

OPTIMAL QUEUING STRATEGIES FOR AN M/G/1 RETRIAL QUEUE SYSTEM WITH RWV AND ISEV POLICIES

Z. CHEN ¹, H. XU ¹ and H. HUO ¹

(Received 10 March, 2022; accepted 15 January, 2024; first published online 19 March, 2024)

Abstract

Designing a reasonable M/G/1 retrial queue system that enhances service efficiency and reduces energy consumption is a challenging issue in Information and Communication Technology systems. This paper presents an M/G/1 retrial queue system incorporating random working vacation (RWV) and improved service efficiency during vacation (ISEV) policies, and examines its optimal queuing strategies. The RWV policy suggests that the server takes random working vacations during reserved idle periods, effectively reducing energy consumption. In contrast, the ISEV policy strives to augment service efficiency during regular working periods by updating, inspecting or maintaining the server on vacations. The system is transformed into a Cauchy problem to investigate its well-posedness and stability, employing operator semigroup theory. Based on the system's stability, steady-state performance measures, such as service efficiency, energy consumption and expected costs, are quantified using the steady-state solution. The paper subsequently demonstrates the existence of optimal queuing strategies that achieve maximum efficiency and minimum expected costs. Finally, two numerical experiments are provided to illustrate the effectiveness of the system.

2020 *Mathematics subject classification*: 35F05.

Keywords and phrases: retrial queue, system dynamics, optimal queuing strategies, asymptotic stability.

1. Introduction

Queuing theory has found extensive applications across various fields, including Information and Communication Technology (ICT) systems, local area networking and service industries. To address diverse requirements, researchers have investigated and implemented different types of queue system models [5, 12, 17]. The retrial queue system, in particular, has gained prominence as a noteworthy model for addressing

¹Department of Mathematics and Statistics, Beijing Institute of Technology, Beijing, PR China
e-mail: chenzhuoqian123@163.com, xuhoubao@bit.edu.cn, huixiahuo@163.com

© The Author(s), 2024. Published by Cambridge University Press on behalf of Australian Mathematical Publishing Association Inc.

the issue of customers having to leave the system due to finite waiting space during peak demand, especially in telecommunication networks and mobile systems [1, 18, 20]. Falin and Templeton [12] have provided a comprehensive overview of the main techniques and outcomes of a retrial queue system, while further research in this field can be found in the works of Artalejo and Gómez-Corral [2–4].

Server vacation policies have been extensively studied as an effective measure for energy conservation in systems that consume a substantial amount of energy during idle periods, such as computer communication, manufacturing, production and inventory systems. From an economic perspective, Burnetas and Economou [6] investigated the relationship between the vacation setup times and energy consumption in a single-server Markovian queue. Meanwhile, the hibernation strategy, studied by Zhang et al. [30], Son et al. [23] and Wu et al. [28], has been shown to be effective for energy-efficient base station vacations in cellular networks during periods of inactivity. Furthermore, in the literature, queuing models with vacation have been proposed by Doshi [9], Takagi [24] and Do [7], which incorporate vacations for checks, maintenance and searching for new work, thereby enhancing server efficiency while achieving additional energy savings.

Although server vacation is an effective energy-saving method to conserve energy, increasing the server's vacation-time may lead to an increased workload and the sojourn time of jobs in the system, particularly in systems providing services such as network, web, file transfer and mail services. To address such issues, the working vacation (WV) policy was introduced by Servi and Finn [22] in an M/M/1 queue system, where the server continues to provide services at a reduced speed during the vacation period rather than ceasing service completely. Servi and Finn [22] studied an M/M/1/WV queue system, and derived the transform formula for the distribution of the number of customers in the system and the sojourn time in the steady-state. Furthermore, Wu and Takagi [27] generalized the results of [22] to an M/G/1/WV queue, while Kim [16] analysed the queue length distribution of the M/G/1/WV. Other notable studies on queue systems with working vacations can be found in the works of Liu and Song [19], Gao and Liu [13], Gao et al. [14] and Zhang and Liu [29].

Aforementioned research on vacation or working vacation policies only considered cases where a vacation commences when there are no customers in the orbit. However, ICT systems that provide services over the internet often consume substantial energy when the server is waiting for customers in the orbit to retry, such as network, file transfer and mail services. To address this problem, a common approach is to close the server when it is idle. Zhang and Wang [31] developed an M/G/1 retrial queue system with reserved idle time and setup time from the perspective of server vacation for energy-saving. Once the server becomes idle, it can be shut down with a certain probability, even if there are customers in the orbit. However, shutting down the server completely when it is idle will increase its workload and decrease its productivity, which cannot meet the requirements of ICT systems. Therefore, there exists a trade-off between energy consumption and service efficiency, which presents important and

challenging problems in developing a reasonable $M/G/1$ retrial queue system for ICT systems and designing optimal queuing strategies for energy-saving and service efficiency improvement.

In this paper, we construct an $M/G/1$ retrial queue system with random working vacation (RWV) and increased service efficiency during vacation (ISEV) policies. The RWV policy, which means the server takes a random working vacation with a given probability during reserved idle periods, is proposed to reduce energy consumption. The ISEV policy is to perform update, inspection or maintenance during working vacations to augment service efficiency in the regular working periods.

