PERFORMANCE ANALYSIS AND OPTIMIZATION OF A PSEUDO-FAULT GEO/GE/O/1 REPAIRABLE QUEUEING SYSTEM WITH \textit{N}-POLICY, SETUP TIME AND MULTIPLE WORKING VACATIONS

ZHANYOU MA
Department of Applied Mathematics, College of Science
Yanshan University, Qinhuangdao 066004, China

PENGCHENG WANG
School of Transportation Science and Engineering
Beihang University, Beijing 100191, China

WUYI YUE
Department of Intelligence and Informatics
Konan University, Kobe 658-8501, Japan

Abstract. In this paper, we consider a discrete time Geo/Geo/1 repairable queueing system with a pseudo-fault, setup time, \textit{N}-policy and multiple working vacations. We assume that the service interruption is caused by pseudo-fault or breakdown, and occurs only when the server is busy. If the pseudo-fault occurs, the server will enter into a vacation period instead of a busy period. At a breakdown instant, the repair period starts immediately and after repaired the server is assumed to be as good as new. Using a quasi birth-and-death chain, we establish a two-dimensional Markov chain. We obtain the distribution of the steady-state queue length by using a matrix-geometric solution method. Moreover, we analyze the considered queueing system and provide several performance indices of the system in steady-state. According to the queueing system, we first investigate the individual and social optimal behaviors of the customer. Then we propose a pricing policy to optimize the system socially, and study the Nash equilibrium and social optimization of the proposed strategy to determine the optimal expected parameters of the system. Finally, we present some numerical results to illustrate the effect of several parameters on the systems.

1. Introduction. The theory of discrete time queueing system has been well investigated and applied in a variety of directions, such as computer systems, computer communications and manufacturing systems. Domestic and foreign scholars are interested in the discrete time queueing system which was first proposed in \[23\]. Bruneel analyzed a discrete-time queueing system with a single server and obtained the probability generating functions (PGF) of the unaccomplished service in \[7\]. Ndreca et al. considered a GI/Geo/1 queueing system with priority and derived the distributions of the mean queue length and the mean waiting time in \[21\]. A
discrete-time queueing system with batch service and multiple working vacations was studied in [11]. During the past decade, the research of discrete time queueing system has been well studied in [3, 2, 11, 28, 29, 34].

The working vacation policy was introduced into the queueing system firstly in [27]. Subsequently, the working vacation policy has been extended to discrete time queueing systems in [6]. Laxmi et al. analyzed queueing system with multiple working vacations, and obtained the steady state queue length distributions by using recursive technique in [16, 17]. Gao et al. considered a discrete-time queue randomized working vacations negative customer, and found the optimum value of the normal service rate by using the parabolic method in [10]. Moreover, discrete time GI/Geo/1, GI/M/1 and Geo/Geo/1 queueing systems with working vacations were considered respectively in [19, 18]. Furthermore, N-policy and setup time were also studied in the considered system. Yadin first introduced N-policy into the queueing system in [33]. Heyman extended N-policy to M/G/1 queueing system in [13]. Zhang et al. studied a GI/M/1 queueing system with a threshold in [35]. Tian et al. analyzed a queueing system with a removable server and N-policy, and derived the equilibrium mixed strategies and the socially optimal strategies in [30]. Using an embedded Markov chain and a trial solution approach, Lim et al. analyzed a GI/Geo/1 queueing system with N-policy and derived the stationary queue length distribution at the arrival instant in [20]. Choudhury presented a bath arrival M/M/1 queueing system with setup time in [8].

Actually, mechanical failure and service interruption often occur while making a product or serving a customer. In order to analyze and simulate the above practical problems better, many scholars studied some queueing system with server breakdown. Failure policy was first introduced into the queueing system in [5]. Subsequently, Kulkarni et al. extended the queueing system with failure in [15]. Towsley et al. studied an M/M/1 queueing system with failure and queue flushing in [31]. Aissani et al. presented an M/G/1 queueing system with server retrial and breakdown in [1], and obtained simplified expressions for the partial generating functions of the server state and the number of customers in the retrial group. Kalidass et al. analyzed an M/M/1 queueing system with an unreliable server and derived the steady-state distribution and analytical expressions for the transient state probabilities of the system length in [14]. Using the supplementary variable method, Wei et al. discussed a Geo/G/1 queueing system with a second optional service and failure and obtained the probability generating function of the mean queue length in [32].

However, the practical problems involved with the Internet, digital communications and industrial manufacturing are usually very complex. In this paper, we introduce the concept of a pseudo-fault, which is totally different from a breakdown or failure, into the considered queueing systems considered in the research of [9, 21, 26]. Self-protection mechanism is a policy adopted by the system, in order to reduce and avoid the damage when the server runs too long or works at a high intensity. On the other hand, the pseudo-fault includes all the cases caused the service to interrupt, except that the server is really breakdown. For instance, when a server runs too long, the server may want a short break to self-protect. Or when a server has other urgent tasks to process, the server may need a short break to deal with the urgent thing. For these, we consider that when the service is interrupted, there will be two possible reasons that are the breakdown and the pseudo-fault.
Moreover, we assume that the server will return to the vacation period rather than the breakdown period when the pseudo-fault appears. When the breakdown resulting from the end of life or server’s improper operation occurs, the server has to stop work and waits to be repaired by the repairman and then the repair period starts immediately. Therefore, in this paper, we present a discrete time Geo/Geo/1 repairable queueing system with setup time, N-policy, pseudo failures and multiple working vacations based on [22]. We also obtain the distribution of the steady-state queue length and provide several performance measures. Finally, we investigate the individual and social optimal behaviors of the customer.

