Mathematics

Research article

Joining strategies under two kinds of games for a multiple vacations retrial queue with N-policy and breakdowns

Zhen Wang[1], Liwei Liu[2,*], Yuanfu Shao[3,*] and Yiqiang Q. Zhao[4]

[1] School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin, 541004, Guangxi, China

[2] School of Science, Nanjing University of Science and Technology, Nanjing 210094, Jiangsu, China

[3] School of Science, Guilin University of Technology, Guilin 541004, Guangxi, China

[4] School of Mathematics and Statistics, Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada

* **Correspondence:** Email: NJUSTliuliwei@163.com, wzailjj@163.com; Tel: +8602584315586.

Abstract: Motivated by cost control and information guidance, in this work, we study a multiple vacations retrial queue with N-policy and breakdowns. This service system has the characteristics that there is no waiting space in front of the server and the waiting list is virtual. If the arriving customer finds that the system is available, he immediately receives the complete service. Otherwise, the customer leaves the system or joins the orbit (virtual waiting list). For cost control, the system is activated only when the current vacation is completed and at least N customers are waiting in the system, otherwise, the server continues to the next vacation until the number of customers in the system is not less than N. Two types of customer joining cases apply to this paper, i.e., non-cooperative customers aim to optimize individual interests, and the social planner in the cooperative case considers the profit of the whole service system. The equilibrium joining strategy for the non-cooperative case and the socially optimal joining strategy for the cooperative case are determined. Since it is difficult to obtain analytical characterization, an improved particle swarm optimization (PSO) algorithm is used to explore the impact of system parameters on the profit of the service provider. At the same time, a large number of numerical experiments visualize the influence of parameters on the system.

1. Introduction

In many service systems, behavioral analysis of the arriving customers' decision whether to join an unavailable system (where service seats are occupied) or leaving forever has become a hot topic in economics in recent years. More specifically, the introduction of customers' strategic behavior in the queuing system makes the analysis of performance indicators more meaningful. Although it is not very difficult to deal with the performance analysis of various M/M/1 systems, it is still difficult to conduct game theory analysis in specific such systems. In fact, the analysis of customer game behavior requires some characteristics of the system, so it is usually very difficult for the Markov process. Furthermore, considering the customer's strategic behavior in the service system can make the customer's subjective behavior theoretical and logical. The joining strategy of customers was first studied by Naor [1]. He studied the M/M/1 system under the observable case, in which the individual's equilibrium joining strategy and social optimal joining strategy were confirmed. Subsequently, the unobservable case of Nash's work was supplemented by Edelson and Hildebrand [2]. In recent years, the research of customer joining strategy in queuing systems has been widely expanded and applied. Such as Shi and Lian [3] presented the joining strategy and socially optimal welfare of passenger-taxi in both observable and unobservable cases, catastrophe or clearing systems with equilibrium [4, 5], and queueing systems with setup or vacation [6–8]. More specifically, the monograph of Hassin and Haviv [9] summarized a large number of related studies on joining strategies and optimal social welfare.

In many queuing service systems, their default rule is that the customers leave the system once they balk and cannot join the system again. Obviously, this rule is often not suitable for the actual service system. For example, when the customer calls the call center for the first time and encounters a busy line, the customer will try again after a random length of time. The retrial queuing system widely exists in other fields such as communication, information processing, storage management, and epidemic situation control. Kosten [10] illustrated the necessity of retrial queuing through practical applications. Jiang [11] applied a tail asymptotics approach to a matching service voting system. Phung-Duc [12] presented an M/G/1 single server retrial queue with setup time. However, there are still a few works of literature on the joining strategy and social welfare of various retrial queue systems. Economou and Kanta [13] studied joining strategies and social welfare in the M/M/1 queue with retrial. Wang and Li [14] proposed a class of cognitive radio networks with retrial behavior. Wang and Zhang [15] studied price strategy and equilibrium of a local area networking retrial queue with delayed vacations.

The conservation and effective use of resources and energy is the key to sustainable economic growth. Barroso and Holzle [16] pointed out that the proportion of energy consumed by standby equipment is also very high, it can reach 60 percent of the normal operation equipment. In the queuing service system, it is necessary to take some measures to control operating costs and labor costs, which can effectively promote the growth of social welfare. For example, threshold strategy, multiple vacations, and adjusted maintenance times are often applied to queuing systems with detailed services. Wang et al. [17] studied a batch service queuing system with finite capacity and gated policy. Guo and Hassin [18] discussed joining strategy and social optimization of a single server queueing system with N-policy. Wang et al. [19] studied customers' strategic behavior and social optimization in a constant retrial queue with the N-policy. Sun et al. [20] presented the strategies of joining and optimal balking in a Markovian queue with multiple vacations. Ye [21] presented a discrete-time Geom/Geom/1 queue with single working vacation and multiple vacations. Gao [22]

considered an unreliable queuing system that adjusts maintenance time and limits idle period. Unlike the previously mentioned literature, the retrial queuing system studied in this paper takes into account the N-policy and multiple vacations. More specifically, this system has the following features: The system is activated only when the current vacation is completed and at least N customers are waiting in the system, otherwise, the server continues to the next vacation until the number of customers in the system is not less than N. Generally, in actual service systems, the server cannot immediately check the system status during the vacation state; or when the number of customers reaches the desired threshold, the server cannot immediately switch to the working state; or when the server ends the vacation, there are no (or few) customers who reach the system. At the same time, it is necessary to consider the operating cycle (lifetime) of the server for the overall benefit of the system. Therefore, the multiple vacations retrial queue with N-policy and breakdowns is considered in this paper; it has a wide range of applications in actual service systems.

There are the cooperative case and the non-cooperative case for the customers to join the system with each other. In the non-cooperative case, the goal of each individual is to obtain their own maximum expected net benefit based on the reward-cost structure. Obviously, if the server is unavailable upon arrival, the customers need to decide whether to join the system or leave forever. If the customer's expected net benefit is positive, then the customer will not hesitate to join the system, otherwise, the customer chooses to balk. If the expected net benefit equals zero, the customer is indifferent to join the orbit or balk. In the cooperative case, the social planner wants the customers to cooperate with each other in order to get the best social welfare for the whole system. Some existing research results ([7,13,19]) indicate that the optimal social joining probability does not exceed the individual's equilibrium joining probability, or the optimal social arrival rate does not exceed the individual's equilibrium arrival rate, i.e., the customers find the server unavailable upon their arrival, the customers in the non-cooperative case are more willing to join the system than the customers in the cooperative case. The result is that social benefits do not reach the expected maximum by the social planner in the non-cooperative case.

The proposed model has a potential application in an order-based production system. In order to avoid equipment wear and tear and unprofitable production, the factory's production lines will only be activated after signing the order quantity sufficient to cover all expenses. A production line is usually in four periods: Vacation, busy, idle, and repair. The vacation period is usually used to accumulate enough orders to cover all expenses, so the production line will not be activated when the order quantity does not accumulate to a certain amount. The idle period is used to prepare production materials for signing orders. The busy period is used to produce orders signed. In the process of production order, equipment may break down due to parts wear or service life. If the equipment breakdown occurs, the equipment is immediately sent for repair, the order which has interrupted service due to breakdown, continues to receive the remaining service after the equipment is repaired. If an idle production line's surplus material can make a new arrival order, the production line can directly complete the order. New orders that arrive during the busy period and the repair period can be joined the waiting list or abandon the factory contract. The new order is independent of other orders and is completed on the first-come, first-served (FCFS) discipline. In this scenario, orders, waiting list, the new order is independent of other orders and is completed on the FCFS principle, equipment breakdown, equipment repair time, the production line of the factory, and order quantity sufficient to cover all expenses corresponding to the customers, the orbit, retrial policy, server breakdown, repair time, the server, and N-policy, respectively,

in the queueing terminology. The characteristics and contributions of this model considered in this paper are as follows:

- This article considers a multiple vacations retrial queue with N-policy and breakdowns, it coincides with the actual service system.
- Two types of customer joining cases apply to this paper, i.e., non-cooperative customers aim to optimize individual interests, and the social planner in the cooperative case considers the profit of the whole service system.
- The equilibrium joining strategy for the non-cooperative case and the socially optimal joining strategy for the cooperative case are presented in this paper.
- The improved PSO algorithm is used to visualize the impact of system parameters on the profit of service providers.
- Both the model itself and the numerical experiment have certain guiding significance for the actual service system, this model is suitable for emergency disaster or epidemic control.

The rest of this paper is organized as follows. Section 2 gives a detailed description and parameter representation of this model. The steady-state distribution of the system and the mean orbit sizes are determined in Section 3. Section 4 obtains equilibrium joining strategies in the non-cooperative case and optimal joining strategies in the cooperative case. Section 5 focuses on the profit of the service provider and gives a simple description of PSO algorithms. Section 6 uses a large number of numerical experiments to show the influence of parameters on the two joining probabilities, social welfare, and the profit of the service provider. Section 7 presents the discussion and further study.

2. Model description

In this paper we consider a constant retrial queuing system with multiple vacations, N-policy, and breakdowns. Assumes the arrival of potential customers according to the Poisson process with rate λ. There is no waiting area in front of the server. Upon arrival, if the server is available (idle state), the arriving customer will immediately receive the service and leaves the system upon the completion of the service. Otherwise, if the server is unavailable (vacation state or busy state or repair state), the arriving customer would enter an infinite capacity orbit with probability q or leave the system forever with probability $\overline{q} = 1-q$, this virtual orbit is similar to a waiting list. The first customer in the orbit asks for service according to the first-come, first-served (FCFS) discipline with a Poisson flow of retrials of rate θ. The idle server will take an exponential time of rate θ to access (or search) the list customer to provide the service. However, if a new customer arrives during the search process, the server will abandon the search and immediately serves this new customer. The server may breakdown due to parts wear or service life during the service of customers. The life time of the server is exponentially distributed with rate α. If the server breakdown occurs, the server is immediately sent for repair, and the repair time follows an exponential distribution with rate γ. The customer who have interrupted service due to breakdown, continue to receive remaining service after the server is repaired. The service time for all (new or repeated) customers is exponentially distributed with the common rate μ, and all service times are independent.