The remainder of this paper is organized as follows. Section 2 formulates the retrial queue system and transforms it into an abstract Cauchy problem in a suitable Banach space. In Section 3, the dynamic behaviour of the retrial queue system is studied and the equilibrium condition for the system to be stable is derived to ensure the feasibility of the system. Based on the stability analysis in Section 3, Section 4 carries out an equilibrium analysis of the system with its special cases to demonstrate the advantages of the proposed system and further obtain the optimal queuing strategies for energy-saving and cost reduction. In Section 5, two numerical experiments are presented to illustrate the effectiveness of the proposed system. Section 6 concludes the paper.

2. Retrial queue system formulation

In ICT systems [17], service disciplines stipulate that unprocessed jobs upon arrival will be resent at a later time, aligning with the concept of retrial queues. For example, within a local area network, stations or processors are interconnected by a single bus (server) that is essential for facilitating communication among these stations. Messages generated by users (primary customers) reach the stations from external sources. Upon receipt of a message, the station verifies whether the bus is idle. If the bus is idle, the message is transmitted via the bus to the target station. Conversely, the message is stored in a buffer (retrial customers) and will attempt to access the bus again after a specified interval, typically referred to as the retrial orbit. The message repeats this process within the retrial orbit until it successfully locates an idle bus.

To optimize utilization efficiency, the bus generally does not maintain a continuous open state; instead, it sets a period of reserved idle time. Specifically, when the bus is idle, it remains in the state for a predetermined duration, waiting for incoming messages (from primary or retrial customers) and initiating transmission immediately upon message arrival. If no messages are received during the reserved idle period, it can be reasonably assumed that the station has few stored messages, and the bus will shut down until the arrival of new messages prompts it to reactivate. This service strategy with reserved idle time and reactivate settings is often used in switched virtual channel connection (SVCCs) services, where VCC dynamically sets up and shuts down virtual channel connections as needed.

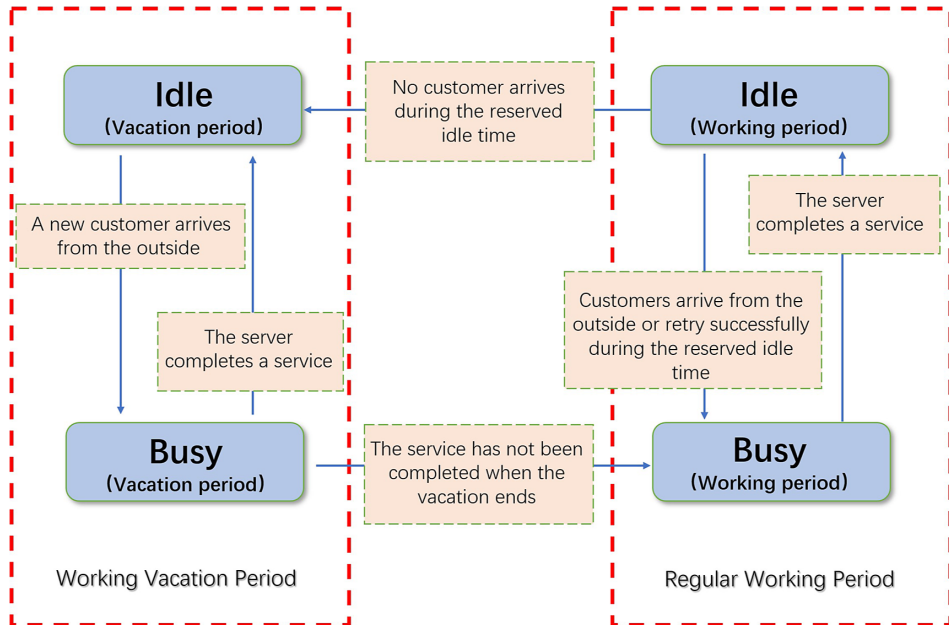


FIGURE 1. Network topology for the model.

However, to minimize the mean sojourn time of messages awaiting bus setup and consequently enhance the efficiency of the system, the bus can transition to an RWV period rather than shutting down when no messages arrive during the reserved idle time. This strategy is referred to as the RWV policy in this paper. Upon the arrival of new messages to the system, the bus operates at a reduced efficiency until the working vacation concludes. Furthermore, the server may undertake additional tasks during the working vacation stage to the ISEV policy within this paper.

Grounded in these principles, this paper presents a mathematical transformation of the aforementioned operations and strategies, and then constructs an M/G/1 retrial queue system with RWV and ISEV policies.

2.1. Model assumption and descriptions Before constructing the mathematical model of a system, in this subsection, we first make reasonable assumptions based on the model description. The network topology for the model is presented in Figure 1. The detailed descriptions and assumptions of the model are as follows.

- Arrival process: The primary customers arrive at the system according to a Poisson process with rate λ .
- Retrial process: We assume that there is no waiting area available, and customers who find the server idle will immediately occupy it. If the server is busy or on a working vacation, the arriving customers will join a retrial orbit. After joining the orbit, customers will attempt to retry the service only after the server completes a

regular service. The retrial time is assumed to follow an exponential distribution with parameter θ .

