The M/M/c Retrial Queueing System with Impatient Customers and Server Working Breakdown

Shengli Lv, *Member, IAENG*, Fengqin Li, Jingbo Li

***Abstract*—This paper analyzes the M/M/c queueing system with server working breakdown and impatient customers under the classical truncated retry strategy. The steady-state equilibrium conditions of the system are obtained using the matrix-geometric solution method. Then the steady-state probabilities and the performance indicators of the system in the steady-state case are obtained using the Markov process theory and Gauss-Seidel iterative algorithm. Numerical experiments are conducted to assess the system's steady-state performance indicators. Finally, we developed an optimization model for the system and derived the optimal parameters from the analysis of the optimization function.**

***Index Terms*—retrial queue, server working breakdown, impatient customers, matrix-geometric solution, system optimization.**

I. Introduction

THE retrial queueing model is commonly implemented in computer science, telecommunications, and various other industries. The retrial queueing model pertains to a scenario in which a customer enters the system and all servers are occupied. In such a case, the customer enters a virtual waiting room for retries. At the same time, each customer can independently retry seeking service after a random period. This retrial model holds practical significance and has attracted scholars' attention. Tien [1] analyzed the application of the retrial queueing system in the dynamic host configuration protocol and conducted simulation experiments. Kim and Kim [2] discussed the retrial queueing model under several different assumptions. Ye and Chen [3] analyzed a retrial queueing model involving working breakdowns. They used matrix-geometric solution and the generalized eigenvalue method to solve the problem, and compared the two methods. Kumar [4] analyzed the M/M/c retrial queueing model for impatient customers with a PH distribution retrial rate and conducted numerical experiments on performance measures. Shin [5] studied several different multi-service desk retrial queueing models, analyzed the effect of the retry rate on the model captain, and verified its monotonicity. Recently, Shweta [6] investigated a retrial queueing model with Bernoulli vacation. The problem was solved using a Markov chain approach and various cost optimization algorithms. Han et al. [7] examined the M/M/1 retrial queueing model. They considered the presence of impatient customers and delayed maintenance, which could occur when the machine is idle. Finally, they performed an equilibrium analysis of the system. Zhang and Wang [8] analyzed an M/M/1 unreliable queueing model with a constant retrial rate. They investigated system indicators for visible and invisible captains and provided the payoff function.

In practical applications, servers may fail due to long-term usage. The traditional fault-repairable model has limitations, so working breakdowns are introduced to be more relevant. With a working breakdown, a server can continue working at a lower rate after a fault occurs, rather than stopping. Karthick and Suvitha [9] conducted a study on fault-repairable queuing systems with vacation strategies, while varying service rates. Lv et al. [10] used matrix-geometric solutions to examine a queuing model with two different fault types and negative customers. They analyzed the performance indicators of the system at steady state conditions and came to a reasonable conclusion. Kalidass and Kasturi [11] initially proposed a working breakdown and solved the model using steady-state conditions, probability functions, and numerical analysis. Lv et al. [12] examined a system with a standby server and startup period, which is repairable in case of complete and incomplete failures. Jing et al. [13] examined a retrial queueing system that had a single server, customer stops, and working breakdowns. They utilized matrix-geometric solutions to obtain clear solutions for the steady-state probabilities. Additionally, they conducted numerical experiments to analyze the performance indexes, applying their conclusions to real-world situations. Lv et al. [14] examined the impact of varying input and failure rates on queuing systems. Ye and Liu [15] analyzed the MAP/M/1 queueing system with a working breakdown. They calculated the steady-state probabilities using the matrix-geometric solution method while analyzing the impact of parameters. Yang et al. [16] discussed the M/M/1/N queuing model with working breakdown. They derived the performance index by solving the steady-state probability and conducted a sensitivity analysis. Zhang and Xu [17] conducted a study on an incomplete breakdown queueing system that involves two types of customers with different arrival times. Once a failure occurs, the machine stops accepting new customers and continues to complete the ongoing service at a reduced rate. They also examined the impact of system parameters on customer behavior.

Manuscript received January 16, 2024; revised June 14, 2024.

This work was supported by the National Nature Science Foundation of China (No. 71971189).

Shengli Lv is an associate professor in the School of Science, Yanshan University, Qinhuangdao, Hebei 066004, PR China. (e-mail: qhdddlsl@163.com).

Fengqin Li is a postgraduate student in the School of Science, Yanshan University, Qinhuangdao, Hebei 066004, PR China. (corresponding author to provide phone: +86-17349800863; e-mail: lifq25455@163.com).

Jingbo Li is an associate professor in the School of Hebei Normal University of Science and Technology, Qinhuangdao, Hebei 066004, PR China. (e-mail: lijingbo668@126.com).

Customers waiting in a queue may become impatient after a long wait. Kumar and Sharma [18] analyzed the M/M/c queueing model with infinite capacity. The system included impatient customers who would enter with a certain probability upon arrival and wait for a certain period before leaving if they became impatient. Finally, the researchers obtained the steady-state probability solution of the system by applying the generating function. Kumar and Raja [19] analyzed the M/M/c retrial queueing model with discouragement and feedback. They provided steady-state conditions and probabilities, analyzed parameter effects on performance metrics, and conducted optimization experiments. Liu and Li [20] studied the MAP/PH/1 queueing model with impatient customers. Shan and Yue [21] conducted a study on an inventory queueing system that has impatient customers and two different vacation policies. They used an iterative approach to derive the steady-state probability and compared the two policies. Finally, they conducted numerical experiments to analyze Performance indicators of the system. Chang et al. [22] analyzed the M/M/c retrial queuing model with impatient customers and feedback. They proposed the classical truncated retrial and constant retrial policies. They also derived the steady-state probability using an iterative algorithm and performed numerical analysis under both retrial policies to assess their practical applicability.

With the advent of the Internet, the traditional way of purchasing tickets through a physical ticketing window is gradually being replaced by online ticketing platforms such as Railway 12306, Meituan, and others. However, when there are too many customers using the system to purchase tickets, the system may crash. In such cases, customers may either refresh the page to repurchase or leave the system altogether. Based On the above literature, this paper describes the M/M/c retrial queueing system with impatient customers and working breakdown. we utilize the matrix-geometric solution method to obtain the steady-state probability vector of the system. Afterwards, we solve system-related performance indexes and provide numerical experiments using MATLAB. Finally, the article subsequently constructs cost and benefit functions, to further optimize the breakdown problem and the impatience problem in the retrial system.

II. Model Description

1) This is an M/M/c queueing model. In this model, customer arrivals follows a Poisson process with arrival rate λ.

2) This queueing system incorporates working breakdowns. During the normal working period, the server's service time is distributed exponentially with the parameter μ. The server may break down during the normal working process. The arrival of the breakdown follows a Poisson process with a breakdown rate α. After a breakdown occurs, the server continues to provide service at a lower rate $\eta\,(\eta < \mu)$. The system features c repairmen who immediately begin work after a service desk failure. Each repairman can only focus on one faulty machine at a time, and once they complete their work, the service desk is restored to a new state. The repair time follows an exponential distribution with parameter β. During the repair period, the server remains to serve at a rate η.