On the other hand, the server in vacation state does not provide any services to arriving customers, and will not be activated until there are no less than N customers in the orbit after completing the current

vacation. If the server finds less than N customers in the orbit at the end of the current vacation, the server will start another new vacation until no less than N customers in the orbit. Once all customers are served, or the system becomes empty, the server starts a vacation of random length V, which is assumed to be exponentially distributed with rate ξ. So we can define this type of queue system as M/M/1/MV constant retrial queueing system with N-policy and breakdowns. Once the server is activated, all customers in the system receive the complete service, and the server turns to vacation state after serving all customers.

Let $S(t)$ be the state of the server at time t,

$$S(t) = \begin{cases} 0, & \text{vacation;} \\ 1, & \text{busy;} \\ 2, & \text{idle,} \\ 3, & \text{repair,} \end{cases}$$

and let $N(t)$ denotes the number of customers in the orbit at time t. Obviously, the stochastic process $\{(S(t), N(t)), t \geq 0\}$ is a continuous-time Markov chain with state space $\Omega = \{(0, i), i \geq 0; (1, j), j \geq 0, (2, k), k \geq 1; (3, h), h \geq 0\}$. The corresponding transition rate diagram of the Markov chain is shown in Figure 1.
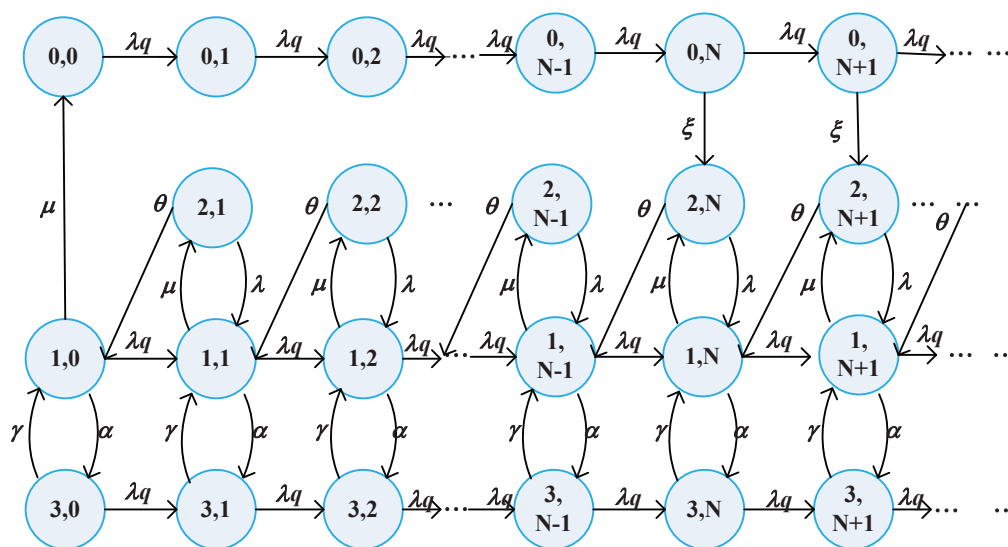

Figure 1. Transition rate diagram of the Markov chain $\{(S(t), N(t)), t \geq 0\}$.

Remark 1. *It is worthwhile to mention that if $q = 0$, then the model degenerates to a system with a single state $(0, 0)$. In this case, all customers will balk. Also, according to the above description, N should be at least equal to 1. Otherwise, the system will be the degenerated one too.*

We consider every customer has a reward-cost structure. When a customer arrives, the customer can decide whether to join the system according to the reward-cost structure. It is assumed that the customers are the same except that the arrival time is unexpected. With this structure, every customer received a reward of R units upon the completion of the service. However, a waiting cost of C per unit time is charged for the waiting in the orbit for each customer. In addition, the server imposes a price p,

announced to all incoming customers, for each customer to enter the system. Assume that customers are risk neutral and want to maximize their expected benefits. Upon arrival, the customer has to weigh the difference between gain and expenditures (waiting cost and announced price), the customers can evaluate the value and decide whether to enter the system or not. Obviously, the customer is more willing to join the system when all the expenses are less than the reward, while all the expenses are more than the reward, the customer does not enter the system. On the other hand, the customers are indifferent if the reward is equal to all expenses. In addition, in a sense, the customer's decision is immutable, which means that customers who join the system cannot exit, and customers who balked can not retry. So, this system can be regarded as a game between the customers. The server pre-announces the price of p in order to know how it affects the customer's behavior. For our analysis to be meaningful, we assume that

$$R > p, \tag{2.1}$$

which ensures that the customers alway join the system if they find that the server is idle upon their arrivals. We can refer to this literature [13] to get the necessary and sufficient stability condition of this system as follows

$$\frac{\lambda q(\lambda + \theta)(\alpha + \gamma)}{\theta\mu\gamma} < 1, \tag{2.2}$$

which is assumed throughout the paper.

3. Steady-state solutions

To study the joining strategies of two kindle of games, we first need to study the performance of the system. If the arriving customer chooses to join the system with probability q when finding the server is unavailable.

Let $\{p(0, i), i \geq 0; p(1, j), j \geq 0; p(2, k), k \geq 1; p(3, h), h \geq 0\}$ be the steady-state probability distribution of the Markov chain $\{(S(t), N(t)), t \geq 0\}$, and let $P_i(z)$, $i = 0, 1, 2, 3$ be the partial generating functions, defined by

$$P_0(z) = \sum_{n=0}^{\infty} z^n p(0, n), \quad P_1(z) = \sum_{n=0}^{\infty} z^n p(1, n), \quad P_2(z) = \sum_{n=1}^{\infty} z^n p(2, n), \quad P_3(z) = \sum_{n=0}^{\infty} z^n p(3, n), \quad |z| \leq 1.$$

Then, we have the following theorem.

Theorem 1. *For the M/M/1/MV constant retrial queueing system with N-policy and breakdowns in the steady state, the probabilities that the server is vacation, busy, idle and repair are given by, respectively,*

$$P_0(1) = \frac{\lambda q + N\xi}{\xi} p(0, 0), \tag{3.1}$$

$$P_1(1) = \frac{\lambda q\gamma(\lambda + \theta)(\lambda q + N\xi)}{\xi(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))} p(0, 0), \tag{3.2}$$

$$P_2(1) = \frac{\lambda q\mu\gamma(\lambda q + N\xi)}{\xi(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))} p(0, 0), \tag{3.3}$$

$$P_3(1) = \frac{\lambda q\alpha(\lambda + \theta)(\lambda q + N\xi)}{\xi(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))} p(0, 0), \tag{3.4}$$

where $p(0,0)$ can be obtained by the normalization condition $P_0(1) + P_1(1) + P_2(1) + P_3(1) = 1$, given by,

$$p(0,0) = \frac{\xi(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))}{\mu\gamma(\lambda q + \theta)(\lambda q + N\xi)}. \tag{3.5}$$

Proof From Figure 1, the balance equations of steady-state distribution are given as follows,

$$\lambda q p(0,0) = \mu p(1,0), \tag{3.6}$$

$$\lambda q p(0,n) = \lambda q p(0,n-1), \qquad 1 \leq n \leq N-1, \tag{3.7}$$

$$(\lambda q + \xi)p(0,n) = \lambda q p(0,n-1), \qquad n \geq N, \tag{3.8}$$

$$(\lambda q + \mu + \alpha)p(1,0) = \theta p(2,1) + \gamma p(3,0), \tag{3.9}$$

$$(\lambda q + \mu + \alpha)p(1,n) = \lambda q p(1,n-1) + \lambda p(2,n) + \theta p(2,n+1) + \gamma p(3,n), \qquad n \geq 1, \tag{3.10}$$

$$(\lambda + \theta)p(2,n) = \mu p(1,n), \qquad 1 \leq n \leq N-1, \tag{3.11}$$

$$(\lambda + \theta)p(2,n) = \xi p(0,n) + \mu p(1,n), \qquad n \geq N \tag{3.12}$$

$$(\lambda q + \gamma)p(3,0) = \alpha p(1,0), \tag{3.13}$$

$$(\lambda q + \gamma)p(3,n) = \alpha p(1,n) + \lambda q p(3,n-1), \qquad n \geq 1. \tag{3.14}$$

From (3.7), we can get

$$p(0,N-1) = \cdots = p(0,1) = p(0,0), \tag{3.15}$$

which leads to

$$\sum_{n=0}^{N-1} z^n p(0,n) = \sum_{n=0}^{N-1} z^n p(0,0) = \frac{1-z^N}{1-z}p(0,0). \tag{3.16}$$

By (3.8) and (3.12), we get

$$p(0,n) = p(0,N-1)\left(\frac{\lambda q}{\lambda q + \xi}\right)^{n-N+1} = p(0,0)\left(\frac{\lambda q}{\lambda q + \xi}\right)^{n-N+1}, n \geq N. \tag{3.17}$$

So,

$$\sum_{n=N}^{\infty} z^n p(0,n) = \sum_{n=N}^{\infty} z^n p(0,0)\left(\frac{\lambda q}{\lambda q + \xi}\right)^{n-N+1} = \frac{\lambda q z^N}{\lambda q(1-z) + \xi}p(0,0). \tag{3.18}$$

Therefore, $P_0(z)$ can be obtained as follows,

$$P_0(z) = \sum_{n=0}^{\infty} z^n p(0,n) = \sum_{n=0}^{N-1} z^n p(0,n) + \sum_{n=N}^{\infty} z^n p(0,n) = \frac{\lambda q(1-z) + \xi(1-z^N)}{(1-z)(\lambda q(1-z) + \xi)}p(0,0). \tag{3.19}$$

Then we consider the partial generating function $P_1(z)$. Multiplying equations (3.9)–(3.14) by z^n and summing up over n, we have that

$$(\lambda q + \mu + \alpha - \lambda q z)P_1(z) = (\lambda + \frac{\theta}{z})P_2(z) + \gamma P_3(z), \tag{3.20}$$

$$(\lambda + \theta)P_2(z) = \mu P_1(z) - \mu p(1,0) + \xi \sum_{n=N}^{\infty} z^n p(0,n), \tag{3.21}$$

$$(\lambda q + \gamma - \lambda qz)P_3(z) = \alpha P_1(z). \tag{3.22}$$

From (3.22), we have

$$P_3(z) = \frac{\alpha}{\lambda q(1 - z) + \gamma} P_1(z). \tag{3.23}$$

Substituting (3.23) into (3.20) yields

$$P_2(z) = \frac{\lambda^2 q^2(1 - z)^2 + \lambda q(1 - z)(\alpha + \mu + \gamma) + \mu\gamma}{(\lambda q(1 - z) + \gamma)(\lambda + \frac{\theta}{z})} P_1(z). \tag{3.24}$$