- RWV policy: Once the server completes a service, it enters an idle state and remains active for a certain period of time, referred to as reserved idle time, which follows an exponential distribution with parameter β . If no customers access the server during the reserved idle time, the server enters a working vacation period and provides service with low working efficiency to any occupying customers until the service is completed or the vacation ends. We assume that the service time during the working vacation period follows an exponential distribution with parameter μ_1 , and the time until the end of the vacation follows an exponential distribution with parameter α , which is also called vacation interruption rate in this paper.
- ISEV policy: The ISEV policy is implemented to reduce the loss due to low service efficiency during the server's working vacation. This involves the server undergoing rejuvenation, inspection or maintenance during the working vacation period to increase its service efficiency when it completes the vacation and enters a regular working period.
- Regular service process: If customers arrive during the idle period, they can immediately occupy the server. It is reasonable to assume that the service time under the ISEV policy follows a general distribution with a hazard rate of $\mu_2(x)$, where the variable $x \in [0, \infty)$ represents the elapsed service time. The hazard rate $\mu_2(x)$ quantifies the instantaneous likelihood of service completion at a specific moment. Therefore, the probability of a service completing in an interval Δx can be approximated by $\mu_2(x)\Delta x + o(\Delta x)$, where $o(\cdot)$ denotes the higher-order infinitesimal quantity. In this paper, the function $\mu_2(x)$ is also referred to as the service completion rate under the ISEV policy. Since the ISEV policy is implemented to increase the server's service efficiency when it completes the vacation and enters a regular working period, thus let $\mu_2^-(x)$ be the service completion rate without the ISEV policy, then $\mu_2(x)$ can be assumed to be

$$\mu_2(x) = (1 + f(\alpha))\mu_2^-(x),$$

where $f(\alpha)$ is a nonnegative, bounded real-value function designed to measure the impact of the vacation interruption rate α on $\mu_2(x)$. As the vacation interruption rate α increases and tends towards infinity, $1/\alpha$ will decrease and approach 0, implying that the mean vacation time of the server decreases to 0. In this situation, the ISEV policy becomes ineffective in improving the server's efficiency when it completes the vacation. Therefore, it is further assumed that $f(\alpha)$ is a monotonically decreasing function with respect to α , and $f(\alpha) \rightarrow 0$ as $\alpha \rightarrow \infty$.

Based on such assumptions above, we denote the state of the server and the number of customers in the orbit at time t by $C(t)$ and $N(t)$, respectively. Then,

$$C(t) = \begin{cases} 0 & \text{if the server is in regular working period} \\ & \text{and the server is idle,} \\ 1 & \text{if the server is in working vacation period} \\ & \text{and the server is busy,} \\ 2 & \text{if the server is in working vacation period} \\ & \text{and the server is idle,} \\ 3 & \text{if the server is in regular working period} \\ & \text{and the server is busy.} \end{cases}$$

In ordinary M/G/1 queues, the stochastic process $\{N(t), t \geq 0\}$ is not Markovian. Thus, in the next subsection, we will employ the random variable $X(t)$, which denotes the elapsed service time of the customers in service at time t to form a Markovian process $\{(C(t), N(t), X(t)) : t \geq 0\}$ and further construct the system.

2.2. System abstraction Based on the model assumptions and descriptions presented in Section 2.1, the retrial queue system can be described as a semi-Markov process $\{(C(t), N(t), X(t)) : t \geq 0\}$ by introducing additional variables. This subsection will derive the Chapman–Kolmogorov forward equations of the system based on the process $\{(C(t), N(t), X(t)) : t \geq 0\}$, and transform it into a Cauchy problem for further analysis.

Let $P_{i,n}(t)$ ($i = 0, 1, 2, n = 0, 1, 2, \dots$) denote the probability that the server is in state i , and there are n customers in the queue at time t . Similarly, $P_{3,n}(t, x) dx$ ($n = 0, 1, 2, \dots$) represents the probability that the server is in state 3, there are n customers in the queue and the elapsed time of the customer being served is between x and $x + dx$ at time t . Then, the state space of the process $\{(C(t), N(t), X(t)) : t \geq 0\}$ can be denoted by

$$\{(0, j), (1, j), (2, j), (3, j, x) \mid 0 \leq j < \infty, 0 \leq x < \infty\}.$$

The corresponding transition rate diagram can be summarized and presented in Figure 2. By relating the states of the system at time t and $t + \Delta t$, we can derive the Chapman–Kolmogorov forward equations [21] and further develop the system model as follows:

$$\begin{cases} \frac{dP_{0,n}(t)}{dt} = (-\lambda - \beta - n\theta)P_{0,n}(t) + \int_0^\infty \mu_2(x)P_{3,n}(t, x) dx, \\ \frac{dP_{1,n}(t)}{dt} = (-\lambda - \alpha - \mu_1)P_{1,n}(t) + \lambda P_{1,n-1}(t) + \lambda P_{2,n}(t), \\ \frac{dP_{2,n}(t)}{dt} = -\lambda P_{2,n}(t) + \beta P_{0,n}(t) + \mu_1 P_{1,n}(t), \\ \frac{\partial P_{3,n}(t, x)}{\partial t} + \frac{\partial P_{3,n}(t, x)}{\partial x} = (-\lambda - \mu_2(x))P_{3,n}(t, x) + \lambda P_{3,n-1}(t, x), \end{cases} \quad (2.1)$$