The rest of this paper is organized as follows. In Section 2, a relevant mathematical model is established according to the quasi birth-and-death chain. In Section 3, the existence condition of the stationary system is presented, and the distribution of the steady-state queue length is obtained. In Section 4, several main performance indices are derived and the system reliability is analyzed. In Section 5, some numerical results relating to the effect of several parameters on the system measures are provided. In Section 6, the individual and social optimal behaviors of the customer are studied. Concluding remarks are given in Section 7.

2. The mathematical model. In this model, let $\bar{x}$ be $1 - x$, for $\forall x \in [0, 1]$. The discrete time Geo/Geo/1 repairable queueing system with pseudo-fault, setup time, N-policy and multiple working vacations is described as follows:

1. A customer arrives at the system in $(n^-, n)$, $n = 0, 1, \ldots$. The inter-arrival time is an independently identically distributed (i.i.d.) sequence, which follows a geometric distribution with parameter $p$ ($0 < p < 1$), namely,
   
   \[ P\{T = j\} = p(\bar{p})^{j-1}, \quad j = 1, 2, \ldots \]

2. The starting and ending of service occur at epoch $n$, $n = 0, 1, \ldots$. The service time $S_b$ is an i.i.d. sequence, which follows a geometric distribution with parameter $\mu_b$ ($0 < \mu_b < 1$) in a regular busy period, namely,
   
   \[ P\{S_b = j\} = \mu_b(\bar{\mu}_b)^{j-1}, \quad j = 1, 2, \ldots \]

   The service time $S_v$ is an i.i.d. sequence, which follows a geometric distribution with parameter $\mu_v$ ($0 \leq \mu_v < 1$, $\mu_v < \mu_b$) in a working vacation period, namely,
   
   \[ P\{S_v = j\} = \mu_v(\bar{\mu}_v)^{j-1}, \quad j = 1, 2, \ldots \]

3. The starting and ending of the vacation occur in $(n, n^+)$, $n = 0, 1, \ldots$. The vacation time $V$ is an i.i.d. sequence, which follows a geometric distribution with parameter $\theta$ ($0 < \theta < 1$), namely,
   
   \[ P\{V = j\} = \theta(\bar{\theta})^{j-1}, \quad j = 1, 2, \ldots \]

4. The beginning and ending of setup time occur in $(n, n^+)$, $n = 0, 1, \ldots$. The setup time $U$ is an i.i.d. sequence, which follows a geometric distribution with parameter $\beta$ ($0 < \beta < 1$) by
   
   \[ P\{U = j\} = \beta(\bar{\beta})^{j-1}, \quad j = 1, 2, \ldots \]

5. When the system becomes empty, the server will start a vacation. If a customer arrives in the system during a vacation period, it will be served with the service rate $\mu_v$. When a working vacation ends, if the number of the customers is less than $N$ (a fixed positive integer, we call this N-policy) in the system, the system will enter into a new working vacation period; otherwise, the system will start a setup period. We assume that the server can not serve customers in
a setup period. When the setup period ends, the server serves customers with the service rate $\mu_b$. During a regular busy period, if service interruptions do not appear in the queue, the server will still work with service rate $\mu_b$ continuously; if the service is interrupted, the queueing system will follow the specific principle as (6) below.

(6) Service interruption occurs only in a regular busy period with probability $q$ ($0 \leq q \leq 1$) in $(n^-, n)$, $n = 0, 1, \ldots$. Subsequently, breakdown and pseudo-fault occur with probabilities $\alpha$ ($0 \leq \alpha \leq 1$) and $\bar{\alpha}$ respectively. If the pseudo-fault occurs, a vacation period is taken. If the breakdown appears, the repair period starts immediately. The repair time $R$ is an i.i.d. random variable and follows a geometric distribution with parameter $\gamma$ $(0 < \gamma < 1)$, i.e. $P\{R = j\} = \gamma \gamma^{j-1}$, $j = 1, 2, \ldots$. Suppose that the repair period begins in $(n^-, n)$ and the repair period ends in $(n, n^+)$. After repair period, the server is assumed to be as good as new one and continues to serve customers, and the served time is still valid.

(7) We assume that inter-arrival times, the probability of service interruptions, repair times, service times, setup times and vacation times are mutually independent. The service discipline is first in first out (FIFO). The queueing system which we considered is illustrated in Figure 1.

\begin{figure}[h]
\centering
\includegraphics{model.png}
\caption{The schematic diagram for the model description.}
\end{figure}

3. Model analysis.

3.1. State transition probability matrix. Let $L_n^+$ be the number of customers in the system at time $n^+$ and $J_n$ be the server state at time $n^+$. From the schematic diagram of the model shown in Figure 1, we assume that a customer who has been served completely or has left in $(n, n^+)$ will not reckon in $L_n^+$.

\[ J_n = \begin{cases} 
0, & \text{the system is in a working vacation and pseudo-fault period at time } n^+ \\
1, & \text{the system is in a setup period at time } n^+ \\
2, & \text{the system is in a regular busy period at time } n^+ \\
3, & \text{the system is in a breakdown period at time } n^+. 
\end{cases} \]