Substituting (3.6) and (3.18) into (3.21) yields

$$(\lambda + \theta)P_2(z) = \mu P_1(z) - \lambda q p(0, 0) + \frac{\lambda q \xi z^N}{\lambda q(1 - z) + \xi} p(0, 0). \tag{3.25}$$

So, substituting (3.24) into (3.25) yields $P_1(z)$ as follows

$$P_1(z) = \frac{\lambda q(\xi z^N - \lambda q(1 - z) - \xi)(\lambda q(1 - z) + \gamma)(\lambda + \frac{\theta}{z})}{(\lambda q(1 - z) + \xi) \cdot A} p(0, 0), \tag{3.26}$$

where

$$A = (\lambda + \theta)(\lambda^2 q^2(1 - z)^2 + \lambda q(1 - z)(\alpha + \mu + \gamma) + \mu\gamma) - \mu(\lambda q(1 - z) + \gamma)(\lambda + \frac{\theta}{z}). \tag{3.27}$$

Substituting (3.26) into (3.24) yields $P_2(z)$ as follows

$$P_2(z) = \frac{\lambda q(\xi z^N - \lambda q(1 - z) - \xi)(\lambda^2 q^2(1 - z)^2 + \lambda q(1 - z)(\alpha + \mu + \gamma) + \mu\gamma)}{(\lambda q(1 - z) + \xi) \cdot A} p(0, 0). \tag{3.28}$$

Substituting (3.26) into (3.23) yields $P_3(z)$ as follows

$$P_3(z) = \frac{\lambda q \alpha(\xi z^N - \lambda q(1 - z) - \xi)(\lambda + \frac{\theta}{z})}{(\lambda q(1 - z) + \xi) \cdot A} p(0, 0). \tag{3.29}$$

Let $z = 1$ in (3.19), (3.26), (3.28) and (3.29), we can get the results (3.1)–(3.4), where $p(0, 0)$ can be obtained by the normalization condition $P_0(1) + P_1(1) + P_2(1) + P_3(1) = 1$, that is (3.5). □

Let M_i be the mean orbit sizes when the server in state i ($i = 0, 1, 2, 3$). From (3.19), (3.26), (3.28) and (3.29), by using $M_i = \frac{d}{dz}P_i(z)|_{z=1}$, and after some algebraic manipulations, we have the following Theorem 2.

Theorem 2. *For the M/M/1/MV constant retrial queueing system with N-policy and breakdowns in the steady state, the mean orbit sizes when the server is on vacation, busy, idle and repair, respectively, are given by*

$$M_0 = \left(\frac{N(N - 1)}{2} + \frac{\lambda q(\lambda q + N\xi)}{\xi^2} \right) p(0, 0), \tag{3.30}$$

$$M_1 = \lambda q p(0, 0) \left(\frac{N(N - 1)\gamma(\lambda + \theta)}{2(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))} \right.$$

$$+ \frac{\gamma(\lambda q(\lambda + \theta) + \lambda \xi)(\lambda q + N\xi)}{(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))\xi^2} + \frac{\lambda q(\gamma(\alpha + \gamma) + \lambda q\alpha)(\lambda + \theta)^2(\lambda q + N\xi)}{(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))^2\xi} \Bigg), \quad (3.31)$$

$$M_2 = \lambda q p(0,0)\Bigg(\frac{\mu\gamma N(N-1)}{2(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))} + \frac{\mu\gamma(\lambda q + \xi)(\lambda q + N\xi)}{(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))\xi^2}$$
$$+ \frac{\lambda q(\lambda q + N\xi)(\lambda q(\lambda + \theta)(\alpha\mu + (\alpha + \gamma)^2) + \lambda\mu\gamma(\alpha + \gamma))}{(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))^2\xi} \Bigg), \quad (3.32)$$

$$M_3 = \lambda q\alpha p(0,0)\Bigg(\frac{N(N-1)(\lambda + \theta)}{2(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))} + \frac{(\lambda q(\lambda + \theta) + \lambda\xi)(\lambda q + N\xi)}{(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))\xi^2}$$
$$+ \frac{\lambda q\theta\mu(\lambda + \theta)(\lambda q + N\xi)}{(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))^2\xi} + \frac{\lambda q(\alpha + \gamma - \lambda q)(\lambda + \theta)^2(\lambda q + N\xi)}{(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))^2\xi} \Bigg), \quad (3.33)$$

where $p(0,0)$ is given in (3.5).

4. Joining strategies in two kinds of games and pricing analysis

If the server is unavailable upon arrival, each customer needs to decide whether to join the system based on the reward-cost structure. If the arriving customer is against the will of the social planner, maximizes the individual's expected net benefit, rather than maximizing the overall benefit from the perspective of the social planner, then the customers are in the non-cooperative case. However, the social planner often wants to encourage customers to cooperate with each other to maximize the social welfare in cooperative case. In this section, we consider the joining strategies of the customers in two kinds of game cases, respectively.

4.1. Equilibrium joining strategy in the non-cooperative case

In the non-cooperative case, the arriving customer is based on a reward-cost structure with the goal of maximizing the individual benefit, when an arriving customer finds that the server is unavailable, he needs to decide whether to enter the system or not. Actually, if the expected net benefit is greater than zero, the customers are more willing to join the orbit; the customers prefer to balk if expected net benefit is less than zero; if the customers' expected net benefit equals zero, they are indifferent to join the orbit or balk. Each customer's different decisions have different impacts on other customers, and it also affects the performance measures of the system. Next, our goal is to explore the symmetric equilibrium joining strategy of the non-cooperative game between these customers.

The expected waiting time of a tagged customer who finds the server is unavailable and decides to join the orbit upon his arrival, denoted by $T(q)$. Then we can give the following Theorem 3 about $T(q)$.

Theorem 3. *For the M/M/1/MV constant retrial queueing system with N-policy and breakdowns in the steady state, the expected waiting time of a tagged customer in the orbit, given that he finds the server is unavailable and decide to join orbit upon his arrival, $T(q)$ is given by*

$$T(q) = \frac{\xi N(N-1)(\lambda q + \theta)}{2\lambda q\theta(\lambda q + N\xi)} + \frac{\mu\gamma(\lambda q + \theta + \xi) + \lambda\xi(\alpha + \gamma)}{\mu\gamma\theta\xi} + \frac{\Phi(q)}{\theta\mu\gamma(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))}, \quad (4.1)$$

where

$$\Phi(q) = \lambda^2 q^2 (\lambda + \theta)(\alpha\mu + (\alpha + \gamma)^2) + \lambda q (\alpha + \gamma)^2 (\lambda + \theta)^2 + \lambda q \mu(\lambda\gamma(\alpha + \gamma) + \alpha\theta(\lambda + \theta)). \quad (4.2)$$

Proof According to the model description, it is only available when the server is idle upon customers arrivals. Therefore, according to the Poisson arrivals see time average (PASTA) property and Theorem 1, the server unavailable probability P_{un} upon arrival of a customer is given by

$$P_{un} = P_0(1) + P_1(1) + P_3(1) = \frac{\theta}{(\lambda q + \theta)}. \quad (4.3)$$

P_{un} is the probability of the server unavailable upon arrival of a customer. So, the total arrival rate of the customers in the orbit can be determined as follows

$$\lambda_{un} = \lambda q P_{un} = \frac{\lambda q \theta}{(\lambda q + \theta)}. \quad (4.4)$$

From Theorem 2, the customers' mean number in the orbit, denoted by M, then

$$M = M_0 + M_1 + M_2 + M_3. \quad (4.5)$$

By using Little' formula, we can get

$$T(q) = \frac{M}{\lambda_{un}}. \quad (4.6)$$

Hence, substituting (4.4) and (4.5) into (4.6), (4.1) can be obtained. □

Now, we obtain the expected waiting time $T(q)$ of a tagged customer who finds the server is unavailable and decides to join orbit upon his arrival. Based on the reward-cost structure previously mentioned, the expected net benefit of a tagged customer, who finds that the server is unavailable upon arrival and decides to join the orbit, denoted by $S_e(q)$, it can be obtained as follows

$$S_e(q) = R - p - CT(q), \quad (4.7)$$

where $T(q)$ is given in (4.1), and the second-order derivative of $T(q)$ in q, given by,

$$T''(q) = \frac{N(N-1)\xi(\lambda^2 q^2(\lambda q + 3\theta) + 3N\lambda q \theta\xi + N^2\theta\xi^2)}{\theta\lambda q^3(\lambda q + N\xi)^3}$$
$$+ \frac{2\lambda^2(\lambda + \theta)((\lambda + \theta)^2(\alpha + \gamma)^3 + (\lambda + \theta)(\alpha + \gamma)(\mu\gamma(\alpha + \gamma) + \theta\mu\gamma) + \theta\mu^2\alpha\gamma)}{(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))^3}. \quad (4.8)$$

Obviously, $T''(q)$ is positive under the condition that the system is stable. So $T(q)$ is strictly convex in the interval $0 \le q < \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)}$. Therefore, there is a unique global minimal point $q_{min} \in [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)})$, which makes $T(q)$ achieve a global minimum in this interval. It follows from (4.7) that $S_e(q)$ has a unique global maximal value in this range of q ($q \in [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)})$).

Remark 2. *Because $T(q)$ is strictly convex under the stable condition, i.e., $q \in [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)})$, and q should satisfy $0 \le q \le 1$. Defined by the set Q as follows,*

$$Q = \begin{cases} Q_1 = [0, 1], & \text{if } \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)} > 1; \\ Q_2 = [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)}), & \text{if } \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)} \le 1. \end{cases} \quad (4.9)$$

So in the following theorems, we will only study the range of $q \in Q$.

We are ready to describe the equilibrium joining strategies for the non-cooperative case.