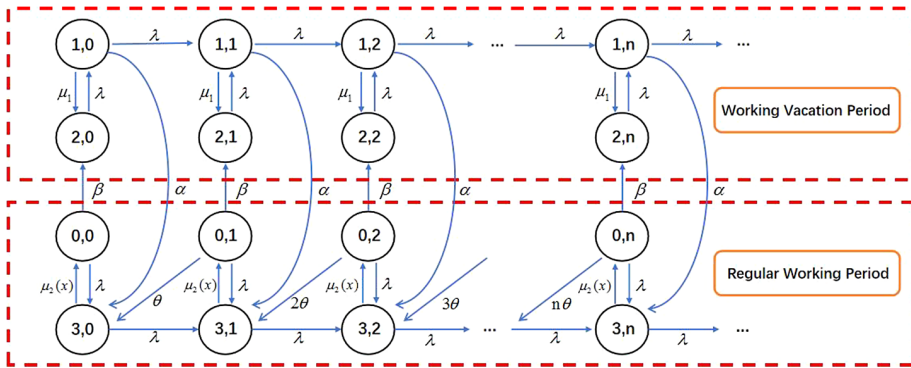


FIGURE 2. Transition rate diagram.

where $n = 0, 1, 2, \dots$, and $P_{1,-1}(t) = 0, P_{3,-1}(t, x) = 0$. The boundary conditions and the initial values are

$$\begin{cases} P_{3,n}(0) = \alpha P_{1,n} + (n + 1)\theta P_{0,n+1} + \lambda P_{0,n}, \\ P_{0,0}(0) = 1, P_{0,j}(0) = 0, \quad j = 0, 1, 2, \dots, \\ P_{i,j}(0) = 0, P_{3,j}(0, x) = 0, \quad i = 1, 2, \quad j = 0, 1, 2, \dots \end{cases}$$

Regarding the practical background, for any fixed α , we assume that

$$\begin{cases} \mu_2(x) \geq 0, \quad \sup_{x \in [0, \infty)} \mu_2(x) < \infty, \\ \int_0^T \mu_2(x) dx < \infty \quad \text{for all } T \in [0, \infty), \quad \int_0^\infty \mu_2(x) dx = \infty. \end{cases}$$

Next, we translate the system into an abstract Cauchy problem in a Banach space. First, let the state space

$$X = \left\{ \vec{P} = (P_0, P_1, P_2, \dots)^T \mid \|\vec{P}\|_X = \sum_{n=0}^\infty \|P_n\| < \infty \right\},$$

where

$$P_n = (P_{0,n}, P_{1,n}, P_{2,n}, P_{3,n}(x))^T \in R^3 \times L^1(R^+), \quad R^+ = [0, \infty),$$

$$\|P_n\| = |P_{0,n}| + |P_{1,n}| + |P_{2,n}| + \|P_{3,n}(x)\|_{L^1}, \quad n = 0, 1, 2, \dots,$$

and $(P_0, P_1, P_2, \dots)^T$ stands for the transpose of vector (P_0, P_1, P_2, \dots) . Note that $(X, \|\cdot\|_X)$ is a Banach space. We define some operators in X for simplicity:

$$A = \text{diag}\{A_0, A_1, A_2, \dots\}, \quad B = \begin{pmatrix} B_0 & 0 & 0 & 0 & \dots \\ B_1 & B_0 & 0 & 0 & \dots \\ 0 & B_1 & B_0 & 0 & \dots \\ 0 & 0 & B_1 & B_0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

$$C\vec{P} = \text{diag}\{C_0P_0, C_1P_1, C_2P_2, \dots\},$$

where

$$A_n = \text{diag}\left\{-\lambda - \beta - n\theta, \quad -\lambda - \alpha - \mu_1, \quad -\lambda, \quad -\frac{d}{dx} - \lambda - \mu_2(x)\right\},$$

$$B_0 = \begin{pmatrix} 0 & & & \\ & 0 & \lambda & \\ \beta & \mu_1 & 0 & \\ & & & 0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 0 & & & \\ & \lambda & & \\ & & 0 & \\ & & & \lambda \end{pmatrix}, \quad C_n P_n = \begin{pmatrix} 0 & & & G_n \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix},$$

$$G_n = \int_0^\infty \mu_2(x) P_{3,n}(x) dx.$$

We denote the domain of operators A , B and C by $D(A)$, $D(B)$ and $D(C)$, respectively. Here,

$$D(A) = \left\{ \vec{P} = (P_0, P_1, P_2, \dots)^T \in X \mid \frac{dP_{3,n}(x)}{dx} \in L^1(\mathbb{R}^+) \right\},$$

where $P_{3,n}(x)$ are continuous functions satisfying

$$P_{3,n}(0) = \alpha P_{1,n} + (n + 1)\theta P_{0,n+1} + \lambda P_{0,n}, \quad n = 0, 1, 2, \dots,$$

and $D(B) = D(C) = X$. Thus, by denoting $\vec{P}(t) = (P_0(t), P_1(t), P_2(t), \dots)^T$ and $P_i(t) = (P_{0,i}(t), P_{1,i}(t), P_{2,i}(t), P_{3,i}(t, x))^T, i = 0, 1, \dots$, the M/G/1 retrial queue system with RWV and ISEV policy (2.1) can be rewritten as an abstract Cauchy problem in the Banach space X as follows:

$$\begin{cases} \frac{d\vec{P}(t)}{dt} = (A + B + C)\vec{P}(t), \\ \vec{P}(0) = ((1, 0, 0, \mathbf{0})^T, (0, 0, 0, \mathbf{0})^T, (0, 0, 0, \mathbf{0})^T, \dots)^T. \end{cases} \tag{2.2}$$