$\{(L_n^+, J_n), n \geq 0\}$ is a discrete time two-dimension Markov Chain (MC) in this system and its state space is given by

\[ \Omega = \{(0, 0)\} \cup \{(i, 1), i \geq N\} \cup \{(i, j), i \geq 1, j = 0, 2, 3\}. \]
Using the lexicographical sequence for the states, the one-step state transition probability matrix of MC can be written as a block tri-diagonal structure as follows:

\[
P = N - 2 \\
N - 1 \\
N \\
N + 1 \\
\vdots \\
\begin{pmatrix}
0 & A_{00} & C_{01} & B_{10} & A_0 & C_0 & 0 & \cdots & \cdots \\
1 & B_0 & A_0 & C_0 & 0 & \cdots & \cdots & \cdots \\
2 & B_0 & A_0 & C_0 & 0 & \cdots & \cdots & \cdots \\
& \vdots & & & \ddots & \ddots & \ddots & \ddots \\
N - 1 & B_0 & A_0 & C_0 & 0 & \cdots & \cdots & \cdots \\
N & B_0 & A_0 & C_0 & 0 & \cdots & \cdots & \cdots \\
N + 1 & B_0 & A_0 & C_0 & 0 & \cdots & \cdots & \cdots \\
& \vdots & & & \ddots & \ddots & \ddots & \ddots \\
\end{pmatrix}
\]  
(1)

where \( A_{00} = \bar{p} \), \( C_{01} = (p, 0, 0) \), \( B_{10} = (\bar{p} \mu_v, \bar{p} \mu_b, 0)^T \), \( T \) denotes matrix transpose, and

\[
A_0 = \begin{pmatrix}
\bar{p} \mu_v + p \mu_v & 0 & 0 & \cdots & \cdots \\
0 & \bar{p} \gamma & \bar{p} \gamma & \cdots & \cdots \\
(\bar{p} \mu_b + p \mu_b) q \alpha & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
(\bar{p} \mu_b + p \mu_b) q \alpha & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
\end{pmatrix}
\]

\[
B_0 = \begin{pmatrix}
\bar{p} \mu_v & 0 & 0 & \cdots & \cdots \\
0 & \bar{p} \gamma & \bar{p} \gamma & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
\end{pmatrix}
\]

\[
C_0 = \begin{pmatrix}
\bar{p} \mu_v & 0 & 0 & \cdots & \cdots \\
0 & \bar{p} \gamma & \bar{p} \gamma & \cdots & \cdots \\
(\bar{p} \mu_b + p \mu_b) q \alpha & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
(\bar{p} \mu_b + p \mu_b) q \alpha & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
\end{pmatrix}
\]

\[
A_1 = \begin{pmatrix}
(\bar{p} \mu_v + p \mu_v) \gamma & 0 & 0 & \cdots & \cdots \\
0 & \bar{p} \beta & \bar{p} \beta & \cdots & \cdots \\
(\bar{p} \mu_b + p \mu_b) q \alpha & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
(\bar{p} \mu_b + p \mu_b) q \alpha & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
\end{pmatrix}
\]

\[
B_1 = \begin{pmatrix}
\bar{p} \mu_v & 0 & 0 & \cdots & \cdots \\
0 & \bar{p} \gamma & \bar{p} \gamma & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
0 & 0 & 0 & \cdots & \cdots \\
\end{pmatrix}
\]

\[
C_1 = \begin{pmatrix}
\bar{p} \mu_v \theta & \bar{p} \mu_v \theta & 0 & 0 & \cdots & \cdots \\
\bar{p} \mu_v \theta & \bar{p} \mu_v \theta & 0 & 0 & \cdots & \cdots \\
0 & 0 & \bar{p} \beta & \bar{p} \beta & \cdots & \cdots \\
0 & 0 & \bar{p} \beta & \bar{p} \beta & \cdots & \cdots \\
0 & 0 & \bar{p} \beta & \bar{p} \beta & \cdots & \cdots \\
0 & 0 & \bar{p} \beta & \bar{p} \beta & \cdots & \cdots \\
0 & 0 & \bar{p} \beta & \bar{p} \beta & \cdots & \cdots \\
0 & 0 & \bar{p} \beta & \bar{p} \beta & \cdots & \cdots \\
0 & 0 & \bar{p} \beta & \bar{p} \beta & \cdots & \cdots \\
0 & 0 & \bar{p} \beta & \bar{p} \beta & \cdots & \cdots \\
\end{pmatrix}
\]

3.2. The steady-state analysis. If MC \( \{(L_n^+, J_n), \ n \geq 0\} \) is positive recurrent, let \( (L, J) \) be the limit of the stationary distribution of \( (L_n^+, J_n) \) and its distribution is given by

\[
\Pi = (\pi_0, \pi_1, \pi_2, \ldots),
\]

\[
\pi_0 = \pi_{00},
\]

\[
\pi_i = (\pi_{i0}, \pi_{i1}, \pi_{i2}, \pi_{i3}), \quad 1 \leq i < N,
\]

\[
\pi_i = (\pi_{i0}, \pi_{i1}, \pi_{i2}, \pi_{i3}), \quad i \geq N,
\]

\[
\pi_{ij} = P\{L = i, J = j\} = \lim_{n \to \infty} P\{L_n^+ = i, J_n = j\}, \quad (i, j) \in \Omega.
\]

The necessary and sufficient condition for the MC \( \{(L_n^+, J_n), \ n \geq 0\} \) to be positive recurrent is that the matrix quadratic equation

\[
R^2 B + RA + C = R
\]