Theorem 4. *For the M/M/1/MV constant retrial queueing system with N-policy and breakdowns in the steady state, for any given price $p < R$, there exists an Nash equilibrium mixed joining strategy, i.e., the arriving customers joining the orbit of the unavailable system with probability q_e, where q_e have the following cases:*

Case I: $\frac{R-p}{C} < T(q_{min})$. *In this case, a unique equilibrium strategy q_e exists: $q_e = 0$.*

Case II: $\frac{R-p}{C} = T(q_{min})$. *In this case, the equilibrium strategy is given by*

    Case II (a): $q \in Q_1$ and $q_{min} > 1$. *Then, a unique equilibrium strategy q_e exists: $q_e = 0$.*

    Case II (b): $q \in Q_1$ and $q_{min} \leq 1$. *Then, two equilibrium strategies q_e exist: $q_e = 0$ and $q_e = q_{min}$.*

    Case II (c): $q \in Q_2$. *Then, two equilibrium strategies q_e exist: $q_e = 0$ and $q_e = q_{min}$.*

Case III: $\frac{R-p}{C} > T(q_{min})$. *Let $q_1 < q_2$ be the roots of $S_e(q) = 0$ (there are two roots because of the strict concavity of $S_e(q)$ in $q \in Q_2$).*

    Case III (a): $q \in Q_1$ and $q_1 > 1$. *Then, a unique equilibrium strategy q_e exists: $q_e = 0$.*

    Case III (b): $q \in Q_1$ and $q_1 = 1$. *Then, two equilibrium strategies q_e exist: $q_e = 0$ and $q_e = 1$.*

    Case III (c): $q \in Q_1$ and $q_1 < 1$. *Then, three equilibrium strategies q_e exist: $q_e = 0$, $q_e = q_1$ and $q_e = min\,(q_2, 1)$.*

    Case III (d): $q \in Q_2$. *Then, three equilibrium strategies q_e exist: $q_e = 0$, $q_e = q_1$ and $q_e = q_2$.*

    Proof   According to the model assumption, the arriving customer encounters the idle state, who accepts the service directly, because the customer who arrives at the idle state does not generate waiting time. However, when a customer arrives and finds that the system is not available (vacation or busy or repair), he needs to decide whether to join the system or not. We have proved the strict concavity of $S_e(q)$ of $q \in [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)})$ in the previously, and let q_{min} ($q_{min} \in [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)})$) be the unique point that globally minimizes $T(q)$ in $q \in [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)})$, i.e., q_{min} be the unique point that globally maximizes $S_e(q)$ in $q \in [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)})$. Obviously, the strategy $q_e = 0$ of balking is always an equilibrium strategy. The reason is that all customers balk, a tagged customer also prefers to balk; otherwise this tagged customer will never get service. In fact, if $q = 0$, the model degenerates to a system with a single state of $(0, 0)$, then the system will never be activated. Specifically, we have the following results:

(I) When $\frac{R-p}{C} < T(q_{min})$, i.e., $S_e(q_{min}) < 0$, which implies $S_e(q)$ is negative for every $q \in Q$. In this case, the best response for the tagged customer is balking, i.e., $q_e = 0$. Thus we have the Case I.

(II) When $\frac{R-p}{C} = T(q_{min})$, i.e., $S_e(q_{min}) = 0$, which implies $S_e(q)$ is negative for every $q \neq q_{min}$. There are three cases, (a): if $q \in Q_1$ and $q_{min} > 1$, which implies $q_{min} \notin Q$, the equilibrium strategy is $q_e = 0$ of balking because of $S_e(q) < 0$ for every q in this subcase. (b): if $q \in Q_1$ and $q_{min} \leq 1$, q_{min} is the unique solution of the equation $S_e(q) = 0$ in this subcase, i.e., q_{min} is a proper mixed strategy $q \in Q$. (c): if $q \in Q_2$, according to the previous assumption, q_{min} must belong to the interval Q_2, so q_{min} is the unique solution of the equation $S_e(q) = 0$ in this subcase, i.e., q_{min} is a proper mixed strategy $q \in Q$. Thus we have the Case II.

(III) When $\frac{R-p}{C} > T(q_{min})$, i.e., $S_e(q_{min}) > 0$. We notice that the two limits $\lim_{q \to 0^+} S_e(q) = -\infty$, $\lim_{q \to (\theta\mu\gamma/\lambda(\lambda+\theta)(\alpha+\gamma))^-} S_e(q) = -\infty$ and $S_e(q)$ is strict concavity in $q \in [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)})$, so the equation $S_e(q) = 0$ must have two roots q_1, q_2 $(q_1 < q_2) \in [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)})$. According to the definition of set Q, there are four cases. (a): if $q \in Q_1$ and $q_1 > 1$, which implies that the equation $S_e(q) = 0$ has no solution in the interval Q_1, so the equilibrium strategy is $q_e = 0$ of balking. (b): if $q \in Q_1$ and $q_1 = 1$, $q = 1$ is the unique solution of the equation $S_e(q) = 0$, so $q_e = 1$ is the equilibrium strategy in this subcase. (c): if $q \in Q_1$ and $q_1 < 1$, similar to the discussion of Case III (b), q_1 is a mixed equilibrium strategy in this subcase because of $S_e(q_1) = 0$. Obviously, if $q_2 < 1$, $q_e = q_2$ is also a mixed equilibrium strategy in this case; if $q_2 \geq 1$, q_2 is not the solution of the equation $S_e(q) = 0$, but $q = 1$ is a equilibrium strategy in this subcase because of $S_e(1) > 0$. (d): if $q \in Q_2$, according to the previous description, q_1 and q_2 must belong to Q_2, so q_1 and q_2 are mixed equilibrium strategies because of $S_e(q_1) = 0$ and $S_e(q_2) = 0$. Thus we have the Case III. □

4.2. Socially optimal joining strategy in the cooperative case

The customers in the non-cooperative case gain the individual's best benefit by selfish means. However, the joining strategy of the non-cooperative case does not match the wishes of social planners. The intention of the social planner is to promote mutual cooperation between customers in order to maximize the social benefits, i.e., the cooperative case. In this subsection, we turn our interest to the analysis of the socially optimal joining strategy in the cooperative case, in which, the social planner maximizes the social welfare through adopting the socially optimal joining probability q^*.

For a given pricing p and joining probability q, the customers join the orbit with probability q when they find the server unavailable upon their arrivals. Social welfare (social net benefit) is the sum of the welfare of customers and server, the benefit of the customers per unit time, denoted by $C(q)$, it is as follows

$$C(q) = \lambda q(P_0(1) + P_1(1) + P_3(1))(R - p - CT(q)) + \lambda P_2(1)(R - p), \qquad (4.10)$$

and the benefit of the server per unit time, denoted by $S(q)$, it is as follows

$$S(q) = (\lambda q(P_0(1) + P_1(1) + P_3(1)) + \lambda P_2(1))p. \qquad (4.11)$$

So, social welfare is the sum of the welfare for all customers and also the server, denoted by $SW(q)$, it is as follows

$$
\begin{aligned}
SW(q) &= C(q) + S(q) \\
&= (\lambda q(P_0(1) + P_1(1) + P_3(1)) + \lambda P_2(1))R - C\lambda q(P_0(1) + P_1(1) + P_3(1))T(q) \\
&= \lambda^* R - C\lambda_{un} T(q), \qquad (4.12)
\end{aligned}
$$

where

$$\lambda^* = \lambda q(P_0(1) + P_1(1) + P_3(1)) + \lambda P_2(1) \qquad (4.13)$$

is the total arrival rate of the system, and

$$\lambda_{un} = \lambda q(P_0(1) + P_1(1) + P_3(1)) = \frac{\lambda q\theta}{\lambda q + \theta} \qquad (4.14)$$

is the total arrival rate to the system when the server is unavailable. By using (4.1), (4.13) and (4.14), we can obtain that

$$SW(q) = \frac{\lambda q(\lambda + \theta)}{\lambda q + \theta}R - C\left(\frac{\xi N(N-1)}{2(\lambda q + N\xi)} + \frac{\lambda q\theta\mu\gamma(\lambda q + \theta + \xi) + \lambda q\theta\lambda\xi(\alpha + \gamma)}{\theta\mu\gamma\xi(\lambda q + \theta)}\right.$$
$$\left. + \frac{\lambda q\theta\Phi(q)}{\theta\mu\gamma(\lambda q + \theta)(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))}\right), \tag{4.15}$$

where $\Phi(q)$ is given in (4.2). Further, the second-order derivative of $SW(q)$ in q as follows,

$$SW''(q) = -\frac{2\lambda^2\theta((\lambda + \theta)R - C)}{(\lambda q + \theta)^3} - C\left(\frac{\lambda^2\xi N(N-1)}{(\lambda q + N\xi)^3} + \frac{2\lambda^2\theta\mu\gamma(\lambda + \theta)((\lambda + \theta)(\alpha + \gamma)^2 + \theta\mu\alpha)}{(\theta\mu\gamma - \lambda q(\lambda + \theta)(\alpha + \gamma))^3}\right). \tag{4.16}$$

Obviously, under the stability condition $0 \le q < \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)}$ (it is also set Q_2), if $(\lambda + \theta)R - C > 0$, i.e., $\frac{R}{C} > \frac{1}{\lambda+\theta}$, then $SW''(q)$ is negative, and it can be further obtained that $SW(q)$ is strictly concave with respect to q in this range (i.e. Q_2) and its maximum q_{max} is unique. From Remark 2, we need to discuss whether q belongs to the set Q_1 or set Q_2. Specifically, the socially optimal joining strategy for the cooperative case is given in Theorem 5.

Remark 3. *If $(\lambda + \theta)R - C < 0$, i.e., $\frac{R}{C} < \frac{1}{\lambda+\theta}$, then $SW''(q)$ is positive under the stability condition $0 \le q < \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)}$, it is difficult to get the maximum value of $SW(q)$ under the stable condition. So in the following theorem, we will only study the range of $(\lambda + \theta)R - C > 0$, i.e., $\frac{R}{C} > \frac{1}{\lambda+\theta}$.*

Theorem 5. *For the M/M/1/MV constant retrial queueing system with N-policy and breakdowns in the steady state. For $\frac{R}{C} > \frac{1}{\lambda+\theta}$ and any given price $p < R$, the number and the type of socially optimal joining strategy for cooperative case depend on the values of q_{max}, and the set Q. Specifically, the customers join the orbit with probability q^* upon their arrival by seeing an unavailable server, the joining probability q^* maximizes social welfare, where q^* is given by:*

Case I: $\frac{R}{C} < \frac{\theta}{\lambda+\theta}T(q_{max})$. *In this case, the socially optimal strategy is $q^* = 0$.*

Case II: $\frac{R}{C} = \frac{\theta}{\lambda+\theta}T(q_{max})$.