3) In this queueing system, the behavior of impatient customers is considered. When the customer arrives, they will be immediately served if there is at least one server available. However, if all servers are busy, the customer has two choices: either wait in a retrial space with a probability of b or leave with a probability of $1-b$.

4) There is a retry space where customers will attempt to request service after a random period. The maximum number of customers allowed to retry is denoted by N. Even though the system itself can handle an infinite number of customers, in reality, there is a limit to the number of customers who can use the system simultaneously. Therefore, we need to apply a truncated retrial strategy. Suppose that, at a certain point in time, there are n customers waiting to retry, and the average number of times that each of them sends a retry request in a unit of time is θ. Then the total number of retry requests in the retry space in a unit of time is $\theta_n = \min\{n, N\}\theta$. If a customer in the retry space sends a retry request and finds that the server is still occupied, they may choose to continue retrying with a probability of r or leave with a probability of $1-r$.

5) Assuming that the customer arrival process, retry process, service process, failure process, and repair process are all independent of each other. The order of service is first-come-first-served (FCFS) and the capacity of the retry space is infinite.

Let $N(t)$ denote the number of customers in the retry space at moment t, $I(t)$ denote the number of servers in working condition at moment t, and $J(t)$ denote the number of servers in working breakdown condition at moment t.

Then $\{N(t), I(t), J(t)\}$ is a three-dimensional Markov process, its state space is

$$\Omega = \{(i,j,k), i \geq 0, c \geq j \geq 0, j \geq k \geq 0\}.$$

The state transition diagram for this system with $c=3$ is shown in Fig. 1.

we can organize the states in dictionary order. This allows us to derive the state transfer rate matrix for the system, expressed as follows:

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{C}_0 & & & & & \\ \boldsymbol{B}_1 & \boldsymbol{A}_1 & \boldsymbol{C} & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & \boldsymbol{B}_{N-1} & \boldsymbol{A}_{N-1} & \boldsymbol{C} & & \\ & & & \boldsymbol{B}_N & \boldsymbol{A}_N & \boldsymbol{C} & \\ & & & & \boldsymbol{B}_N & \boldsymbol{A}_N & \boldsymbol{C} \\ & & & & & \ddots & \ddots & \ddots \end{pmatrix},$$

where

$$\boldsymbol{A}_0 = \begin{pmatrix} \boldsymbol{E}_0 & \boldsymbol{F}_1 & & & & \\ \boldsymbol{D}_1 & \boldsymbol{E}_1 & \boldsymbol{F}_2 & & & \\ & \boldsymbol{D}_2 & \boldsymbol{E}_2 & \boldsymbol{F}_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \ddots \\ & & & \boldsymbol{D}_{c-1} & \boldsymbol{E}_{c-1} & \boldsymbol{F}_c \\ & & & & \boldsymbol{D}_c & \boldsymbol{E}_c \end{pmatrix}.$$

$$\boldsymbol{F}_i(k,k) = \lambda, 1 \leq k \leq i. \tag{1}$$

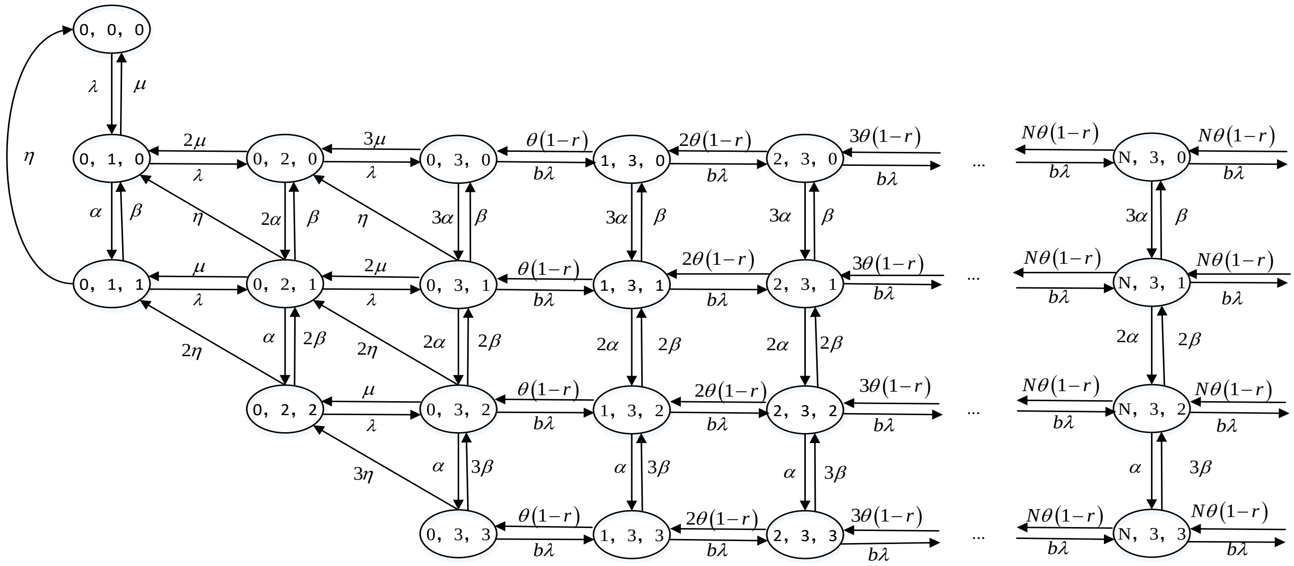

Fig. 1. The state transition diagram of the system

$$\begin{cases} \boldsymbol{D}_i\left(k,k\right)=\left(i-k+1\right)\mu, 1\le k\le i, \\ \boldsymbol{D}_i\left(k+1,k\right)=k\eta, 1\le k\le i. \end{cases} \tag{2}$$

When $1\le i\le c-1$, $1\le k\le i$,

$$\begin{cases} \boldsymbol{E}_i\left(k,k+1\right)=(\mathrm{i}-\mathrm{k}+1)\alpha, \\ \boldsymbol{E}_i\left(k+1,k\right)=k\beta, \\ \boldsymbol{E}_i\left(k,k\right)=-(\mathrm{i}-\mathrm{k}+1)\mu- \\ (\mathrm{i}-\mathrm{k}+1)\alpha-\lambda-\mathrm{k}\eta-\mathrm{k}\beta. \end{cases} \tag{3}$$

