximum social welfare U_s^* can be denoted as follows

$$\lambda_s^* = \underset{0<\lambda_s<1}{\arg\max} \{\lambda_s[(\mu_s/\lambda_s)R_0 - w_s C] - P_1 C_p\}.$$

In Figure 3, taking $N = 6$, $R_1 = 15$, $C_1 = 1.2$, when n is unchanged, $U_{\mathrm{I}1}$ increases with the increase of λ firstly, then with the increase of λ, $U_{\mathrm{I}1}$ decreases. When λ is unchanged, with the increase of n, $U_{\mathrm{I}1}$ decreases. And obviously, in the same condition, the bigger of the value μ_2 is, the more benefit we can get. We also can find that when $n = 2$,

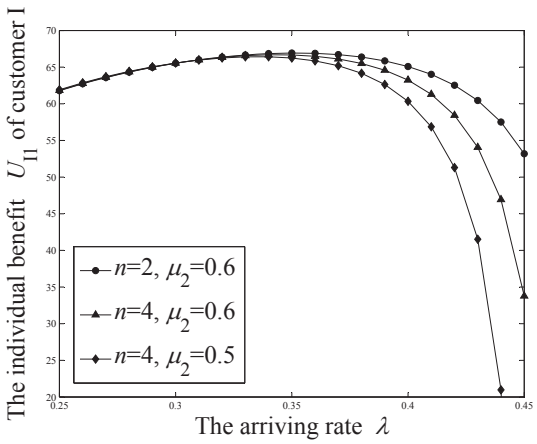

FIGURE 3. The relation of $U_{\mathrm{I}1}$ with λ and n

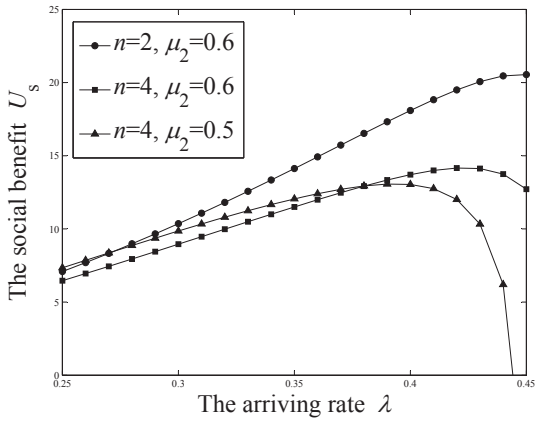

FIGURE 4. The relation of U_s with λ and n

$\mu_2 = 0.6$, the value of $\lambda = 0.36$ can make the individual benefit U_{I1} reach maximum. When $n = 4$, $\mu_2 = 0.5$, the value of $\lambda = 0.35$ can make the individual benefit U_{I1} reach maximum. When $n = 4$, $\mu_2 = 0.6$, the value of $\lambda = 0.34$ can make the individual benefit U_{I1} reach maximum. In Figure 4, taking $N = 8$, when n is unchanged, U_s increases with the increase of λ firstly, then with the increase of λ, the social benefit decreases. When λ is unchanged, with the increase of n, the social benefit decreases. This is mainly because with the increase of n, the frequency of server transformation between two types of customers increases; hence, the cost will increase and the benefit will decrease. We also can find that when $n = 2$, $\mu_2 = 0.6$, the value of $\lambda = 0.44$ is the best parameter, when $n = 4$, $\mu_2 = 0.6$, the value of $\lambda = 0.42$ is the best parameter, and when $n = 4$, $\mu_2 = 0.5$, the value of $\lambda = 0.39$ is the best parameter.

6. **Conclusions.** In this paper, the discrete time Geom/Geom/1 queue with (N, n)-preemptive priority discipline and multiple working vacation was studied. By the analysis of the state transition, a discrete time three-dimension MC was built, the stationary distribution of the queue length was derived by matrix-geometric solution method. We obtained some numerical examples to illustrate the effect of the parameters on system measures. By building a utility function, we could reasonably set parameters for the system and obtain the best parameters which could maximize the individual and social benefits. For the future, this model can be used in cognitive radio networks, peer-to-peer network and communication system.

Acknowledgment. This work was supported in part by the National Natural Science Foundation of China (No. 61472342), Natural Science Foundation of Hebei Province (No. A2014203096), Foundation for Young Teacher of Yanshan University (No. 13LGA017), and Youth Foundation of Higher Education Science and Technology Research of Hebei Province (No. QN2016016).

REFERENCES

[1] Y. Guo and Y. Liu, A law of iterated logarithm for multiclass queues with preemptive priority service discipline, *Queueing System*, vol.79, no.3, pp.251-291, 2015.

[2] E. A. Peköz, Optimal policies for multi-server non-preemptive priority queues, *Queueing Systems*, vol.42, no.1, pp.91-101, 2002.

[3] K. Kim, (N, n)-Preemptive priority queues, *Performance Evaluation*, vol.68, no.7, pp.575-585, 2011.

[4] L. Servi and S. Finn, M/M/1 queues with working vacations (M/M/1/WV), *Performance Evaluation*, vol.50, no.1, pp.41-52, 2002.

[5] G. Zhao and R. Tian, Equilibrium strategy for M/M/1 queueing system with delayed working vacation, *ICIC Express Letters*, vol.9, no.10, pp.2843-2849, 2015.

[6] V. Goswami and G. Mund, Analysis of a discrete-time GI/Geo/1/N queue with multiple working vacations, *Journal of Systems Science and Systems Engineering*, vol.19, no.3, pp.367-384, 2010.

[7] J. Li, N. Tian and W. Liu, Discrete-time GI/Geo/1 queue with multiple working vacations, *Queueing Systems*, vol.56, no.1, pp.53-63, 2007.

[8] W. Sun and S. Li, Equilibrium and optimal behavior of customers in Markovian queues with multiple working vacations, *Top*, vol.22, no.2, pp.694-715, 2014.

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