(2)
has a minimal non-negative solution $R$ and the spectral radius $SP(R) < 1$, and the 
$(3N + 2)$ dimensional stochastic matrix

$$B[R] = \begin{pmatrix}
0 & A_0 & C_0 \\
1 & B_{10} & A_0 & C_0 \\
2 & \vdots & \vdots & \vdots \\
N - 1 & B_0 & A_0 & C_1 \\
N & B_1 & A & RB
\end{pmatrix} \quad \text{(3)}$$

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

$$\begin{cases}
\pi_i = \pi_N R_i^{i-N}, & i \geq N + 1 \\
(\pi_0, \pi_1, ..., \pi_N) &= (\pi_0, \pi_1, ..., \pi_N) B[R] \\
\pi_0 + \sum_{k=1}^{N-1} \pi_k e_1 + \pi_N (I - R)^{-1} e_2 = 1
\end{cases} \quad \text{(4)}$$

where $e_1 = (1, 1, 1)^T$, $e_2 = (1, 1, 1, 1)^T$.

The proof of Eq. (4) can be obtained by using equilibrium equation $\Pi P = \Pi$ and matrix-geometric solution method presented in [25], and the main step is given as follows. Firstly, according to the structure of $P$ and $B[R]$, we can prove that $B[R]$ is a $(3N + 2)$-dimensional stochastic matrix. Secondly, we show that $\pi_i = \pi_N R_i^{i-N}, i \geq N + 1$, and $(\pi_0, \pi_1, ..., \pi_N) B[R] = (\pi_0, \pi_1, ..., \pi_N)$. Finally, according to the normalizing condition $\Pi e = 1$, we can obtain $\pi_0 + \sum_{k=1}^{N-1} \pi_k e_1 + \pi_N (I - R)^{-1} e_2 = 1$.

### 3.3. Approximation of algorithm of rate matrix $R$

In this paper, since we use a matrix-geometric solution to express the stationary probability vectors, we have to calculate the minimal non-negative solution $R$ of Eq. (2).

In fact, only when matrices $A$, $B$, $C$ have the special structural forms of an upper triangular matrix and a lower triangular matrix or their structures are relatively simple, we can obtain the analytical expression of $R$. However, in this paper, matrices $A$, $B$, $C$ have general matrix form and their expressions are relatively complex. Therefore, getting an analytic expression of the rate matrix $R$ directly becomes more difficult. Therefore, we usually need to derive the recursion expression of the rate matrix $R$ and calculate the numerical solution by using Matlab program. From Eq. (2), the recursion expression is derived as follows

$$R_{n+1} = (R_n^2 B + C)(I - A)^{-1}, \quad n = 0, 1, ..., \quad \text{(5)}$$

Recursive principle is described as follows

1. We present an error precision $\varepsilon$ whose value is set as greater than 0 but close to 0 infinitely.
2. We provide an initial value $R$ and let $R_0 = 0$ generally.
3. We calculate the expression of $R_{n+1}$ by Eq. (3).
4. Let $\|R\|_{\text{max}} = \|R_{n+1} - R_n\|_{\text{max}} = \max \{ |R_{ij}| \}$, if the largest of the absolute values of all the elements in $R^*$ is not less than $\varepsilon$, the third step is repeated; whereas, iteration is completed, and the value of $R_{n+1}$ is outputted.

So far, the numerical solution of rate matrix $R$ is the output of $R_{n+1}$. 
4. **Performance measures.** In this section, we obtain a serial of main performance measures of the considered queueing system. Specifically, when \( N = 1 \), the considered system degenerates into the discrete time Geo/Geo/1 repairable queueing system with pseudo-fault, setup time and multiple working vacations. When the setup time is taken as zero, the considered system degenerates into the discrete time Geo/Geo/1 repairable queueing system with pseudo-fault, \( N \)-policy and multiple working vacations. Moreover, when the setup time is set as zero and \( N = 1 \), the considered system degenerates into a Geo/Geo/1 repairable queueing system with pseudo-fault and multiple working vacations.

4.1. **Queue indicators.** The expected queue length is given by

\[
E[L] = \sum_{i=0}^{\infty} \sum_{j=0}^{3} i \pi_{ij} = \sum_{i=0}^{N-1} i \pi_{i} e_1 + \sum_{i=N}^{\infty} \sum_{j=0}^{3} i \pi_{i} e_2.
\]

The expected waiting queue length is given by

\[
E[L_q] = \sum_{i=1}^{\infty} \sum_{j=0}^{3} (i-1) \pi_{ij} = \sum_{i=0}^{N-1} (i-1) \pi_{i} e_1 + \sum_{i=N}^{\infty} (i-1) \pi_{i} e_2 = E[L] + \pi_{00} - 1.
\]

The expected waiting time is given by

\[
E[W] = \frac{1}{p} E[L_q] = \frac{1}{p} \sum_{i=1}^{\infty} \sum_{j=0}^{3} (i-1) \pi_{ij} = \frac{E[L] + \pi_{00} - 1}{p}.
\]

The state probability that the system is in a setup period is given by

\[
P_U = P \{ L \geq N, j = 1 \} = \lim_{n \to \infty} P \{ J_n = 1 \} = \sum_{i=N}^{\infty} \pi_{i1}.
\]

The state probability that the system is in a pseudo-fault period is given by

\[
P_{q2} = \bar{\alpha} P_{q1} / \alpha = (1/\alpha - 1) \sum_{i=1}^{\infty} \pi_{i3}.
\]