    Case II (a): $q \in Q_1$ *and* $q_{max} > 1$. *Then, the socially optimal strategy is $q^* = 0$.*

    Case II (b): $q \in Q_1$ *and* $q_{max} = 1$. *Then, the socially optimal strategy is $q^* = 1$.*

    Case II (c): $q \in Q_1$ *and* $q_{max} < 1$. *Then, the socially optimal strategy is $q^* = q_{max}$.*

    Case II (d): $q \in Q_2$. *Then, the socially optimal strategy is $q^* = q_{max}$.*

Case III: $\frac{R}{C} > \frac{\theta}{\lambda+\theta}T(q_{max})$. *Then, the socially optimal strategy is $q^* = min(q_{max}, 1)$.*

    Proof According to the previous description, under the stability condition ($0 \le q < \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)}$), if $(\lambda + \theta)R - C > 0$, $SW''(q)$ is strictly concave for $q \in [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)})$ and its maximum q_{max} is unique in this interval (i.e. Q_2). According to the reward-cost structure, we have the following results:

(I) When $\frac{R}{C} < \frac{\theta}{\lambda+\theta}T(q_{max})$, which implies $SW(q_{max}) < 0$, then $SW(q)$ is negative for every q ($q \in Q$). So, it has no positive social welfare for any q. In this case, the socially optimal strategy $q^* = 0$ (balking). Thus we have the Case I.

(II) When $\frac{R}{C} = \frac{\theta}{\lambda+\theta}T(q_{max})$, which implies $SW(q_{max}) = 0$, then $SW(q) \leq 0$ of $q \in Q$. So, $SW(q)$ is negative for each $q \neq q_{max}$. There are four cases, (a): if $q \in Q_1$ and $q_{max} > 1$, which implies $SW(q) < 0$ for every $q \in Q$ because of $q_{max} \notin Q$, i.e., there is no point such that $SW(q) = 0$. In this subcase, the socially optimal strategy $q^* = 0$ (balking). (b): if $q \in Q_1$ and $q_{max} = 1$, since $SW(1) = 0$, we have that the maximum value of $SW(q)$ in $q \in Q = Q_1$ is 0, so the socially optimal strategy is attained $q^* = 1$. (c): if $q \in Q_1$ and $q_{max} < 1$, similar to the discussion of Case II (b), the socially optimal strategy is attained $q^* = q_{max}$. (d): if $q \in Q_2$, since q_{max} must belong to the set Q_2, and $SW(q_{max}) = 0$, so a unique socially optimal strategy is attained at $q^* = q_{max}$, there is no other point such that $SW(q) = 0$ of $q \in Q = Q_2$. Thus we have the Case II.

(III) When $\frac{R}{C} > \frac{\theta}{\lambda+\theta}T(q_{max})$, which implies $SW(q_{max}) > 0$, i.e. the maximum value of $SW(q)$ in $q \in [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)}))$ is attained at some point. Since the strict concavity of $SW(q)$ as a function of $q \in [0, \frac{\theta\mu\gamma}{\lambda(\lambda+\theta)(\alpha+\gamma)}))$, we can get the maximum value q_{max} is unique. When $q \leq q_{max}$, $SW(q)$ is strictly increasing; while $SW(q)$ is strictly decreasing for $q \geq q_{max}$. So the maximum value of $SW(q)$ is obtained at $q^* = q_{max}$, if $q_{max} \leq 1$, otherwise the maximum value of $SW(q)$ is obtained at $q^* = 1$. Thus we have the Case III. □

5. Optimal pricing of the service provider

The problem of maximizing the profit of the server is also considered in the actual service system. As mentioned before, all customers who join the system are charged an admission fee p (announced price, i.e., service price) by the server. On the other hand, running the server will generate consumption costs. In fact, the profit of the server is dynamic, and the server can maximize the profit by adjusting the admission fee p and the service rate μ.

Assume that the server has no consumption cost in the vacation state and repair state. However, the consumption costs of the server in idle and busy states are C_i and C_b, respectively. Then, the profit of the service provider per time unit under equilibrium is given by

$$S(\mu, p) = (\lambda q_e(P_0(1) + P_1(1) + P_3(1)) + \lambda P_2(1))p - (C_b P_1(1) + C_i P_2(1)), \tag{5.1}$$

where q_e is the equilibrium joining probability, which is given in Theorem 4, $P_0(1)$, $P_1(1)$, $P_2(1)$ and $P_3(1)$ are given in Theorem 1.

The expression $S(\mu, p)$ of the profit of the service provider per time unit under equilibrium is complex, so it is usually difficult to obtain analytical characterization. Therefore, we can give some numerical analysis by Particle Swarm Optimization (PSO) algorithm to overcome the difficulty of obtaining analytical characterization.

PSO algorithm is a global random search algorithm based on swarm intelligence proposed by [23], inspired by the results of artificial life research, by simulating the migration and swarming behavior of birds swarm foraging. The basic idea of the PSO algorithm comes from mutual cooperation and information sharing between groups to obtain the optimal solution. Because of its simple operation and fast convergence speed, the PSO algorithm has been widely used in many fields such as function optimization, image processing, and geodesy.

The PSO algorithm is first initialized as a group of random particles (random solution), and then the optimal solution is found through multiple iterations. In each iteration, the particles continuously

update the optimal values of velocity V_{id} and position X_{id} and finally obtain the global optimum through multiple iterations, where velocity V_{id} represents the speed of movement, position X_{id} represents the direction of movement, the number of particles represents the dimension of the solutions. The position of each particle represents a potential optimal solution. V_{id} and X_{id} are updated by the following formulas,

$$V_{id} = \omega * V_{id} + c_1 * rand() * (pBest_i - X_{id}) + c_2 * rand() * (gBest_i - X_{id}), \tag{5.2}$$

and

$$X_{id} = X_{id} + V_{id}, \tag{5.3}$$

where i is the number of particles, d is dimension, $rand()$ is a random number in $(0, 1)$, c_1 and c_2 are learning factor and ω is inertia factor.

6. Numerical examples

In this section, the theoretical knowledge we have obtained based on the previous research contains many parameters, so we explore the impact of these parameters on the two joining strategies (q_e and q^*) and social welfare $SW(q)$ through a series of numerical experiments. On the other hand, due to the complexity of the expression $S(\mu, p)$ of the profit of the service provider per time unit under equilibrium, we numerically analyze how the service provider sets up the price p and service rate μ to maximize its own benefit. To address these issues, through a large number of numerical experiments on various parameters, regardless of the choice of the parameters, we find that the qualitative results of the numerical experiments are similar. So we present some typical numerical scenarios to illustrate the findings of these numerical experiments. The first set of numerical experiments presents the trends of the two joining strategies (q_e and q^*) on the parameters λ, μ, θ, ξ, α, γ and N, respectively, in Figures 2–5, respectively.

6.1. The impact of parameters on the two joining strategies q_e and q^*

Figure 2 (a) shows that q_e and q^* have a downward trend with arrival rate λ, and q_e is always above q^*. The reason is that as λ increases, more and more customers join the system, and the orbit becomes more crowded, the waiting cost of the later arriving customers is increasing, which will lead to a gradual decline in the enthusiasm of customers to join the system. However, as the service rate μ increases, q_e and q^* show an increasing trend as shown in Figure 2 (b). This is because as the service rate μ increases, the customers on the orbit get fast service, and waiting costs will be reduced. Therefore, no matter the cooperation case or non-cooperation case, the arriving customers are more willing to join the unavailable system.
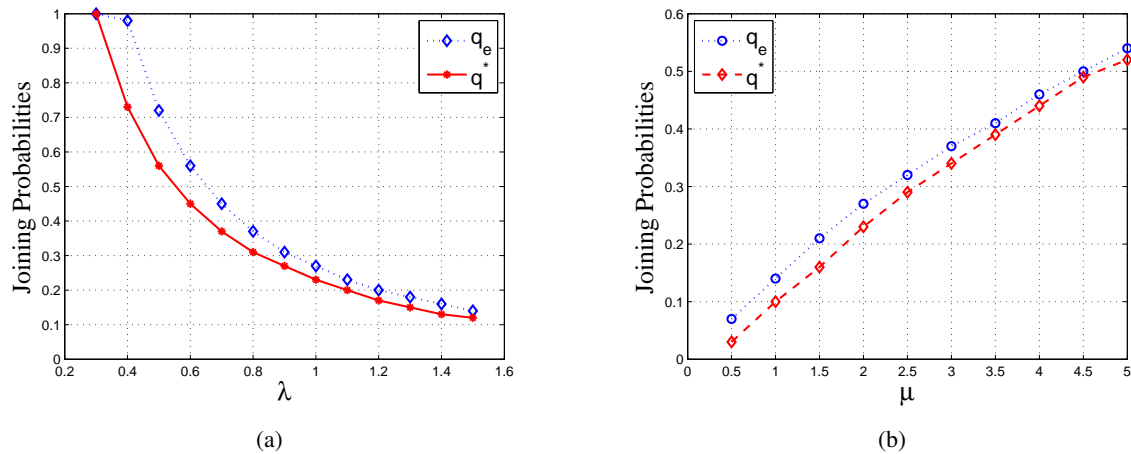
Figure 2. (a) q_e and q^* vs. λ when $R = 5$, $p = 1$, $C = 1$, $\mu = 2$, $\xi = 1$, $N = 5$, $\theta = 0.6$, $\alpha = 0.3$, $\gamma = 0.8$. (b) q_e and q^* vs. μ when $R = 5$, $p = 1$, $C = 1$, $\lambda = 1$, $\xi = 1$, $N = 5$, $\theta = 0.6$, $\alpha = 0.3$, $\gamma = 0.8$.