Following this, studying the evolution of stochastic process $\{(C(t), N(t), X(t)) \mid t \geq 0\}$ can be transformed into the problem of analysing the dynamic behaviour of the system (2.2). Therefore, in the next section, we will utilize operator semigroup theory to study the system operator $A + B + C$, thereby investigating the dynamic behaviour of the system.

3. Dynamic behaviour of the retrial queue system

The well-posedness and asymptotic stability of a system are important characteristics that describe the dynamic behaviour of the system, and they are also necessary conditions for conducting research on optimal queuing strategies for the system. Based on Cauchy problem (2.2), this section will investigate the well-posedness and asymptotic stability of the system to study the dynamic behaviour of the system by using C_0 -semigroup theory and spectral theory.

A system is considered well-posed if and only if it possesses a unique nonnegative time-dependent solution. According to Engel et al. [10, page 151, Definition 6.8 and Corollary 6.9], system (2.2) can be regarded as well-posed if the system operator

$A + B + C$ can generate a contractive C_0 semigroup in the system space X . With reference to the Hille–Yosida theorem [21], it is evident that the operator $A + B + C$ can indeed generate a C_0 semigroup on the space X . We summarize the main theorem as follows.

THEOREM 3.1. *The M/G/1 retrial queue system with RWV and ISEV policy (2.2) has a unique nonnegative time-dependent solution $\vec{P}(t)$, which can be expressed as $\vec{P}(t) = T(t)\vec{P}_0$, where $\{T(t)\}_{t \geq 0}$ are the C_0 semigroups generated by the system operator $A + B + C$ and \vec{P}_0 is the initial condition. Moreover, for any $t \geq 0$, $\vec{P}(t)$ satisfies that*

$$\|\vec{P}(t)\|_X = 1.$$

The well-posedness of system (2.2) indicates that the system possesses a unique nonnegative dynamic solution. A similar proof to that of Theorem 3.1 can be found in [8, 25, 26]. In the following, we investigate the spectrum distribution of the system operator $A + B + C$ and demonstrate that the dynamic solution of the system strongly converges to the steady-state solution under necessary conditions.

THEOREM 3.2. *If $\lambda \int_0^\infty e^{-\int_0^x \mu_2(\xi) d\xi} dx < 1$ and $\lambda \leq \alpha$, the time-dependent solution of the M/G/1 retrial queue system with RWV and ISEV policy (2.2), $\vec{P}(t)$ strongly converges to its steady-state solution \vec{P}^* as $t \rightarrow \infty$.*

PROOF. According to the asymptotic stability theorem introduced by Gupur [15, page 48], we can divide the proof into three steps as follows.

- *Step 1:* We prove that $\{\gamma \in C \mid \operatorname{Re} \gamma > 0 \text{ or } \gamma = ia, a \in R \setminus \{0\}\}$ belongs to the resolvent set of system operators $A + B + C$.

According to the closed operator and inverse operator theorems, $\{\gamma \in C \mid \operatorname{Re} \gamma > 0 \text{ or } \gamma = ia, a \in R \setminus \{0\}\}$ belongs to the resolvent set of system operators $A + B + C$, which is equivalent to the existence of the bounded invertible operator $\gamma I - (A + B + C)$. So, by verifying that the operator equation $[\gamma I - (A + B + C)]\vec{P} = \vec{Y}$ has a unique nonzero solution, for any

$$\gamma \in \{\gamma \in C \mid \operatorname{Re} \gamma > 0 \text{ or } \gamma = ia, a \in R \setminus \{0\}\},$$

and $\vec{Y} \in X$, we can accomplish the proof of Step 1.

- *Step 2:* We prove that 0 is an eigenvalue of the system operator $A + B + C$ and the eigenvector of 0 is positive if

$$\lambda \int_0^\infty e^{-\int_0^x \mu_2(\xi) d\xi} dx < 1 \quad \text{and} \quad \lambda \leq \alpha.$$

It is difficult to solve the equation corresponding to eigenvalue 0 as an infinite-dimensional problem under the threshold condition. In this paper, we resolve this difficulty by flexibly utilizing several properties of the series; the details of the proof are given in Appendix A.

- *Step 3:* We verify that 0 is a simple eigenvalue of operator $(A + B + C)^*$, and the algebraic multiplicity of 0 is one, where $(A + B + C)^*$ is the conjugate operator of $(A + B + C)$.

By calculating the expression of operator $(A + B + C)^*$ and combining with the definition of algebraic multiplicity, it can be proved that 0 is a simple eigenvalue of operator $(A + B + C)^*$, and the algebraic multiplicity of 0 is one. The details of the proof are also given in Appendix B.