The state probability that the system is in a regular busy period is given by

\[
P_B = P \{ L \geq 1, j = 2 \} = \lim_{n \to \infty} P \{ J_n = 2 \} = \sum_{i=1}^{\infty} \pi_{i2}.
\]

The state probability that the system is in a working vacation and pseudo-fault period is given by

\[
P_0 = P \{ L \geq 0, j = 0 \} = \lim_{n \to \infty} P \{ J_n = 0 \} = \sum_{i=0}^{\infty} \pi_{i0}.
\]

4.2. **Fault analysis.** During the busy period, service interruptions which resulted from breakdowns and pseudo-faults may occur with a certain possibility. But how to take the appropriate measures to deal with service interruptions effectively becomes particularly important. In this subsection, we provided the probability of a pseudo-fault period and breakdown period in the steady-state.

The state probability that the system is in a breakdown period is given by

\[
P_{q1} = P \{ L \geq 1, j = 3 \} = \lim_{n \to \infty} P \{ J_n = 3 \} = \sum_{i=1}^{\infty} \pi_{i3}.
\]

The state probability that the system is in a pseudo-fault period is given by

\[
P_{q2} = \bar{\alpha} P_{q1} / \alpha = (1/\alpha - 1) \sum_{i=1}^{\infty} \pi_{i3}.
\]
The state probability that the system is in a working vacation period is given by

\[ P_V = P_0 - P_{q2} = \sum_{i=0}^{\infty} \pi_{i0} - \frac{1}{\alpha - 1} \sum_{i=1}^{\infty} \pi_{i3}. \] (14)

5. Numerical illustration. Based on the above analysis, we obtained several expressions for the main performance indices. In this section, we provide some numerical results to describe the effect of system parameters on performance measures. Specifically, Figures 2-5 show the curves of the different parameters and system indices by taking \( q = 0.4, \gamma = 0.6, \theta = 0.8 \). When \( p = 0.3, \mu_v = 0.2, \alpha = 0.7, \beta = 0.7, N = 15 \), Tables 1-2 illustrate the change relation of performance indices with the relevant parameters.

Figure 2 describes the influence of the parameters \( \mu_v \) and \( N \) on the probability \( P_B \) by taking \( p = 0.3, \mu_b = \alpha = \beta = 0.7 \). When \( N \) is fixed, \( P_B \) decreases with the increase of \( \mu_v \). This is mainly because \( \mu_v \) increases, the departing customers increase the vacation period, and the probability that the system is in a busy period becomes small. If \( \mu_v \) is fixed, \( P_B \) decreases with the increasing value of \( N \). This is mainly because the larger \( N \) becomes, the slower the chance that the system changes into a regular busy period. Hence, the probability that the system is in a busy period becomes small.

When \( \mu_v = 0.2, \mu_b = \alpha = \beta = 0.7 \), Figure 3 shows the relation of \( E[L] \) with \( p \) and \( N \). When \( p \) is fixed, \( E[L] \) increases with the increasing value of \( N \). When the value of \( p \) becomes large, then \( E[L] \) increases. This is mainly because the larger \( p \) is, the more the arrival customers are. Hence, the queue length also becomes larger.

Figure 4 illustrates the effect of the waiting queueing length \( E[L_q] \) with \( \mu_b \) and \( \beta \) when \( p = 0.3, \mu_v = 0.1, \alpha = 0.7, N = 15 \). If \( \beta \) is a constant, \( E[L_q] \) decreases with the increasing value of \( \mu_b \). This is mainly because the service time reduces with an increase in the service rate \( \mu_b \), and the mean waiting time \( E[L_q] \) decreases. When \( \mu_b \) is unchanged, \( E[L_q] \) decreases with the increase of \( \beta \). That is mainly because the larger the probability \( \beta \) is, the longer the busy period is. Hence, \( E[L_q] \) decreases.
The expected queue length $E[L]$ to $p$ and $N$.

The expected waiting queue length $E[L_q]$ to $\mu_b$ and $\beta$.

The relation of $P_{q2}$ to $p$ and $\alpha$. 
Figure 5 shows the change relation of the probability $P_{q_2}$ with the arrival rate $p$ and $\alpha$ when $\mu_v = 0.2$, $\mu_b = \beta = 0.7$, $N = 15$. When $\alpha$ is fixed, $P_{q_2}$ increases with the increasing value of $p$. This is mainly because of the probability that the system is in a regular busy period increasing along with an increase of the arrival rate $p$, therefore the probability $P_{q_2}$ increases. When $p$ is fixed, $P_{q_2}$ decreases with the increasing value of $\alpha$. From the analysis of the results in Figure 5, it is believed that the pseudo-fault phenomenon really exists and cannot be ignored, and it is affected by the size of the relevant parameters.