When $1\le k\le i$,

$$\begin{cases} \boldsymbol{E}_c\left(k,k+1\right)=(\mathrm{i}-\mathrm{k}+1)\alpha, \\ \boldsymbol{E}_c\left(k+1,k\right)=k\beta, \\ \boldsymbol{E}_c\left(k,k\right)=-(\mathrm{i}-\mathrm{k}+1)\mu \\ -(\mathrm{i}-\mathrm{k}+1)\alpha-\lambda-\mathrm{k}\eta-\mathrm{k}\beta-\mathrm{b}\lambda. \end{cases} \tag{4}$$

$$\begin{cases} \boldsymbol{A}_i\left(k,k+1\right)=(\mathrm{c}+1-\mathrm{k})\alpha, \\ \boldsymbol{A}_i\left(k+1,k\right)=k\beta, 1\le k\le c, \\ \boldsymbol{A}_i\left(k,k\right)=-(\mathrm{k}-1)\beta \\ -\mathrm{b}\lambda-(\mathrm{c}+1-\mathrm{k})\alpha-\mathrm{i}\theta(1-\mathrm{r}), 1\le k\le c+1. \end{cases}$$

The other elements of the matrix $\boldsymbol{D}_i$, $\boldsymbol{E}_c$, $\boldsymbol{F}_i$, $\boldsymbol{A}_i$ are 0.

$$\boldsymbol{E}_0=\left(-\lambda\right), \boldsymbol{C}_0=\begin{pmatrix} \boldsymbol{C}_1 \\ \boldsymbol{C} \end{pmatrix},$$

$$\boldsymbol{C}=\begin{pmatrix} b\lambda & & & & & \\ & b\lambda & & & & \\ & & b\lambda & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & b\lambda \end{pmatrix},$$

$$\boldsymbol{B}_i=\begin{pmatrix} i\theta\left(1-r\right) & & & & \\ & i\theta\left(1-r\right) & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & i\theta\left(1-r\right) \end{pmatrix}.$$

From the structure of the $\boldsymbol{Q}$ matrix, it appears that the Markov process is a QBD process.

III. Steady-state Conditions

Theorem 1. If $N\theta(1-r)>b\lambda$, the system is stationary.

Proof Assuming $\boldsymbol{H}=\boldsymbol{A}_N+\boldsymbol{B}_N+\boldsymbol{C}$ is an infinitesimal generator, $\boldsymbol{X}$ is a stationary probability vector for $\boldsymbol{H}$, then the sufficient and necessary condition to ensure the existence of the steady-state probability distribution of $\{N\left(t\right), I\left(t\right), J\left(t\right), t\ge 0\}$ is $\boldsymbol{X}\boldsymbol{B}_N\boldsymbol{e}>\boldsymbol{X}\boldsymbol{C}\boldsymbol{e}$.

$$\boldsymbol{H}=\begin{pmatrix} -c\alpha & c\alpha & & & \\ \beta & -\beta-(c-1)\alpha & (c-1)\alpha & & \\ & & & \ddots & \\ & & \ddots & \ddots & \alpha \\ & & & c\beta & -c\beta \end{pmatrix}.$$

Assuming

$$\boldsymbol{X}=\{x_1, x_2, \cdots\cdots, x_N\},$$

$$\begin{cases} \boldsymbol{X}\boldsymbol{H}=0, \\ \boldsymbol{X}\boldsymbol{e}=1. \end{cases} \tag{5}$$

Taking $\boldsymbol{H}$ and $\boldsymbol{X}$ into Eq. (5) obtain:

$$x_1=\frac{1}{1+\sum\limits_{i=1}^{c}\prod\limits_{j=1}^{i}\frac{(c+1-j)\alpha}{j\beta}}, \tag{6}$$

$$x_i=\frac{\prod\limits_{j=1}^{i-1}\frac{(c+1-j)\alpha}{j\beta}}{1+\sum\limits_{i=1}^{c}\prod\limits_{j=1}^{i}\frac{(c+1-j)\alpha}{j\beta}}, 2\le i\le N. \tag{7}$$

Form $\boldsymbol{X}\boldsymbol{B}_N\boldsymbol{e}>\boldsymbol{X}\boldsymbol{C}\boldsymbol{e}$ obtain:

$$N\theta(1-r)>b\lambda.$$

Then the system is stationary.

IV. Steady-State Probability

If the matrix equation $\boldsymbol{R}^2\boldsymbol{B}_N + \boldsymbol{R}\boldsymbol{A}_N + \boldsymbol{C} = 0$ has a minimal non-negative solution R, and its spectral radius $SP(\boldsymbol{R})<1$, the QBD process $\{N(t), I(t), J(t), t \geq 0\}$ is positive recurrence. Additionally the system of linear homogeneous equations $(\boldsymbol{P}_0, \boldsymbol{P}_1, \cdots, \boldsymbol{P}_N)\,\boldsymbol{B}[\boldsymbol{R}] = 0$ has a positive solution.

When the Markov process is positive recurrence, the steady-state probability is defined as:

$$P_{i,j,k} = \lim_{t\to\infty} P\{N(t)=i, I(t)=j, J(t)=k\}, (i,j,k)\in\Omega.$$

The steady-state probability vector is

$$\boldsymbol{P} = (\boldsymbol{P}_0, \boldsymbol{P}_1, \boldsymbol{P}_2, \cdots),$$

where

$$\begin{cases} \boldsymbol{P}_0 = (P_{0,0,0}, P_{0,1,0}, P_{0,1,1}, P_{0,2,0}, P_{0,2,1}, \cdots, P_{0,c,c}), \\ \boldsymbol{P}_l = (P_{l,c,0}, P_{l,c,1}, P_{l,c,2}, \cdots, P_{l,c,c}), l>0. \end{cases}$$

Taking the steady-state probability vector into the equilibrium equation, then

$$\begin{array}{l} \boldsymbol{P}_0\boldsymbol{A}_0 + \boldsymbol{P}_1\boldsymbol{B}_1 = 0, \\ \boldsymbol{P}_0\boldsymbol{C}_0 + \boldsymbol{P}_1\boldsymbol{A}_1 + \boldsymbol{P}_2\boldsymbol{B}_2 = 0, \\ \boldsymbol{P}_1\boldsymbol{C} + \boldsymbol{P}_{l+1}\boldsymbol{A}_{l+1} + \boldsymbol{P}_{l+2}\boldsymbol{B}_{l+2} = 0, 1\leq l\leq N-1, \\ \boldsymbol{P}_l\boldsymbol{C} + \boldsymbol{P}_{l+1}\boldsymbol{A}_N + \boldsymbol{P}_{l+2}\boldsymbol{B}_N = 0, l\geq N. \end{array} \tag{8}$$

To solve for the boundary probability vector, a $(c+1)\times(c+2)/2+N\times(c+1)$-dimensional random matrix is constructed as follows:

$$\boldsymbol{B}[\boldsymbol{R}] = \begin{pmatrix} \boldsymbol{A}_0 & \boldsymbol{C}_0 & & & \\ \boldsymbol{B}_1 & \boldsymbol{A}_1 & \boldsymbol{C} & & \\ & \ddots & \ddots & \ddots & \\ & & \boldsymbol{B}_{N-1} & \boldsymbol{A}_{N-1} & \boldsymbol{C} \\ & & & \boldsymbol{B}_N & \boldsymbol{R}\boldsymbol{B}_N + \boldsymbol{A}_N \end{pmatrix}.$$