With Step 1–Step 3, we derive that

$$\sigma_p(A + B + C) \cap iR = \{0\} \quad \text{and} \quad \{\gamma \in C \mid \gamma = ia, a \neq 0, a \in R\}$$

belong to the resolvent set of $A + B + C$; the algebraic multiplicity of 0 in X^* is one. Furthermore, according to the asymptotic stability theorem of Gupur [15, page 48], time-dependent solution of the system $\vec{P}(t)$ strongly converges to its steady-state solution \vec{P}^* as $t \rightarrow \infty$. □

Theorems 3.1 and 3.2 ensure the well-posedness and stability of the M/G/1 retrial queue system with RWV and ISEV policy, which means that the system (2.1) considered in this study has a unique time-dependent solution, and the solution asymptotically converges to the steady-state solution. Therefore, based on the theoretical analysis in Section 3, we can study the steady-state indices of the system and investigate the optimal queuing strategy.

4. Optimal queuing strategies

In this section, we investigate optimal queuing strategies for the M/G/1 retrial queue system with RWV and ISEV policies. First, we calculate the steady-state indices of the system, including its special cases. Then, based on these indices, we define and analyse the system’s performance measures, including energy consumption, service efficiency and expected cost. Finally, we demonstrate the existence of optimal queuing strategies from the point of efficiency and expected cost. We also show that the optimal strategy can be achieved by adjusting the vacation length.

4.1. Steady-state indices calculation Since 0 is the eigenvalue of the system operator, the steady-state solution of the system is equivalent to the eigenvector corresponding to 0, which can be derived from equation $(A + B + C)\vec{P} = 0$, that is,

$$\begin{cases} (\lambda + \beta + n\theta)P_{0,n} - \int_0^\infty \mu_2(x)P_{3,n}(x) dx = 0, \\ (\lambda + \alpha + \mu_1)P_{1,n} - \lambda P_{2,n} - \lambda P_{1,n-1} = 0, \\ \lambda P_{2,n} - \beta P_{0,n} - \mu_1 P_{1,n} = 0, \\ \frac{d}{dx}P_{3,n}(x) + (\lambda + \mu_2(x))P_{3,n}(x) - \lambda P_{3,n-1}(x) = 0, \end{cases} \tag{4.1}$$

with $P_{3,n}(0) = \alpha P_{1,n} + (n + 1)\theta P_{0,n+1} + \lambda P_{0,n}$, where

$$P_{i,n} = \lim_{t \rightarrow \infty} P_{i,n}(t), \quad i = 0, 1, 2, \quad P_{3,n}(x) = \lim_{t \rightarrow \infty} P_{3,n}(t, x), \quad P_{1,-1} = 0, \quad P_{3,-1}(x) = 0.$$

It is possible for us to use the generation function method [31] to calculate the important indices. Let

$$\Pi_i(z) = \sum_{j=0}^{\infty} P_{i,j} z^j, \quad i = 0, 1, 2 \quad \text{and} \quad \Pi_3(z, x) = \sum_{j=0}^{\infty} P_{3,j}(x) z^j$$

for all complex variables $|z| < 1$. Such definitions of $\Pi_i(z)$ and $\Pi_3(z, x)$ at $z = 1$ are reasonable, since the stability of the system has been derived in Section 3. Then, by multiplying (4.1) by z^n and summing over n , the equation can be rewritten as

$$\begin{cases} (\lambda + \beta)\Pi_0(z) + \theta z \Pi'_0(z) = \int_0^{\infty} \mu_2(x) \Pi_3(z, x) dx, \\ (\lambda + \alpha + \mu_1)\Pi_1(z) = \lambda \Pi_2(z) + \lambda z \Pi_1(z), \\ \lambda \Pi_2(z) = \beta \Pi_0(z) + \mu_1 \Pi_1(z), \\ \left(\frac{\partial}{\partial x} + \lambda + \mu_2(x)\right) \Pi_3(z, x) = \lambda z \Pi_3(z, x), \end{cases} \tag{4.2}$$

with the boundary condition

$$\Pi_3(z, 0) = \theta \Pi'_0(z) + \alpha \Pi_1(z) + \lambda \Pi_0(z). \tag{4.3}$$

Solving (4.2) with (4.3), we have

$$\begin{cases} \Pi_1(z) = \frac{\beta}{\lambda + \alpha - \lambda z} \Pi_0(z), \quad \Pi_2(z) = \left(\frac{\beta}{\lambda} + \frac{\mu_1 \beta}{\lambda(\lambda(1-z) + \alpha)}\right) \Pi_0(z), \\ \Pi_3(z, x) = \left[\theta \Pi'_0(z) + \left(\lambda + \frac{\beta \alpha}{\lambda + \alpha - \lambda z}\right) \Pi_0(z)\right] \exp\left(-\int_0^x \lambda(1-z) + \mu_2(\xi) d\xi\right), \\ \Pi_0(z) = C \exp\left\{-\int_z^1 \frac{(\lambda(\lambda + \alpha - \lambda y) + \alpha \beta)h(y) - (\lambda + \beta)(\lambda(1-y) + \alpha)}{\theta(\lambda + \alpha - \lambda y)(y - h(y))} dy\right\}, \end{cases}$$

where

$$h(y) = \int_0^{\infty} \mu_2(x) e^{-\int_0^x \lambda(1-y) + \mu_2(\xi) d\xi} dx,$$