Figure 3 (a) indicates that q_e and q^* increase as ξ increases, but the growth trend is not very obvious. Since as the vacation time becomes shorter, the number of customers accumulating on the orbit is reduced, so the arriving customers are more likely to join the system when it is unavailable. Figure 3 (b) shows that both q_e and q^* increase with the retrial rate θ. The reason is that as the retrial rate continues to increase, the customers accumulating in the orbit get fast service, which leads to the newly arriving customers who are more willing to join the unavailable system.
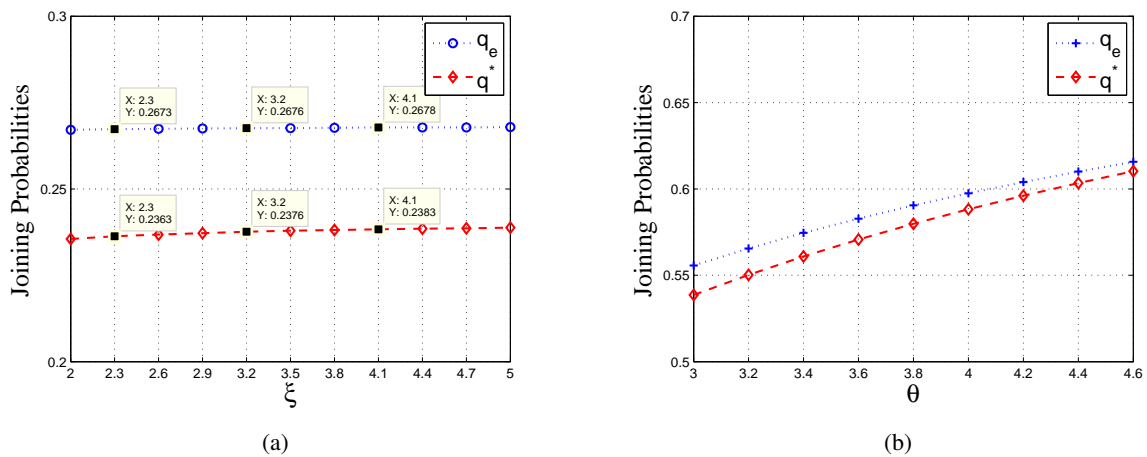
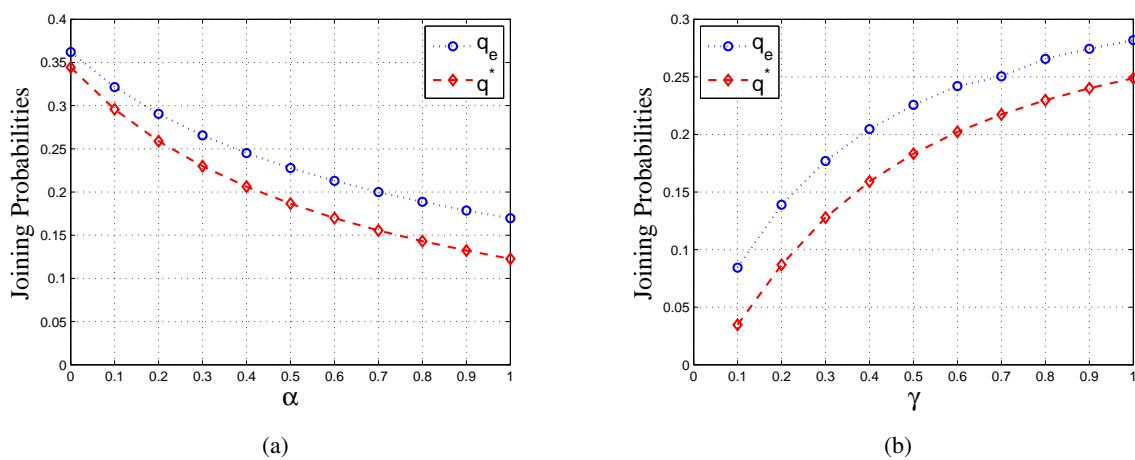**Figure 3.** (a) q_e and q^* vs. ξ when $R = 5$, $p = 1$, $C = 1$, $\lambda = 1$ $\mu = 2$, $N = 5$, $\theta = 0.6$, $\alpha = 0.3$, $\gamma = 0.8$. (b) q_e and q^* vs. θ when $R = 5$, $p = 1$, $C = 1$, $\lambda = 1$, $\mu = 2$, $\xi = 1$, $N = 5$, $\alpha = 0.3$, $\gamma = 0.8$.

Obviously, the life cycle of the server effectively affects the joining probability of the arriving customers in the cooperation case and non-cooperation case. In Figure 4 (a), q_e and q^* show a downward trend with the increase of α. This is because the shortened lifetime of the server leads

to an increase in the number of server repairs within a certain period, which will increase the waiting cost of customers in the orbit. So the newly arriving customers are more reluctant to join the system when they find it unavailable. However, the shortening of the server maintenance time, i.e., increase γ, it will effectively increase the joining probability of the customers in the two cases, which can be shown in Figure 4 (b). This is exactly the opposite of Figure 4 (a) because the increase in γ reduces the waiting costs of the customers in the orbit.

Figure 4. (a) q_e and q^* vs. α when $R = 5$, $p = 1$, $C = 1$, $\lambda = 1$ $\mu = 2$, $\xi = 1$, $N = 5$, $\theta = 0.6$, $\gamma = 0.8$. (b) q_e and q^* vs. γ when $R = 5$, $p = 1$, $C = 1$, $\lambda = 1$, $\mu = 2$, $\xi = 1$, $N = 5$, $\theta = 0.6$, $\alpha = 0.3$.

Figure 5 (a) indicates that q_e and q^* are inconsistent as functions of the threshold N. This is because as N increases, the enthusiasm of selfish customers for joining the system gradually decreases, i.e., q_e decreases with N. However, as N increases, accumulating more customers on the waiting list is completely coincide with the wishes of the social planner. Figure 5 (b) shows that the increase in announcement price p leads to a gradual decline in the enthusiasm of selfish customers, i.e., the customers in non-cooperative case, but p has no effect on the customers in the cooperation case. Obviously, as the service price p increases, the probability q_e of the arriving customers (the customers in non-cooperative case) to join the unavailable system must be reduced. However, as discussed in Theorem 5, since the service price p is only a transfer relationship between the customers and the server or the service provider, so q^* remains unchanged.
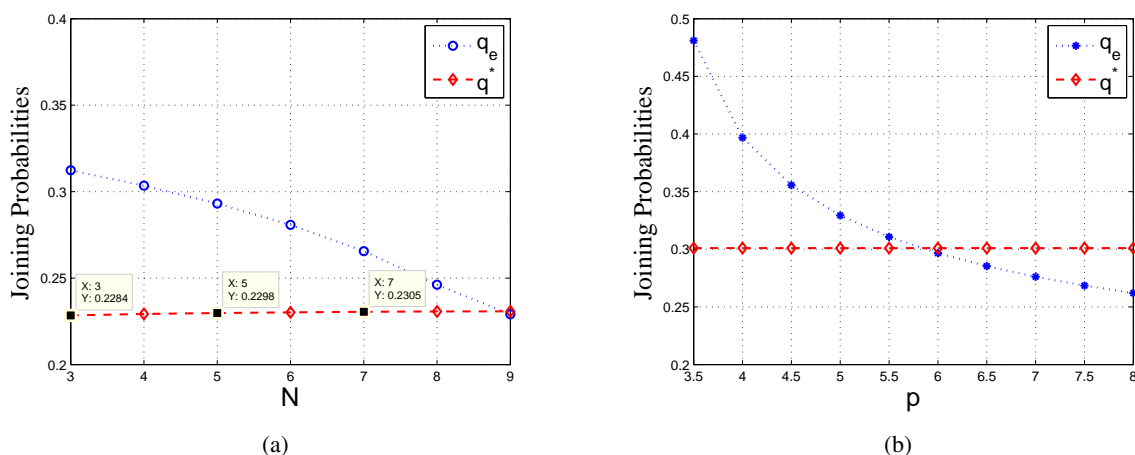
(a)

(b)

Figure 5. (a) q_e and q^* vs. N when $R = 5$, $p = 1$, $C = 1$, $\lambda = 1$ $\mu = 2$, $\xi = 1$, $\theta = 0.6$, $\alpha = 0.3$, $\gamma = 0.8$. (b) q_e and q^* vs. p when $R = 10$, $C = 1$, $\lambda = 2$, $\mu = 2$, $\xi = 1$, $N = 5$, $\theta = 0.6$, $\alpha = 0.3$, $\gamma = 0.8$.

6.2. Exploring the impact of system parameters on the maximum social welfare

Social planners are usually more concerned about the social welfare of the service system, especially the maximum social welfare. This multiple vacation retrial queue with N-policy and breakdowns studied in this paper has more parameters, which effectively affect the trend of social welfare. Through a large number of numerical experiments, we find that the qualitative results of the influence of these parameters on social welfare are similar. Figures 6–8 present typical numerical experiments illustrating the respective effects of parameters R, C, λ, μ, θ, N, α, γ and ξ on the maximum social welfare $SW(q^*)$ of this service system.

From Figure 6 (a), we can see that the social planner increases the reward R of the customers will promote the growth of maximum social welfare, because as R increases, it will encourage more and more customers to join the unavailable system, which has a positive stimulating effect on the net profit of this service system. Conversely, increasing waiting cost of C per unit time will reduce the enthusiasm of customers to join the unavailable system, which will lead to a decrease in the maximum social welfare as the waiting cost of C per unit time increases, which as shown in Figure 6 (a).