Theorem 2. If $\{N(t), I(t), J(t), t\geq 0\}$ is a positive recurrence, the steady-state distribution of the system satisfies the following equations:

$$\begin{cases} (\boldsymbol{P}_0, \boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)\boldsymbol{B}[\boldsymbol{R}] = 0, \\ \boldsymbol{P}_0\boldsymbol{e}_1 + \sum\limits_{l=1}^{N-1}\boldsymbol{P}_l\boldsymbol{e}_2 + \boldsymbol{P}_N(\boldsymbol{I}-\boldsymbol{R})^{-1}\boldsymbol{e}_2 = 1, \\ \boldsymbol{P}_l = \boldsymbol{P}_N\boldsymbol{R}^{l-N}, l\geq N, \end{cases} \tag{9}$$

where $\boldsymbol{e}_1 = (1,1,\cdots,1)^T$ is a $(c+1)(c+2)/2$ dimensional column vector, $\boldsymbol{e}_2 = (1,1,\cdots,1)^T$ is a $(c+1)$ dimensional column vector, $\boldsymbol{I}$ is a $(c+1)$ dimensional unit matrix.

Proof 1) Proving that

$$\boldsymbol{P}_l = \boldsymbol{P}_N\boldsymbol{R}^{l-N}, l\geq N. \tag{10}$$

When $l=N$, $\boldsymbol{P}_l = \boldsymbol{P}_N\boldsymbol{R}^{l-N}$ clearly holds. When $l>N$, the hypothesis test is applied. Assuming that $\boldsymbol{P}_l = \boldsymbol{P}_N\boldsymbol{R}^{l-N}$ holds, take it into the equilibrium equation Eq. (8), then

$$\begin{array}{l} \boldsymbol{P}_l C + \boldsymbol{P}_{l+1}\boldsymbol{A}_N + \boldsymbol{P}_{l+2}\boldsymbol{B}_N \\ = \boldsymbol{P}_N\boldsymbol{R}^{l-N}\boldsymbol{C} + \boldsymbol{P}_N\boldsymbol{R}^{l+1-N}\boldsymbol{A}_N + \boldsymbol{P}_N\boldsymbol{R}^{l+2-N}\boldsymbol{B}_N \\ = \boldsymbol{P}_N\boldsymbol{R}^{l-N}(\boldsymbol{C} + \boldsymbol{R}\boldsymbol{A}_N + \boldsymbol{R}^2\boldsymbol{B}_N) \\ = 0, \end{array}$$

so $\boldsymbol{P}_l = \boldsymbol{P}_N\boldsymbol{R}^{l-N}, l\geq N$ holds.

2) Proving that

$$(\boldsymbol{P}_0, \boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)\boldsymbol{B}[\boldsymbol{R}] = 0.$$

Bringing the matrix $\boldsymbol{B}[\boldsymbol{R}]$ into the above equation, we can obtain:

$$\begin{array}{l} (\boldsymbol{P}_0, \boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)\boldsymbol{B}[\boldsymbol{R}] \\ = (\boldsymbol{P}_0\boldsymbol{A}_0 + \boldsymbol{P}_1\boldsymbol{B}_1, \cdots, \boldsymbol{P}_{N-1}\boldsymbol{C} + \boldsymbol{P}_N(\boldsymbol{R}\boldsymbol{B}_N + \boldsymbol{A}_N)) \\ = (\boldsymbol{P}_0\boldsymbol{A}_0 + \boldsymbol{P}_1\boldsymbol{B}_1, \cdots, \boldsymbol{P}_{N-1}\boldsymbol{C} + \boldsymbol{P}_N\boldsymbol{A}_N + \boldsymbol{P}_N\boldsymbol{R}\boldsymbol{B}_N), \end{array} \tag{11}$$

substituting Eq. (10) into Eq. (11), then

$$\begin{array}{l} (\boldsymbol{P}_0, \boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)\boldsymbol{B}[\boldsymbol{R}] \\ = (\boldsymbol{P}_0\boldsymbol{A}_0 + \boldsymbol{P}_1\boldsymbol{B}_1, \cdots, \boldsymbol{P}\boldsymbol{P}_{N-1}\boldsymbol{C} + \boldsymbol{P}_N\boldsymbol{A}_N + \boldsymbol{P}_{N+1}\boldsymbol{B}_N). \end{array}$$

From the equilibrium equation Eq. (8) obtain:

$$(\boldsymbol{P}_0, \boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_N)\boldsymbol{B}[\boldsymbol{R}] = 0.$$

3) Proving that

$$\boldsymbol{P}_0\boldsymbol{e}_1 + \sum_{l=1}^{N-1}\boldsymbol{P}_l\boldsymbol{e}_2 + \boldsymbol{P}_N(\boldsymbol{I}-\boldsymbol{R})^{-1}\boldsymbol{e}_2 = 1.$$

By the regularisation condition of $\boldsymbol{P}\boldsymbol{e} = 1$ and the theorem of $\boldsymbol{P}_l = \boldsymbol{P}_N\boldsymbol{R}^{l-N}, l\geq N$, we have

$$\begin{array}{l} \boldsymbol{P}\boldsymbol{e} \\ = \boldsymbol{P}_0\boldsymbol{e}_1 + \sum\limits_{l=1}^{\infty}\boldsymbol{P}_l\boldsymbol{e}_2 \\ = \boldsymbol{P}_0\boldsymbol{e}_1 + \boldsymbol{P}_1\boldsymbol{e}_2 + \cdots + \boldsymbol{P}_N\boldsymbol{e}_2 + \boldsymbol{P}_{N+1}\boldsymbol{e}_2 + \boldsymbol{P}_{N+2}\boldsymbol{e}_2 + \cdots \\ = \boldsymbol{P}_0\boldsymbol{e}_1 + \sum\limits_{l=1}^{N-1}\boldsymbol{P}_l\boldsymbol{e}_2 + \boldsymbol{P}_N\boldsymbol{e}_2 + \boldsymbol{P}_N\boldsymbol{R}\boldsymbol{e}_2 + \boldsymbol{P}_N\boldsymbol{R}^2\boldsymbol{e}_2 + \cdots \\ = \boldsymbol{P}_0\boldsymbol{e}_1 + \sum\limits_{l=1}^{N-1}\boldsymbol{P}_l\boldsymbol{e}_2 + \boldsymbol{P}_N(\boldsymbol{I} + \boldsymbol{R} + \boldsymbol{R}^2 + \cdots)\boldsymbol{e}_2. \end{array}$$

Since $SP(\boldsymbol{R})<1$, $\boldsymbol{I} + \boldsymbol{R} + \boldsymbol{R}^2 + \cdots$ converges to $(\boldsymbol{I}-\boldsymbol{R})^{-1}$.