C is a constant and $\Pi_0(1) = C$. We define $\Pi_3(z) = \int_0^{\infty} \Pi_3(x, z) dx$. Then, applying the normalizing condition $\Pi_0(1) + \Pi_1(1) + \Pi_2(1) + \Pi_3(1) = 1$, we obtain

$$C = \frac{\lambda \alpha (1 - \lambda \beta_1)}{\lambda \alpha + \alpha \beta + \lambda \beta + (1 - \lambda \beta_1) \beta \mu_1}, \quad \beta_1 = \int_0^{\infty} e^{-\int_0^x \mu_2(\xi) d\xi} dx.$$

Let $K(t)$ be the number of customers in the system. Then the probability generating function of $K(t)$ is $\Pi(z) = \Pi_0(z) + z \Pi_1(z) + \Pi_2(z) + z \Pi_3(z)$, and the mean number of

customers in the system is $E[K] = \partial\Pi(z)/\partial z|_{z=1}$. By denoting the probabilities of the system in state 0, 1, 2 and 3 as P_0, P_1, P_2 and P_3 , respectively, we derive

$$\left\{ \begin{aligned} P_0 &= \frac{\lambda\alpha(1 - \lambda\beta_1)}{\lambda\alpha + \alpha\beta + \lambda\beta + (1 - \lambda\beta_1)\beta\mu_1}, & P_1 &= \frac{\lambda\beta(1 - \lambda\beta_1)}{\lambda\alpha + \alpha\beta + \lambda\beta + (1 - \lambda\beta_1)\beta\mu_1}, \\ P_2 &= \frac{\beta(1 - \lambda\beta_1)(\alpha + \mu_1)}{\lambda\alpha + \alpha\beta + \lambda\beta + (1 - \lambda\beta_1)\beta\mu_1}, & P_3 &= \frac{\lambda\beta_1(\lambda\alpha + \alpha\beta + \lambda\beta)}{\lambda\alpha + \alpha\beta + \lambda\beta + (1 - \lambda\beta_1)\beta\mu_1}, \\ E_0[K] &= \frac{\alpha(\lambda\alpha + \alpha\beta + \lambda\beta)(\beta_1(1 - \lambda\beta_1) + \lambda\beta_2)}{\alpha(1 - \lambda\beta_1)[\lambda\alpha + \alpha\beta + \lambda\beta + (1 - \lambda\beta_1)\beta\mu_1]} + \frac{\beta + \alpha\beta\beta_1 + \lambda\alpha\beta_1}{\alpha\theta(1 - \lambda\beta_1)} \\ &\quad + \frac{[\beta(\lambda + \alpha) + \mu_1\beta(1 - \lambda\beta_1)](1 - \lambda\beta_1)}{\alpha(1 - \lambda\beta_1)[\lambda\alpha + \alpha\beta + \lambda\beta + (1 - \lambda\beta_1)\beta\mu_1]}, \end{aligned} \right. \tag{4.4}$$

where $\beta_2 = \int_0^\infty xe^{-\int_0^x \mu_2(\xi)d\xi} dx$.

To show the effectiveness of the proposed system, we consider three special cases of the system as follows.

Case 1: $f(\alpha) \equiv 0, \mu_1 \neq 0$ and $\beta \neq 0$. This means that the system does not perform the ISEV policy, the M/G/1 retrial queue system with RWV and ISEV policies will degrade into an M/G/1 retrial queue system with RWV policy. We denote the stationary probabilities of the system in state 0, 1, 2 and 3 as $P_0^{(1)}, P_1^{(1)}, P_2^{(1)}$ and $P_3^{(1)}$, respectively. Then, the stationary probabilities of the system and the mean number of customers in the system $E_1[K]$ can be obtained according to (4.4), that is,

$$\left\{ \begin{aligned} P_0^{(1)} &= \frac{\lambda\alpha(1 - \lambda\beta'_1)}{\lambda\alpha + \alpha\beta + \lambda\beta + (1 - \lambda\beta'_1)\beta\mu_1}, & P_1^{(1)} &= \frac{\lambda\beta(1 - \lambda\beta'_1)}{\lambda\alpha + \alpha\beta + \lambda\beta + (1 - \lambda\beta'_1)\beta\mu_1}, \\ P_2^{(1)} &= \frac{\beta(1 - \lambda\beta'_1)(\alpha + \mu_1)}{\lambda\alpha + \alpha\beta + \lambda\beta + (1 - \lambda\beta'_1)\beta\mu_1}, & P_3^{(1)} &= \frac{\lambda\beta_1(\lambda\alpha + \alpha\beta + \lambda\beta)}{\lambda\alpha + \alpha\beta + \lambda\beta + (1 - \lambda\beta'_1)\beta\mu_1}, \\ E_0[K] &= \frac{\alpha(\lambda\alpha + \alpha\beta + \lambda\beta)(\beta'_1(1 - \lambda\beta'_1) + \lambda\beta_2)}{\alpha(1 - \lambda\beta'_1)[\lambda\alpha + \alpha\beta + \lambda\beta + (1 - \lambda\beta'_1)\beta\mu_1]} + \frac{\beta + \alpha\beta\beta'_1 + \lambda\alpha\beta'_1}{\alpha\theta(1 - \lambda\beta'_1)} \\ &\quad + \frac{[\beta(\lambda + \alpha) + \mu_1\beta(1 - \lambda\beta'_1)](1 - \lambda\beta'_1)}{\alpha(1 - \lambda\beta'_1)[\lambda\alpha + \alpha\beta + \lambda\beta + (1 - \lambda\beta'_1)\beta\mu_1]}, \end{aligned} \right. \tag{4.5}$$