### Table 1. The relation of $E[W]$ to $q$ and $\theta$.

| $\theta$ | $q = 0$ | $q = 0.05$ | $q = 0.1$ | $q = 0.15$ | $q = 0.2$ | $q = 0.25$ | $q = 0.3$ |
|----------|----------|----------|----------|----------|----------|----------|----------|
| 0.3      | 20.207   | 22.379   | 24.622   | 26.885   | 29.148   | 31.430   | 33.798   |
| 0.5      | 19.942   | 22.019   | 24.140   | 26.250   | 28.318   | 30.351   | 32.396   |
| 0.7      | 19.843   | 21.884   | 23.958   | 26.009   | 28.006   | 29.947   | 31.875   |

In Table 1 when $\mu_b = 0.7$, $\gamma = 0.6$, we provide the change relation of the mean waiting time $E[W]$ with $q$ and $\theta$. If $\theta$ is fixed, $E[W]$ increases with the increasing value of $q$. That is mainly because the higher the probability $q$ is, the more waiting customers are, therefore $E[W]$ will increase. When $q$ is a constant, $E[W]$ decreases with the increase of $\theta$. This is mainly because the probability that the system is in a regular busy period increases with the increasing value of $\theta$, therefore $E[W]$ decreases. Especially, when $q = 0$, it indicates that a service interruption does not occur during the busy period.

### Table 2. The relation of $E[L]$ to $\mu_b$ and $\gamma$.

| $\gamma$ | $\mu_b = 0.6$ | $\mu_b = 0.65$ | $\mu_b = 0.7$ | $\mu_b = 0.75$ | $\mu_b = 0.8$ | $\mu_b = 0.85$ | $\mu_b = 0.9$ |
|----------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 0.4      | 32.139       | 17.645       | 14.184       | 12.449       | 11.387       | 10.673       | 10.165       |
| 0.6      | 15.069       | 12.820       | 11.560       | 10.757       | 10.207       | 9.809        | 9.508        |
| 0.8      | 13.100       | 11.697       | 10.830       | 10.250       | 9.838        | 9.529        | 9.288        |

When $q = 0.4$, $\theta = 0.8$, Table 2 presents the influence of $\mu_b$ and $\gamma$ on the mean queue length $E[L]$. We can see that $E[L]$ decreases with the increase of $\mu_b$, when $\gamma$ is a constant. If $\mu_b$ is fixed, $E[L]$ decreases with the increasing value of the repair rate $\gamma$. This is mainly because the higher $\gamma$ is, the faster the stationary chance that the system turns from a breakdown period to a busy period. Hence, $E[L]$ becomes small.

### 6. Optimization and pricing policy.

Considering the discrete time queueing system proposed in this paper, we first investigate the individual and social optimal behaviors of the customer, then we propose a pricing policy to optimize the system socially. In order to study the Nash equilibrium and social optimization of the proposed strategy, we give some assumptions as follows:

1. Customers can adjust their arrival rate according to the system feedback information. In this section, we consider the arrival rate of customer ranging from $p_1$ to $p_2$.
2. The cost per unit waiting time is $C_W$ for a customer stayed in the system.
3. The reward of a customer served is $R$. 
6.1. Individual optimal and social optimal behavior. We note that every customer is individually selfish and tries to receive service to get a reward. We define the individual benefit function $U_I$ for a customer as follows

$$U_I = R - C_W E[W]$$  \hspace{1cm} (15)$$

where $E[W]$ is the mean waiting time of a customer given in Eq. (8).

When the arrival rate $p$ of customers increases, the mean waiting time $E[W]$ of a customer will increase monotonically. Therefore, the individual benefit for a customer is a decreasing function about the arrival rate $p$. Because all customers are personally selfish, they all try their best to get service. Provided the benefit is positive, the arrival rate of customers will be as high as possible. If there is at least one solution for the inequality $U_I \geq 0$ within the closed interval $[p_1, p_2]$, the maximal value of the solutions is the individual optimal arrival rate $p^e$. Otherwise, $p^e = p_1$.

With the change of the individual optimal arrival rate $p^e$, the system will reach a Nash equilibrium state. No customer has any incentive to deviate unilaterally from the individual optimal arrival rate. We discuss the Nash equilibrium and social optimization of the proposed strategy as follows.

(1) For any $p \in [p_1, p_2]$, if $R < C_W E[W]$, the individual benefit function $U_I$ for a customer is less than zero. Namely, even if all the customers arrive at the system with the lowest arrival rate, the individual benefit for a customer is negative. Therefore, $p^e = p_1$ is an equilibrium arrival rate and no other equilibrium arrival rate exists. Accordingly, the dominant strategy is that customers arrive at the lowest level.

(2) For any $p \in [p_1, p_2]$, if $R \geq C_W E[W]$, the individual benefit function $U_I$ for a customer is no less than zero. Namely, even if all the customers arrive at the system with the maximum arrival rate, the individual benefit for a customer is non-negative, they can get non-negative individual benefit. Therefore, $p^e = p_2$ is an equilibrium arrival rate and no other equilibrium arrival rate exists. Accordingly, the dominant strategy is that customers arrive at the highest level.

(3) On the closed interval $[p_1, p_2]$, if there are two kinds of inequality $R > C_W E[W]$ and $R < C_W E[W]$ at the same time, there exists a unique equilibrium strategy $p_1 < p^e < p_2$ which satisfies $R = C_W E[W]$. Therefore, $p^e$ is a mixed equilibrium arrival rate.

We investigate the monotonicity of the individual benefit $U_I$ for a customer with numerical experiments. Referencing to [12], we set the parameters as follows, $p_1 = 0.1$, $p_2 = 0.4$, $\mu_0 = 0.8$, $\mu_v = 0.25$, $q = 0.2$, $\beta = 0.3$, $\gamma = 0.6$, $R = 250$, $C_W = 25$, and setting $\theta = 0.2, 0.4, 0.8$ respectively, we show the change trend of individual benefit $U_I$ for a customer versus the arrival rate $p$ of customers in Figure 6.