Then we have

$$\boldsymbol{P}_0\boldsymbol{e}_1 + \sum_{l=1}^{N-1}\boldsymbol{P}_l\boldsymbol{e}_2 + \boldsymbol{P}_N(\boldsymbol{I}-\boldsymbol{R})^{-1}\boldsymbol{e}_2 = 1.$$

The analysis mentioned above is based on the matrix geometric solution theory. To analyze the various performance indexes of the system, it is essential to solve the exact expression of R. However, due to the large dimensionality and complex structure of the matrix, finding the specific expression of $\boldsymbol{R}$ is difficult. Therefore, we employ the Gauss-Seidel iterative algorithm to obtain an approximate solution for the rate matrix $\boldsymbol{R}$. To ensure the convergence of the algorithm, we set the accuracy of the algorithm $\varepsilon(\varepsilon = 10^{-10})$. The detailed steps of the algorithm are shown in Table I.

V. System Steady-state Performance Measures

1) The expected number of customers in the retry space

$$E(O) = \sum_{i=1}^{\infty} i\boldsymbol{P}_i = \sum_{i=1}^{\infty}\sum_{j=1}^{c}\sum_{k=0}^{j} iP_{i,j,k}.$$

2) The steady-state mean queue length of the system

$$E(L) = \boldsymbol{P}_0\boldsymbol{u}_1 + \sum_{i=1}^{\infty} i\boldsymbol{P}_i = \boldsymbol{P}_0\boldsymbol{u}_1 + \sum_{i=1}^{\infty}\sum_{k=0}^{c} i\boldsymbol{P}_i,$$

where

$$\boldsymbol{e}_3 = (0,1,1,2,2,2,3,3,3,3,\cdots c,c,\cdots,c)^T.$$

Table I. Gauss-Seidel iterative algorithm.

Step	Operation
Step 1	Set b, r, N, c, η, λ, μ, ε and $R = 0$
Step 2	Input $\boldsymbol{A}_N$, $\boldsymbol{B}_N$, $\boldsymbol{C}$
	$\boldsymbol{R}_n = \boldsymbol{R}$
Step 3	Define $n = n + 1$
	$\boldsymbol{R}_n = -(\boldsymbol{R}_{n-1}^2 \boldsymbol{B}_N + \boldsymbol{C})\boldsymbol{A}_N^{-1}$
Step 4	If $\lVert \boldsymbol{R}_{i+1} - \boldsymbol{R}_i \rVert > \varepsilon$, Go to Step 3;
	else, Go to Step 5
Step 5	$\boldsymbol{R}_n = \boldsymbol{R}$

3) The expected number of breakdown servers

$$E(BD) = \boldsymbol{P}_0 \boldsymbol{u}_2 + \sum_{i=1}^{\infty} \boldsymbol{P}_i \boldsymbol{u}_3 = \boldsymbol{P}_0 \boldsymbol{u}_2 + \sum_{i=1}^{\infty} \sum_{k=0}^{c} \boldsymbol{P}_i \boldsymbol{u}_3,$$

where

$$\boldsymbol{u}_2 = (0, 0, 1, 0, 1, 2, 0, 1, 2, 3, \cdots 0, 1, 2, 3, \cdots c)^T,$$

$$\boldsymbol{u}_3 = (0, 1, 2, 3 \cdots c-1, c)^T.$$

4) The number of impatient customers leaving the system

$$E(L_F) = (1-r)(c+i) \sum_{i=1}^{\infty} \sum_{k=0}^{c} P_{i,j,k}.$$

5) The average waiting time of customers in the system

$$E(W) = \frac{E(L)}{\lambda}.$$

VI. Numerical Experiments

A. Performance Indicator Analysis

After obtaining the steady-state probability through the iterative algorithm and the system of steady-state equations mentioned earlier, we can use MATLAB to plot the trend of the system's performance indicators.

Assuming that $b = 0.7, r = 0.7, \theta = 5, N = 30, \beta = 2, \alpha = 0.5, \mu = 2.5, \eta = 1.6$. Fig. 2 depicts the relationship between λ and $E(O)$ for different numbers of servers. As the number of servers c remains fixed, $E(O)$ increases with an increase in λ, but the rate at which it increases gradually slows down after a certain point.

Assuming $\lambda = 20$ and $\eta = 1$, according to Fig. 2, Fig. 3 displays the correlation between μ and $E(O)$ in different numbers of servers. As μ increases, $E(O)$ gradually decreases. This phenomenon occurs because as μ increases, the probability of the customers in the retry space entering the service area by retrying also increases. As a result, $E(O)$ decreases accordingly.

For different numbers of service stations, Fig. 4 shows the effect of b and c on $E(O)$ in the retry space. When c is constant, $E(O)$ increases as b increases and when b is constant, $E(O)$ decreases as c increases.

Fig. 5 reveals the significant influence of λ and β on $E(BD)$. When λ is fixed, $E(BD)$ decreases as β increases. Conversely, when β is constant, $E(BD)$ increases as λ

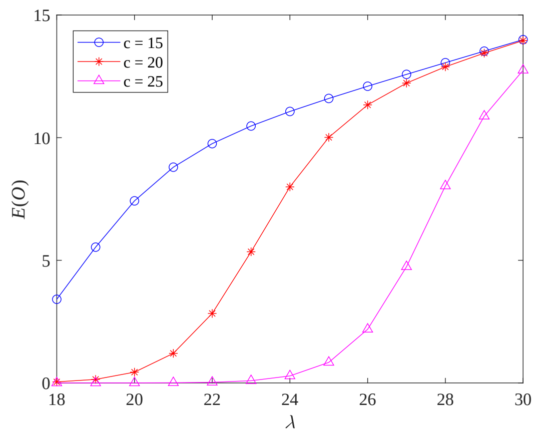

Fig. 2. The trend of $E(O)$ versus λ and c ($\mu = 2.5$, $\beta = 2$, $\alpha = 0.5$, $\eta = 1.6$, $b = 0.7$, $r = 0.7$, $\theta = 5$ and $N = 30$).

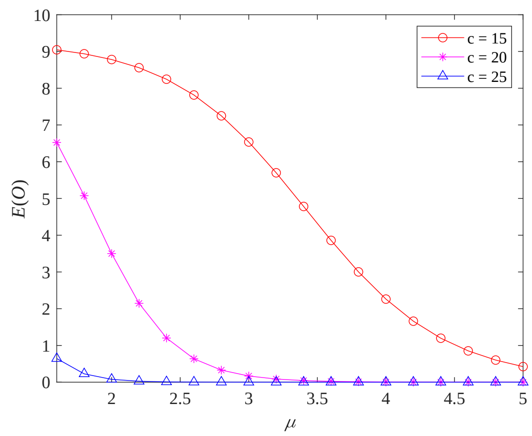

Fig. 3. The trend of $E(O)$ versus μ and c ($\lambda = 20$, $\eta = 1$, $\beta = 2$, $\alpha = 0.5$, $b = 0.7$, $r = 0.7$, $\theta = 5$ and $N = 30$).