Figure 6 (b) shows that the growth of both λ and μ can increase the maximum social welfare. Although the probabilities q_e and q^* of the two joining cases decrease as the arrival rate λ increases, which can be seen in Figure 2 (a), the maximum social welfare $SW(q^*)$ may not be significantly affected. This is because when the arriving customers find that the system is available, the arriving customers quickly join the system with a probability of 1, which effectively promotes the increase in social welfare. The increase in service rate μ allows the customers in the orbit to get fast service, and the waiting costs of the customers are reduced. As shown in Figure 2 (b), the customers are more willing to join the unavailable system. Therefore, the increase in service rate μ effectively promotes the growth of social welfare.
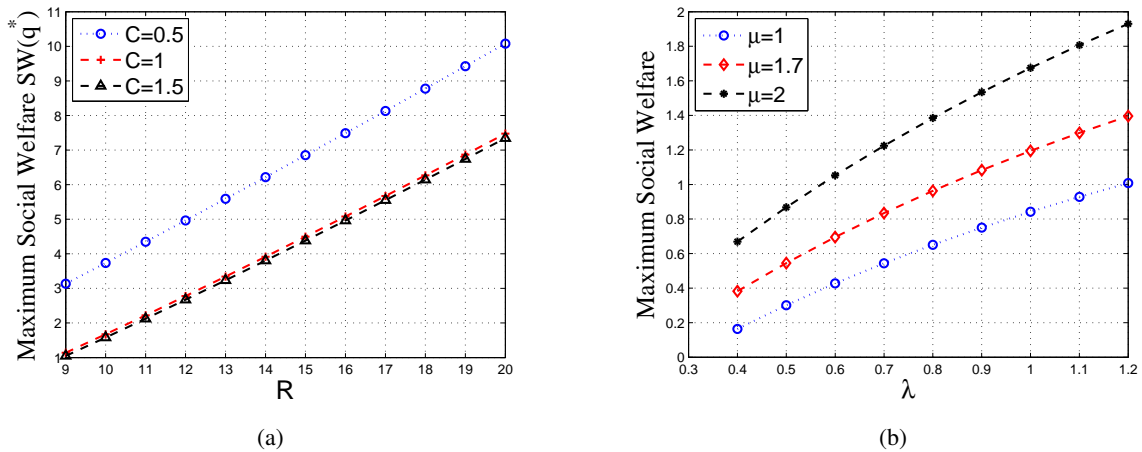
Figure 6. (a) The maximum social welfare vs. R for different values of C when $p = 1$, $\lambda = 1$ $\mu = 2$, $\xi = 1$, $N = 5$, $\theta = 0.6$, $\alpha = 0.3$, $\gamma = 0.8$. (b) The maximum social welfare vs. λ for different values of μ when $R = 10$, $p = 1$, $C = 1$, $\xi = 1$, $N = 5$, $\theta = 0.6$, $\alpha = 0.3$, $\gamma = 0.8$.

As shown in Figure 4 (a), the increase of α shortens the available time of the server, which leads to a decrease in the probabilities of q_e and q^* of the two joining cases. As a result, social welfare must decline, which can be seen in Figure 7 (a). On the contrary, as shown in Figure 4 (b), the shortened repair time increases the available time of the server, and the customers under the two joining cases are more willing to join the unavailable system. As a result, Figure 7 (a) shows that the maximum social welfare increases as γ increases. The social planner hopes to reduce the server's vacation time to reduce the number of customers on the waiting list, and the probability q_e and q^* of the customers also conform to that shown in Figure 3 (a). As shown in Figure 7 (b), this will contribute to the growth of social welfare.
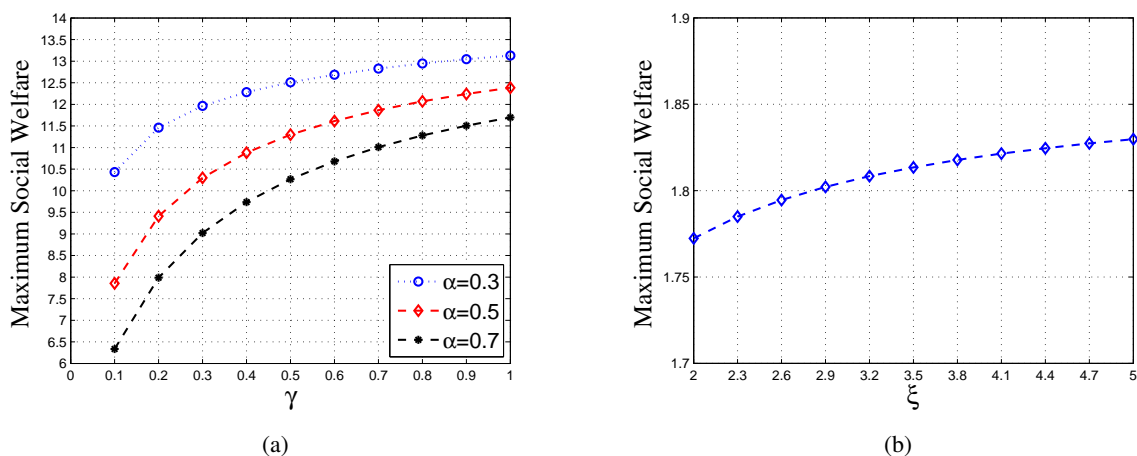

Figure 7. (a) The maximum social welfare vs. γ for different values of α when $R = 10$, $p = 1$, $C = 1$, $\lambda = 2$, $\mu = 2$, $\xi = 1$, $N = 5$, $\theta = 0.6$. (b) The maximum social welfare vs. ξ when $R = 10$, $p = 1$, $C = 1$, $\lambda = 1$, $\mu = 2$, $N = 5$, $\theta = 0.6$, $\alpha = 0.3$, $\gamma = 0.8$.

Figure 8 (a) illustrates that the maximum social welfare $SW(q^*)$ is an increasing function of the retrial rate θ. As shown in Figure 3 (b), as the retrial rate θ increases, the retrial frequency of the customers in the orbit is accelerated, and the arriving customers are more willing to join the system when the system is unavailable. Because the waiting cost of the customers is reduced, which encourages more customers to join the unavailable service system, which promotes the growth of maximum social welfare $SW(q^*)$. In Figure 8 (b), the social welfare $SW(q^*)$ decreases with N. This is because, with the increase of N, the more customers accumulate in the vacation state, the customers who later arrive at the system are more reluctant to join the system, which ultimately leads to a decline in the net income of the whole service system.
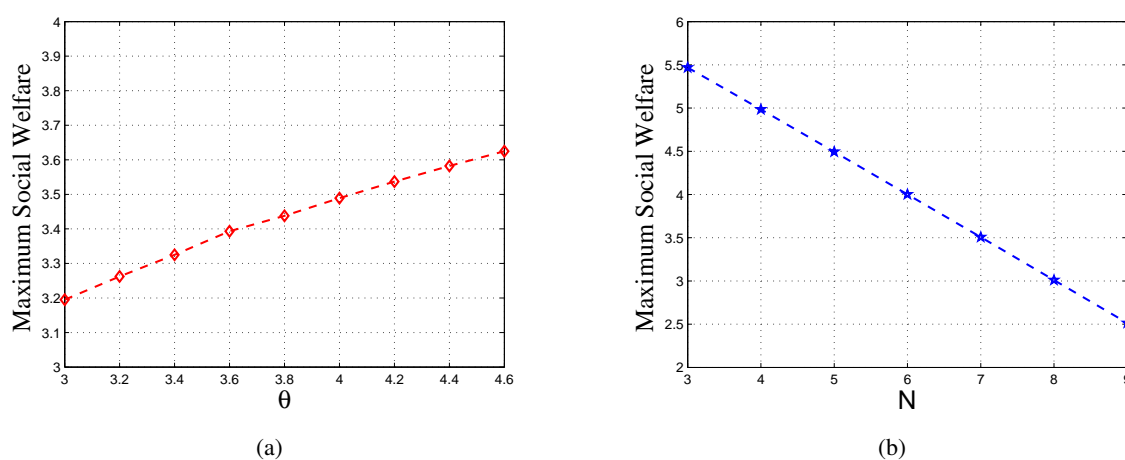

(a)

(b)

Figure 8. (a) The maximum social welfare vs. θ when $R = 10$, $p = 1$, $C = 1$, $\lambda = 1$, $\mu = 2$, $\xi = 1$, $N = 5$, $\alpha = 0.3$, $\gamma = 0.8$. (b) The maximum social welfare vs. N when $R = 15$, $p = 1$, $C = 1$, $\lambda = 1$, $\mu = 2$, $\xi = 1$, $\theta = 0.6$, $\alpha = 0.3$, $\gamma = 0.8$.

6.3. The profit of service provider

From the discussion in Section 5, we know that the profit of the service provider is dynamic, and the server can obtain the maximum profit by adjusting the entrance fee p and the service rate μ. It is difficult for us to obtain an explicit or closed-form representation of $S(\mu, p)$, so we use the PSO algorithm introduced earlier to explore how the server adjusts the price p and service rate μ to obtain maximum profit under equilibrium. We have run a large number of numerical experiments to find that their qualitative results are similar, so we illustrate the results through typical numerical scenarios. The key procedure of applying the PSO algorithm to find an optimal solution $(S(\mu, p))$ is illustrated in Algorithm 1, where the velocity V_{id} and position X_{id} are given in (5.2) and (5.3), respectively. Specifically, Figure 9 gives an example of finding the optimal solution of $S(\mu, p)$ through the improved PSO algorithm.
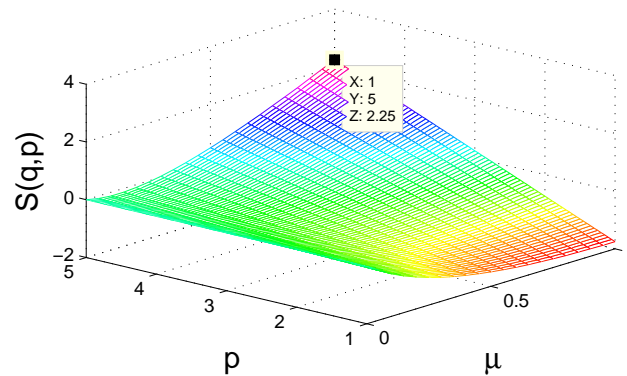
Figure 9. An example of searching for the optimal solution of $S(\mu, p)$ when $R = 5, C = 1,$ $\lambda = 1.5, \xi = 2, N = 5, \theta = 0.6, \alpha = 0.3, \gamma = 0.8, C_b = 3, C_i = 2.$

Algorithm 1 A example of searching for the maximum profit $S(\theta, p)$ of the server

---

Require: $R, C, \lambda, \xi, N, \theta, \alpha, \gamma, C_b, C_i$;
Ensure: μ, $p, S(\theta, p)$;

1: **for** each particle i
2:     Initializing velocity V_{id} and position X_{id} for each particle i
3:     Evaluating particle i and setting $pBest_i = X_{id}$
4: **end for**
5: $gBest_i$=min $\{pBest_i\}$
6: **while** not stop
7:   **for** $i = 1$ to M
8:     Updating the velocity and position of particle i
9:     Evaluating particle i
10:    **if** fit $(X_{id}) <$ fit $(pBest_i)$
11:      $pBest_i = X_{id}$
12:    **if** fit $(pBest_i) <$ fit $(gBest_i)$
13:      $gBest_i = pBest_i$
14:   **end for**
15: **end while**
16: print $gBest_i$
17: **end**