Figure 6 shows that with the parameters set above, all the individual benefits $U_I$ decrease as the arrival rate $p$ of customers increases. We find that all the individual benefits $U_I$ go through $U_I = 0$, i.e., there are always values of $p^e$ subjected to $U_I = 0$. Therefore, an individual optimal behavior for our proposed strategy exists.

System design may need to consider many factors, not only individual benefit, but also the socially optimal behavior of customers. Therefore, we define the social benefit function $U_S$ as follows

$$U_S = p(R - C_W E[W]).$$  \hspace{1cm} (16)$$
By maximizing the social benefit, we can obtain the socially optimal arrival rate \( p^* \). So \( p^* \) is given as follows

\[
p^* = \arg \max_{p \in [p_1, p_2]} U_S.
\]  

(17)

With the same parameters used in Figure 6, we show how the social benefit function \( U_S \) changes with the arrival rate \( p \) of customers in Figure 7.

In Figure 7, we find that there is always a socially optimal arrival rate \( p^* \) and a maximal social benefit \( U_S^* \) for all the vacation parameters \( \theta \). And combining the results given in Figures 6 and 7, we summarize the individually optimal arrival rate \( p^e \) and the socially optimal arrival rate \( p^* \) in Table 3.

From Table 3, we conclude that optimizing the individual benefit leads to a higher arrival rate of customers than that socially desired. This can be addressed by imposing an appropriate admission fee for customers.
Table 3. Comparison of individually and socially optimal arrival rate.

| Vacation parameter $\theta$ | Individually optimal arrival rate $p^e$ | Socially optimal arrival rate $p^*$ |
|-----------------------------|----------------------------------------|--------------------------------------|
| 0.2                         | 0.348                                  | 0.204                                |
| 0.4                         | 0.378                                  | 0.228                                |
| 0.8                         | 0.392                                  | 0.240                                |

6.2. Pricing policy. One approach that oblige the customers to adopt the socially optimal arrival rate is to charge a fee to customers. We assume the management of the system acts as a pricing agent and imposes an admission fee on all customers. Therefore, we present a pricing policy, when an admission fee $f$ is imposed, the individual benefit $U^f_I$ for a customer is given as follows

$$U^f_I = R - C_W E[W] - f.$$ (18)

Substituting the arrival rate $p$ of customers with socially optimal arrival rate $p^*$ of customers given in Table 3 and letting $U^f_I = 0$, we can calculate the admission fee $f$ as follows

$$f = R - C_W E[W].$$ (19)

With the change of the socially optimal arrival rates $p^*$ given in Table 3, we calculate the mean waiting time using Eq. (8). Therefore, we can give the admission fee $f$ using Eq. (19). For different vacation parameters $\theta$, we summarize the socially maximum benefit $U^*_S$ and the admission fee $f$ in Table 4.

Table 4. Numerical results for admission fee.

| Vacation parameter $\theta$ | Socially maximum benefit $U^*_S$ | Admission fee $f$ |
|-----------------------------|---------------------------------|-------------------|
| 0.2                         | 28.726                          | 140.813           |
| 0.4                         | 30.796                          | 135.071           |
| 0.8                         | 33.101                          | 133.754           |

7. Conclusions. In this paper, we presented a system model and analyses to evaluate the performance of a pseudo-fault Geo/Geo/1 repairable queueing system with $N$-policy, setup time and multiple working vacations. By using the quasi birth-and-death chain and matrix-geometric solution method, we established a two-dimensional Markov chain and obtained the distribution of the steady-state queue length. Based on the distribution of the steady-state queue length, we analyzed the expected queue length and waiting queue length of the system, and the expected waiting time of a customer. Furthermore, we obtained the probabilities that the system is in a setup period, a regular busy period, a working vacation period or pseudo-fault period, respectively and performance measures such as the expected queue lengths and waiting time. Then, we provided numerical results to describe the effect of system parameters on performance measures. Finally, we investigated the individual and social optimal behaviors of the customer, and by proposing a pricing policy to optimize the system socially to discuss the Nash equilibrium and social optimization of the proposed strategy to determine the optimal expected parameters of the system.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (No. 61472342), Hebei Province National Science Foundation (No. A2014203096), Youth Foundation of Higher Education Science and Technology Research of Hebei Province (No. QN2016016), and was supported in part by MEXT, Japan.
REFERENCES