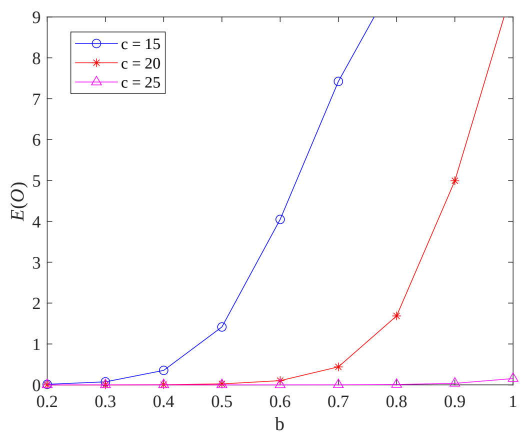

Fig. 4. The trend of $E(O)$ versus b and c ($\lambda = 20$, $\eta = 1$, $\beta = 2$, $\alpha = 0.5$, $\mu = 2.5$, $r = 0.7$, $\theta = 5$ and $N = 30$).

increases. This is largely because servers are more likely to break down when they are busy. As λ increases, more servers get busy, resulting in an increase in $E(BD)$.

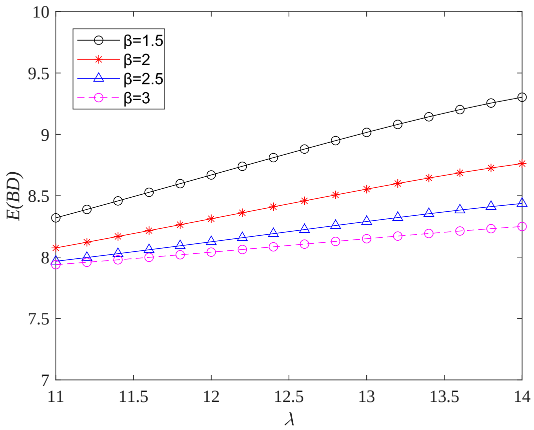

Fig. 5. The trend of $E(BD)$ versus λ and β ($\mu = 2.5$, $\alpha = 0.5$, $\eta = 1.6$, $b = 0.7$, $r = 0.7$, $\theta = 5$, $c = 15$ and $N = 30$).

For different numbers of service stations, Fig. 6 illustrates the effect of θ on E(BD). When c is constant, E(BD) increases with θ. The rate at which E(BD) enhances with θ is greater with larger c.

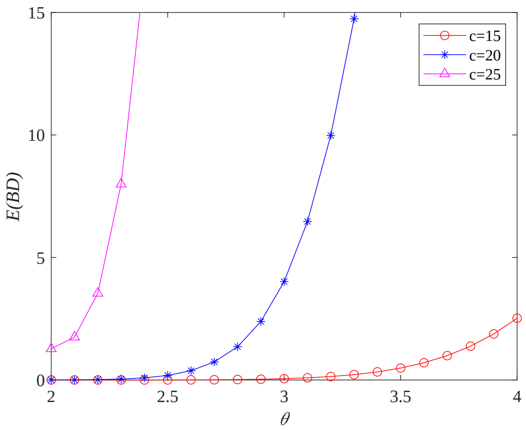

Fig. 6. The trend of $E(BD)$ versus θ and c ($\lambda = 20$, $\mu = 2.5$, $\alpha = 0.5$, $\eta = 1.6$, $b = 0.7$, $r = 0.7$, $\beta = 2$ and $N = 30$).

Fig. 7 illustrates the impact of α and β on $E(L_I)$. When β is constant, $E(L_I)$ increases as α increases. As the failure rate rises, so does the number of faulty servers, resulting in a decline in service rate and a greater influx of customers in the retry space. However, fewer customers enter the service area by retrying, causing $E(L_I)$ increases. When α is constant, $E(L_I)$ increases as β decreases. Additionally, when β decreases, the number of servers transitioning from low to normal service rates also decreases. Consequently, the low service rate persists, preventing customers from moving from the retry space to the service area, leading to more impatient customers.

Fig. 8 depicts $E(O)$ versus θ for different service rates μ. As θ increases, $E(O)$ decreases. The probability of a customer leaving the retry space and entering the service area increases as θ increases, while the number of impatient

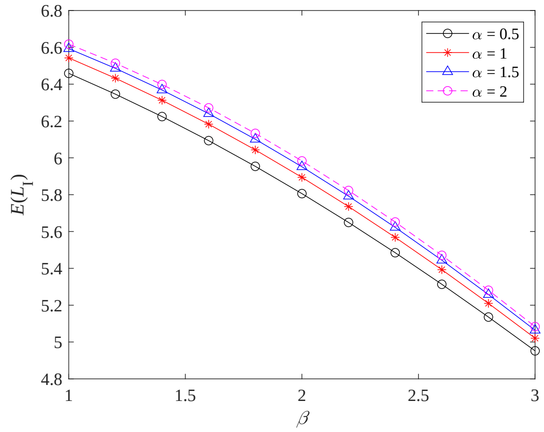

Fig. 7. The trend of $E(L_I)$ versus α and β ($\lambda = 20$, $\eta = 1.6$, $\mu = 2.5$, $b = 0.7$, $r = 0.7$, $\theta = 5$, $c = 15$ and $N = 30$).

customers decreases, leading to a smaller $E(O)$. The $E(O)$ decreases faster with a larger value of μ.

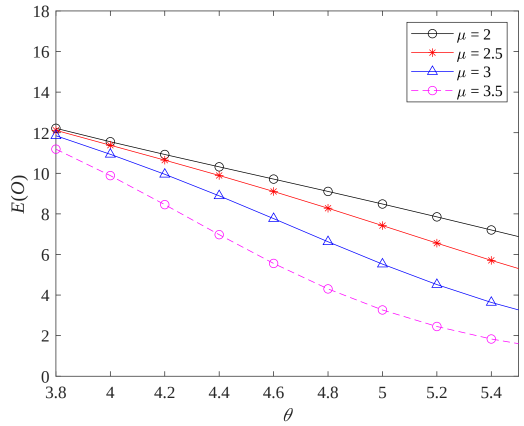

Fig. 8. The trend of $E(O)$ versus θ and μ ($\lambda = 20$, $\eta = 1.6$, $\beta = 2$, $\alpha = 0.5$, $b = 0.7$, $r = 0.7$, $c = 15$ and $N = 30$).

Fig. 9 gives a three-dimensional plot of failure rate α and arrival rate λ against $E(L)$. As λ increases, $E(L)$ also increases, while $E(L)$ decreases as α decreases.