Now, we apply the PSO algorithm to explore how the server adjusts the service price p and service rate μ to maximize profit under the case of $q_e \neq 0$ and $\mu \neq 0$. On the basis of the service price p and service rate μ as control variables, as shown in Figure 10 (a), the profit $S(\mu, p)$ of the service provider decreases with respect to the arrival rate λ. This is because as the arrival rate λ increases, more and more customers accumulate in the orbit. At this time, the server needs to increase the service rate μ to reduce congestion, and the increase in service price p is used to reduce service costs. Ultimately, because of the increase in the arrival rate λ, the congestion of the system, and the increase in waiting costs have led to a decline in the profit $S(\mu, p)$ of the server. However, as shown in Figure 10 (b), the

profit $S(\mu, p)$ of the service provider increases with respect to the retrial rate θ. The reason is that the increase in the retrial rate θ helps reduce the length of the orbit, which promotes the rapid operation of the system, the result is an increase in the profits of the service provider.
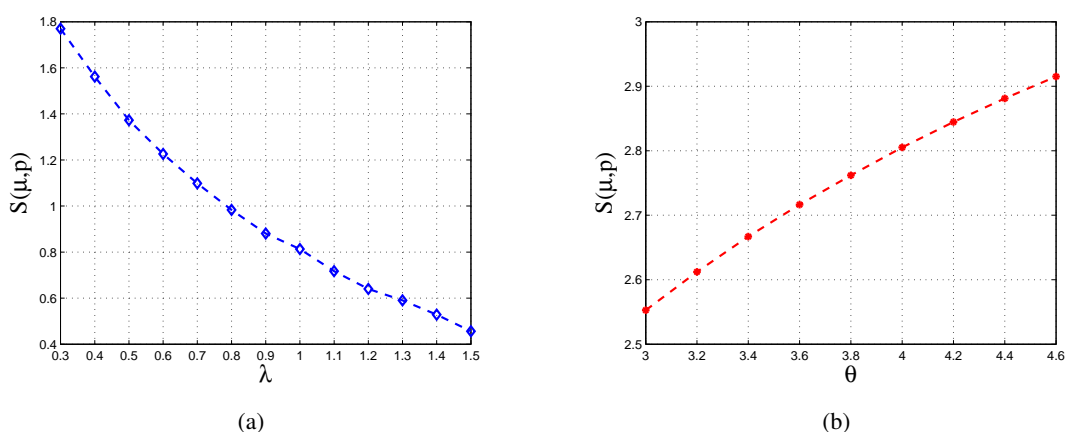

Figure 10. (a) The profit $S(\mu, p)$ of service provider vs. λ when $R = 5$, $C = 1$, $\xi = 1$, $N = 5$, $\theta = 0.6$, $\alpha = 0.3$, $\gamma = 0.8$, $C_b = 3$, $C_i = 1$. (b) The profit $S(\mu, p)$ of service provider vs. θ when $R = 5$, $C = 1$, $\lambda = 1$, $\xi = 1$, $N = 5$, $\alpha = 0.3$, $\gamma = 0.8$, $C_b = 3$, $C_i = 1$.

Figure 11 (a) illustrates that the profit $S(\mu, p)$ of the server decreases with respect to the consumption cost C_i of the server per time unit in the idle state. This is because as the increase of C_i, the unit time consumption cost of the idle state keeps increasing, and the social planner will shorten the idle time of the server to reduce the consumption cost. The shortening of idle time reduces the number of customers who join the system with probability 1 when the system is available. In the case of increased costs and a decrease in the number of customers who join the system with probability 1, the profit of the service provider is ultimately reduced. Similarly, as shown in Figure 11 (b), with the increase of C_b, the unit time consumption of the busy state will continue to increase. In order to reduce costs, the server will shorten the working time of the server, which will cause more and more customers to accumulate in the orbit, the customers who arrive later are more reluctant to join the system, which will eventually lead to a decrease in the effective arrival rate, which will result in a decrease in the server profit $S(\mu, p)$.
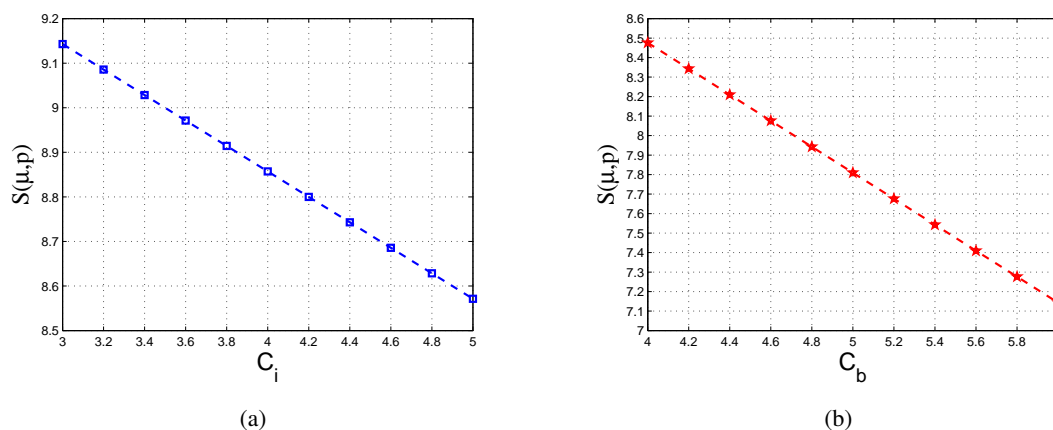
Figure 11. (a) The profit $S(\mu, p)$ of service provider vs. C_i when $R = 5$, $C = 1$, $\lambda = 2$, $\xi = 1$, $N = 5$, $\theta = 3.6$, $\alpha = 0.3$, $\gamma = 0.8$, $C_b = 3$. (b) The profit $S(\mu, p)$ of service provider vs. C_b when $R = 5$, $C = 1$, $\lambda = 2$, $\xi = 1$, $N = 5$, $\theta = 3.6$, $\alpha = 0.3$, $\gamma = 0.8$, $C_i = 3$.

7. Conclusions and further research

Motivated by the cost control of service and information, a multiple vacations retrial queue with N-policy and breakdowns is studied in this paper. Two types of customer joining cases apply to this paper, i.e., the non-cooperative customers aim to optimize individual interests, and the social planner in the cooperative case consider the profit of the whole service system. For this service system, we presented the equilibrium joining strategy for the non-cooperative case and the socially optimal joining strategy for the cooperative case.

For the maximum profit of the server, it is difficult for us to obtain an explicit or closed-form representation of $S(\mu, p)$, so we use the improved PSO algorithm to explore how the server obtains maximum profit. A large number of numerical experiments show that the qualitative results of the parameters are similar to the system, so we use some typical numerical scenarios to explain the influence of the parameters on the two joining probabilities, and how the parameters affect the changing trend of maximum social welfare. Both the model itself and the numerical experiment have certain guiding significance for the actual service system. Because there is no waiting space in front of the server and the virtuality of the waiting list, this model is suitable for emergency disaster or epidemic control. Changes in the parameters assumed by the model, resulting in changes in the joining strategy are interesting topics for future research.

Acknowledgements

Conflict of interest

No potential conflict of interest was reported by the authors.

References

1. P. Naor, The Regulation of Queue Size by Levying Tolls, *Econometrica*, **37** (1969), 15–24.

2. N. Edelson, D. Hilderbrand, Congestion tolls for Poisson queuing processes, *Econometrica*, **43** (1975), 81–92.

3. Y. Shi, Z. Lian, Optimization and strategic behavior in a passenger–taxi service system, *Eur. J. Oper. Res.*, **249** (2016), 1024–1032.

4. O, Boudali, A. Economou, Optimal and equilibrium balking strategies in the single server markovian queue with catastrophes, *Eur. J. Oper. Res.*, **218** (2012), 708–715.

5. A. Economou, A. Manou, Equilibrium balking strategies for a clearing queueing system in alternating environment, *Ann. Oper. Res.*, **208** (2013), 489–514.

6. A. Burnetas, A. Economou, Equilibrium customer strategies in a single server markovian queue with setup times, *Queueing Syst.*, **56** (2007), 213–228.

7. Y. Zhang, J. Wang, Equilibrium pricing in an M/G/1 retrial queue with reserved idle time and setup time, *Appl. Math. Model.*, **49** (2017), 514–530.

8. J. Chang, J. Wang, Unreliable M/M/1/1 retrial queues with set-up time, *Qual. Technol. Quantit. M.*, **15** (2018), 589–601.

9. R. Hassin, M. Haviv, *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems*, Kluwer Academic Publishers, Boston, 2003.

10. L. Kosten, *Stochastic theory of service systems*, Pergamon Press, University of California, 1973.

11. T. Jiang, Tail asymptotics for a batch service polling system with retrials and nonpersistent customers, *J. Math. Anal. Appl.*, **459** (2018), 893–905.

12. T. Phung-Duc, Single server retrial queues with setup time, *J. Ind. Manag. Optim.*, **13** (2017), 75–78.

13. A. Economou, S. Kanta, Equilibrium customer strategies and social–profit maximization in the single-server constant retrial queue, *Nav. Res. Log.*, **58** (2011), 107–122.

14. J. Wang, W. Li, Noncooperative and cooperative joining strategies in cognitive radio networks with random access, *IEEE T. Veh. Technol.*, **65** (2015), 5624–5636.

15. J. Wang, F. Zhang, Monopoly pricing in a retrial queue with delayed vacations for local area network applications, *IMA J. Manag. Math.*, **27** (2016), 315–334.

16. L. Barroso, U. Holzle, The case for energy-proportional computing, *Computer*, **40** (2007), 33–37.

17. Z. Wang, L. Liu, Y. Shao, X. Chai, B. Chang, Equilibrium joining strategy in a batch transfer queuing system with gated policy, *Methodol. Comput. Appl.*, **22** (2020), 75–99.

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

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

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

21. Q. Ye, L. Liu, The analysis of discrete time Geom/Geom/1 queue with single working vacation and multiple vacations (geom/geom/1/swv+ mv), *RAIRO-Operations Res.*, **52** (2018), 95–117.

22. S. Gao, H. Dong, X. Wang, Equilibrium and pricing analysis for an unreliable retrial queue with limited idle period and single vacation, *Oper. Res.*, (2018), 1–23.

23. R. Eberhart, J. Kennedy, A new optimizer using particle swarm theory, *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, (1995), 39–43.

AIMS Press