[1] A. Aissani and J. Artalejo, On the single server retrial queue subject to breakdowns. *Queueing Systems*, 30 (1998), 309–321.
[2] I. Atencia, I. Fortes, S. Nishimura and S. Sánchez, A discrete-time retrial queueing system with recurrent customers, *Computers & Operations Research*, 37 (2010), 1167–1173.
[3] I. Atencia, I. Fortes and S. Sánchez, A discrete-time retrial queueing system with starting failures, Bernoulli feedback and general retrial times. *Computers & Industrial Engineering*, 57 (2009), 1291–1299.
[4] I. Atencia and A. Pechinkin, A discrete-time queueing system with optional LCFS discipline. *Annals of Operations Research*, 202 (2013), 3–17.
[5] B. Avi-Itzhak and P. Naor, Some queuing problems with the service station subject to breakdown. *Operations Research*, 11 (1963), 303–320.
[6] Y. Baba, Analysis of a GI/M/1 queue with multiple working vacations. *Operations Research Letters*, 33 (2005), 201–209.
[7] H. Bruneel, Performance of discrete-time queueing systems. *Computers & Operations Research*, 20 (1993), 303–320.
[8] G. Choudhury, On a batch arrival Poisson queue with a random setup time and vacation period. *Computers & Operations Research*, 25 (1998), 1013–1026.
[9] C. Cramer, A seismic hazard uncertainty analysis for the New Madrid seismic zone. *Engineering Geology*, 62 (2001), 251–266.
[10] S. Gao and J. Wang, On a discrete-time GI\(X/G/1/N-G\) queue with randomized working vacations and at most J vacations. *Journal of Industrial and Management Optimization*, 11 (2015), 779–806.
[11] V. Goswami and G. Mund, Analysis of discrete-time batch service renewal input queue with multiple working vacations. *Computers & Industrial Engineering*, 61 (2011), 629–636.
[12] R. Hassin and M. Haviv, *Queue or not to Queue: Equilibrium Behavior in Queueing Systems*. Kluwer Academic Publishers, 2003.
[13] D. Heyman, The T-policy for the M/G/1 queue. *Management Science*, 23 (1977), 775–778.
[14] K. Kalidass, J. Gnanaraj, S. Gopinath and R. Kasturi, Transient analysis of an M/M/1 queue with a repairable server and multiple vacations. *International Journal of Mathematics in Operational Research*, 6 (2014), 193–216.
[15] V. Kulkarni and B. Choi, Retrial queues with server subject to breakdowns and repairs. *Queueing Systems*, 7 (1990), 191–208.
[16] P. Laxmi and S. Demie, Performance analysis of renewal input (a,c,b) policy queue with multiple working vacations and change over times. *Journal of Industrial and Management Optimization*, 10 (2014), 839–857.
[17] P. Laxmi, S. Indira and K. Jyothsna, Ant colony optimization for optimum service times in a Bernoulli schedule vacation interruption queue with balking and reneging. *Journal of Industrial and Management Optimization*, 12 (2016), 1199–1214.
[18] J. Li, N. Tian and M. Liu, Discrete-time GI/G/1 queue with multiple working vacations. *Queueing Systems*, 56 (2007), 53–63.
[19] J. Li, N. Tian and Z. Ma, Performance analysis of GI/M/1 queue with working vacations and vacation interruption. *Applied Mathematical Modelling*, 32 (2008), 2715–2730.
[20] D. Lim, D. Lee, W. Yang and K. Chae, Analysis of the GI/G/1 queue with N-policy. *Applied Mathematical Modelling*, 37 (2013), 4643–4652.
[21] Z. Ma, P. Wang, G. Cui and Y. Hao, The discrete time Geom/Geom/1 repairable queueing system with pseudo-fault and multiple vacations, *Journal of Information and Computational Science*, 11 (2014), 4667–4678.
[22] Z. Ma, P. Wang and W. Yue, The pseudo-fault Geo/Geo/1 queue with setup time and multiple working vacation. In *Proceedings of the 10th International Conference on Queueing Theory and Network Applications*, 383 (2015), 105–112.
[23] T. Meisling, Discrete-time queueing theory. *Operations Research*, 6 (1958), 96–105.
[24] S. Ndreca and B. Scoppola, Discrete time GI/Geom/1 queueing system with priority. *European Journal of Operational Research*, 189 (2008), 1403–1408.
[25] M. Neuts, *Matrix-geometric Solution in Stochastic Model: An Algorithmic Application*. The Johns Hopkins University Press, 1981.
[26] E. Papatheou, G. Manson, R. J. Bartheorpe and K. Worden, The use of pseudo-faults for damage location in SHM: An experimental investigation on a Piper Tomahawk aircraft wing, *Journal of Sound and Vibration*, 333 (2014), 971–990.

[27] L. Servi and S. Finn, M/M/1 queues with working vacations (M/M/1/WV), *Performance Evaluation*, 50 (2002), 41–52.

[28] W. Sun, H. Zhang and N. Tian, The discrete-time Geom/G/1 queue with multiple adaptive vacations and server setup/closedown times, *International Journal of Management Science and Engineering Management*, 2 (2007), 289–296.

[29] N. Tian and Z. Zhang, The discrete-time GI/Geo/1 queue with multiple vacations *Queueing Systems*, 40 (2002), 283–294.

[30] R. Tian, D. Yue and W. Yue, Optimal balking strategies in an M/G/1 queueing system with a removable server under N-policy, *Journal of Industrial and Management Optimization*, 11 (2015), 715–731.

[31] D. Towsley and S. Tripathi, A single server priority queue with server failures and queue flushing, *Operations Research Letters*, 10 (1991), 353–362.

[32] C. Wei, Z. Zou and Y. Qin, Discrete time Geom/G/1 queue with second optional service and server breakdowns, *Fuzzy Engineering and Operations Research*, 147 (2012), 557–568.

[33] M. Yadin and P. Naor, Queueing systems with a removable service station, *Operations Research Quarterly*, 14 (1963), 393–405.

[34] Z. Zhang and N. Tian, Discrete time Geo/G/1 queue with multiple adaptive vacations, *Queueing Systems*, 38 (2001), 419–429.

[35] Z. Zhang and N. Tian, The N-threshold policy for the GI/M/1 queue, *Operations Research Letters*, 32 (2004), 77–84.

Received September 2015; 1st revision March 2016; final revision June 2016.

E-mail address: mzhy55@ysu.edu.cn
E-mail address: kdwpc@126.com
E-mail address: yue@konan-u.ac.jp