B. Cost Analysis

After conducting the aforementioned analysis, it has become apparent that each parameter has a significant impact on the system. To optimize the model, we can alter the parameter values. To determine the most effective way to manage the system and assess its economic feasibility, we will introduce a cost function F. This function will allow us to minimize costs and expenses, thereby enhancing the system's efficiency. The system cost parameters are as follows:

1) The cost of a customer's stay in the retry space c_h.
2) The cost per unit of time that a server is busy c_b.
3) The unit time cost of a server failure c_d.
4) The fixed cost of providing a server c_f.
5) The cost per unit of time to serve customers c_s.
6) The unit time cost of providing repairs c_r.

Then the cost function of the system per unit of time is

$$F = c_h E(O) + c_b E(B) + c_d E(BD) + c_f c + c_s \mu + c_r \beta.$$

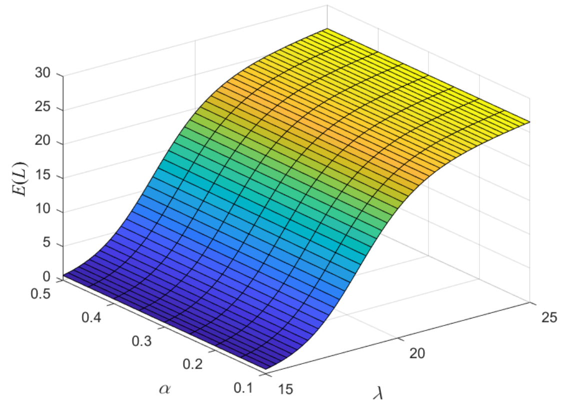

Fig. 9. The trend of $E(L)$ versus α and λ ($\mu = 2.5$, $\beta = 2$, $\eta = 1.6$, $b = 0.7$, $r = 0.7$, $\theta = 5$, $c = 15$ and $N = 30$).

Table II. The cost function F for various values of β and μ.

(β, μ)	$c = 6$	$c = 8$	$c = 10$	$c = 12$
(1.5,1.6)	773.3	204.5	292.4	331.0
(2,2.5)	1013.1	222.2	259.9	267.8
(2,3.5)	1177.4	210.9	224.9	227.9
(3,3.5)	1235.0	203.9	213.5	216.1

Table III. The cost function F for various values of θ and λ.

(θ, λ)	$c = 6$	$c = 8$	$c = 10$	$c = 12$
(4,6)	1521.5	183.6	186.4	188.5
(5,6)	1574.5	184.0	186.5	188.5
(5,8)	1571.8	217.9	225.5	228.2
(5,10)	1013.1	222.3	259.9	267.8
(6,10)	1439.9	241.0	263.1	268.1

Table IV. The cost function F for various values of b and r.

(b, r)	$c = 6$	$c = 8$	$c = 10$	$c = 12$
(0.3,0.9)	439.0	170.2	245.4	265.5
(0.8,0.9)	70.0	80.1	90.9	110.6
(0.8,0.5)	1782.7	250.5	264.5	267.2
(0.8,0.2)	2104.5	257.2	265.4	267.3

Assuming $c_h = 2, c_b = 4, c_d = 2.5, c_f = 1, c_s = 2$ and $c_r = 1.5$, the following table shows the effect of changes in the system parameters on the cost.

Based on Table II, we can conclude that the cost of the system decreases with the increase of β, and decreases with the increase of μ. The analysis reveals that when β increases, the number of faulty machines reduces, and the total system service rate goes up. An increase in μ leads to more customers successfully entering the system, fewer people in the retry space, and a lower average retry rate, fewer customers leave due to impatience and lower costs. However, with a fixed arrival rate, if the number of servers is too small, the cost will increase and the system revenue will be reduced.

Table III gives the effect of θ and λ on the cost F. As the arrival and retry rates increase, the cost of maintaining the retry space gradually goes up. Furthermore, when the number of people in the retry space increases due to a higher λ, and is coupled with an increase in θ, the cost of maintaining the retry space in the system also goes up even further.

Table IV shows the effect of b and r on the cost expense F. The combination of the three tables shows that the system has the lowest cost when $b = 0.8$, and $r = 0.9$. With fixed parameter $\lambda = 10$, the number of service stations is set to $c = 8$, while $\beta = 3$ and $\mu = 3.5$, resulting in the lowest cost expense.

The above analysis shows that to maximize the benefits and minimize the costs of the system, the appropriate number of service stations should be set according to the arrival rate. It's important to monitor the status of machines promptly to reduce the failure rate of servers. Additionally, improving the repair rate can enhance the service efficiency of the servers. These measures can collectively ensure that the system is running smoothly and efficiently.

C. *Benefit Analysis*

In this section, we construct benefit functions from both individual and societal perspectives.

Assuming that the personal benefit to the customer for a completed service is Z, the expenditure per unit of time that the customer stays in the system is G, and the fee that the customer pays for accessing the system is f. Assuming that U_I denotes the customer's personal benefit, then

$$U_I = Z - f_1 E(W) - f_2.$$

Assuming that $b = 0.7, r = 0.7, \theta = 5, N = 30, \lambda = 15, \alpha = 0.5, \eta = 1, c = 15, Z = 50, f_1 = 3$, and $f_2 = 2$. Fig. 10 depicts the impact of restoration β and μ on U_I. When β is constant, U_I increases with μ. As μ increases, customers experience reduced wait times within the system, resulting in lower average costs per unit of time spent. This, in turn, leads to enhanced individual benefits. Similarly, when μ is constant, U_I decreases as β decreases.

In order to devise the most effective strategy for the benefit of society, we define the social benefit function as:

$$U_S = \lambda(Z - f_1 E(W) - f_2).$$

Based on the given assumptions, Fig. 11 depicts U_S in relation to β and μ. It can be observed that when μ remains constant, it increases as β increases, and when β remains constant, it increases as μ increases. However, if both parameters are decreased simultaneously, U_S will decrease and may even become negative. Therefore, it is crucial to appropriately increase β and μ in case of server breakdown to avoid negative social benefits and ensure the maximization of social benefits in the system.

VII. Conclusion

In this paper, we analyze the M/M/c retrial queueing system with impatient customers and server working breakdown.

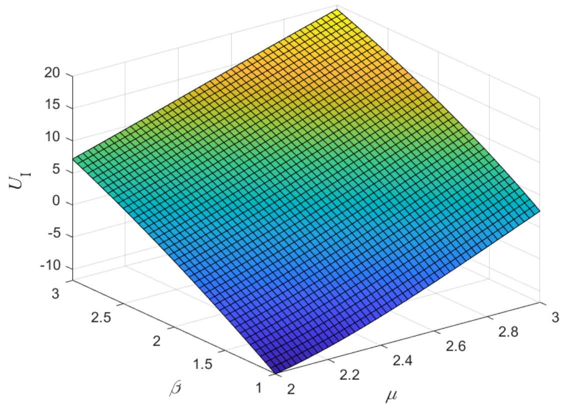

Fig. 10. The trend of U_I versus β and μ ($b = 0.7$, $r = 0.7$, $\theta = 5$, $N = 30$, $\lambda = 15$, $\alpha = 0.5$, $\eta = 1$, $c = 15$, $Z = 50$, $f_1 = 3$, $f_2 = 2$).

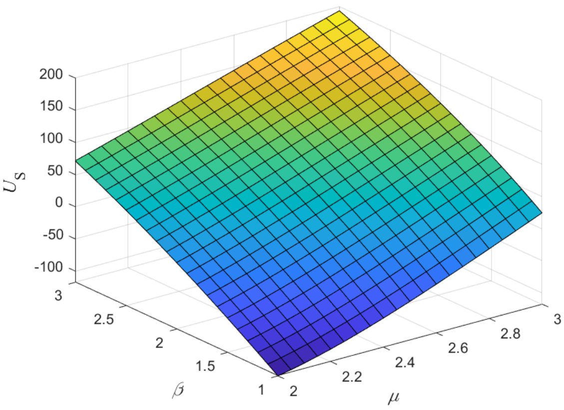

Fig. 11. The trend of U_S versus β and μ ($b = 0.7$, $r = 0.7$, $\theta = 5$, $N = 30$, $\lambda = 15$, $\alpha = 0.5$, $\eta = 1$, $c = 15$, $Z = 50$, $f_1 = 3$, $f_2 = 2$).

We use matrix-geometric solutions and iterative algorithms to compute the steady-state probabilities and determine the performance indicators of the system. To assess the impact of each parameter on the indicators, we conduct numerical experiments using MATLAB. Finally, we optimize the model using cost and benefit functions to determine the appropriate parameters. In practical applications, it is crucial to establish suitable servers based on arrival and service rates, along with regular server maintenance to minimize failure rates. The results are significant to various systems such as the Internet ticketing system and after-sales service center system.

References

[1] T. V. Do, "An Efficient Solution to a Retrial Queue for the Performability Evaluation of DHCP," *Computers & Operations Research*, vol. 37, pp. 1191-1198, 2010.

[2] J. Kim and B. Kim, "Survey of Retrial Queueing Systems," *Ann Oper Res*, vol. 247, pp. 3-36, 2016.

[3] Q. Q. Ye and Y. Chang, "Performance Analysis of M/M/1 Retrial Queuing System with Working Breakdowns," *Journal of Henan Normal University (Natural Science Edition)*, vol. 51, no. 3, pp. 82-89, 2023.

[4] M. S. Kumar, K. Sohraby and K. Kim, "Delay Analysis of Orderly Reattempts in Retrial Queueing System with Phase Type Retrial Time," *IEEE Communications Letters*, vol. 17, no. 5, 2013.

[5] Y. M. Shin, "Monotonicity Properties in Various Retrial Queues and Their Applications," *Queueing Systems*, vol. 53, pp. 147-157, 2006.

[6] S. Upadhyaya, R. Sharma, D. Agarwal and G. Malik, "Convexity Analysis and Cost Optimization of a Retrial Queue with Bernoulli Vacation and Delayed Phase Mending," *International Journal System Assurance Engineering and Management*, vol. 14, no. 5, pp. 1671-1690, 2023.

[7] Y. N. Han, R. L. Tian, X. Y. Wu and L. Q. He, "On a Retrial Queue with Negative Customers, Passive Breakdown, and Delayed Repairs," *Probability in the Engineering and Informational Sciences*, pp. 1-20, 2023.

[8] Y. Zhang and J. T. Wang, "Equilibrium Analysis in The Retrial Queue with an Unreliable Server," *Operations Research Transactions*, vol. 26, no. 2, pp. 1-15, 2022.

[9] V. Karthick and V. Suvitha, "An analysis of three servers markovian multiple vacation queueing system with servers breakdown," *IAENG International Journal of Applied Mathematics*, vol. 53, no. 2, pp. 670-677, 2023.

[10] S. L. Lv, M. Yang and J. Y. Wen, "The M/M/1 Repairable Queueing System with Two Types of Server Breakdowns and Negative Customers," *Engineering Letters*, vol. 32, no. 1, pp. 77-83, 2024.

[11] K. Kalidass and R. Kasturi, "A Queue with Working Breakdowns," *Computers & Industrial Engineering*, vol. 63, pp. 779-783, 2012.

[12] S. L. Lv, H. L. Fu and J. B. Li, "Two Types of Failures M/M/1 Queueing System with One Standby Service Station and Start-up Time," *Engineering Letters*, vol. 31, no. 4, pp. 1847-1852, 2023.

[13] T. Jing, Q. Q. Ye, X. L. Zhang and J. S. Liu, "The Performance Analysis of a Retrial M/M/1 Queue with Working Breakdowns and Balking," *Journal of Chongqing Normal University (Natural Science)*, vol. 37, no. 5, pp. 1-9, 2020.

[14] S. L. Lv, Y. Y. Wu, J. B. Li and R. Y. Wang, "The M/M/1 Repairable Queueing System with Variable Input Rates and Failure Rates," *IAENG International Journal of Applied Mathematics*, vol. 52, no. 4, pp. 784-790, 2022.

[15] Q. Q. Ye and L. W. Liu, "Analysis of MAP/M/1 Queue with Working Breakdowns," *CommunicationS In Statistics—Theory And Methods*, vol. 47, no. 13, pp. 3073-3084, 2018.

[16] X. J. Yang, Z. X. Li, S. P. Li and F. Wu, "Performance Analysis of M/M/1/N Queue with Setup Time and Working Breakdown," *Control Theory & Applications*, vol. 36, no. 4, pp. 561-569, 2019.

[17] S. T. Zhang and X. L. Xu, "Equilibrium Analysis of Queueing System with Two Types of Parallel Customers and Partial Failures," *Journal of Systems Science and Mathematical Sciences*, vol. 39, no. 4, pp. 637-647, 2019.

[18] R. Kumar and S. Sharma, "Transient Analysis of an M/M/c Queuing System with Balking and Retention of Reneging Customers," *Communications in Statistics - Theory and Methods*, vol. 47, no. 6, pp. 1318-1327, 2017.

[19] B. Krishna Kumar and J. Raja, "On Multiserver Feedback Retrial Queues with Balking and Control Retrial Rate," *Ann Oper Res*, vol. 141, pp. 211-232, 2006.

[20] H. L. Liu and Q. L. Li, "Performance Analysis of an MAP/PH/1 Queueing System with Impatient Customers," *Journal of Systems Science and Mathematical Sciences*, vol. 43, no. 7, pp. 1819-1836, 2023.

[21] Q. Q. Shan and D. Q. Yue, "Queueing-inventory System with Impatient Customers and Multiple Working Vacations," *Mathematics in Practice and Theory*, vol. 51, no. 4, pp. 133-144, 2021.

[22] F. M. Chang, T. H. Liu and J. C. Ke, "On an unreliable-Server Retrial Queue with Customer Feedback and Impatience," *Applied Mathematical Modelling*, vol. 55, pp. 171-182, 2018.