where $\beta'_1 = \int_0^\infty e^{-\int_0^x \mu_2^-(\xi)d\xi} dx$ and $\beta'_2 = \int_0^\infty xe^{-\int_0^x \mu_2^-(\xi)d\xi} dx$.

Case 2: $f(\alpha) \equiv 0, \mu_1 = 0$ and $\beta \neq 0$. The system does not utilize the ISEV policy, and the working vacation state changes to a normal vacation. In this case, the M/G/1 retrial queue system with RWV and ISEV policies will degrade to an M/G/1 retrial queue system with reserved idle time and setup time, which has been previously studied by Zhang and Wang [31]. Furthermore, the states of the server $C(t) = k$,

where $k = 0, 1, 2, 3$ at time t , will transform to indicate idle, setup, off and busy states, respectively. We denote the stationary probabilities of the system being idle, setup, off and busy as $P_0^{(2)}, P_1^{(2)}, P_2^{(2)}$ and $P_3^{(2)}$, respectively. The stationary probability of the system and the mean number of customers in the system $E_2[K]$ can be expressed as follows:

$$\left\{ \begin{aligned} P_0^{(2)} &= \frac{\lambda\alpha(1 - \lambda\beta'_1)}{\lambda\alpha + \alpha\beta + \lambda\beta}, & P_1^{(2)} &= \frac{\lambda\beta(1 - \lambda\beta'_1)}{\lambda\alpha + \alpha\beta + \lambda\beta}, \\ P_2^{(2)} &= \frac{\alpha\beta(1 - \lambda\beta'_1)}{\lambda\alpha + \alpha\beta + \lambda\beta}, & P_3^{(2)} &= \lambda\beta'_1, \\ E_2[K] &= \frac{\alpha(\lambda\alpha + \alpha\beta + \lambda\beta)(\beta'_1(1 - \lambda\beta'_1) + \lambda\beta'_2) + \beta(\lambda + \alpha)(1 - \lambda\beta'_1)}{\alpha(1 - \lambda\beta'_1)(\lambda\alpha + \alpha\beta + \lambda\beta)} \\ &\quad + \frac{\beta + \alpha\beta\beta'_1 + \lambda\alpha\beta'_1}{\alpha\theta(1 - \lambda\beta'_1)}, \end{aligned} \right. \tag{4.6}$$

which agrees with the conclusions of Zhang and Wang [31].

Case 3: $f(\alpha) \equiv 0, \mu_1 = 0$ and $\beta = 0$. This means that the server will never take a vacation, so the M/G/1 retrial queue system with RWV and ISEV policy will degrade into a classical M/G/1 retrial queue system. The probability of the system in idle $P_0^{(3)}$ and in busy $P_1^{(3)}$ and the mean number of customers in the system $E_3[K]$ can be derived, that is,

$$\left\{ \begin{aligned} P_0^{(3)} &= 1 - \lambda\beta'_1, & P_1^{(3)} &= \lambda\beta'_1, \\ E_3[K] &= \frac{\beta'_1(\lambda + \theta) + \lambda\theta\beta'_2 - \lambda\theta\beta'^2_1}{\theta(1 - \lambda\beta'_1)}, \end{aligned} \right. \tag{4.7}$$

which also agrees with the conclusion of Falin and Templeton [12].

4.2. System performance measures analysis Energy consumption, service efficiency and expected cost are the three most critical performance measures used to evaluate the quality of a queue system. In this subsection, we define and discuss the system’s performance measures for the retrial queue system proposed in this paper, as well as its special cases. Before delving into this, we introduce some notation for the systems considered in Section 4.1 for convenience.

- *sys.0:* The M/G/1 retrial queue system with RWV and ISEV policy proposed in this paper.
- *sys.1:* The M/G/1 retrial queue system with RWV policy in Case 1.
- *sys.2:* The M/G/1 retrial queue system with a reserved idle time and setup time in Case 2, which was studied by Zhang and Wang [31].
- *sys.3:* The M/G/1 retrial queue system in Case 3, which was studied by Falin and Templeton [12].

In this paper, we measure the service efficiency of the systems using the mean number of customers in the system. As indicated in Section 4.1, the mean number of customers in the system are denoted by $E_i[K]$, where $i = 0, 1, 2, 3$. These values are given by (4.4)–(4.7).

To discuss the energy consumption and expected cost of the systems, we introduce the following definitions:

- $C_i, i = rwb, rwi$, represents the energy consumption of the server in the regular working period when the server is busy and idle, respectively;
- $C_j, j = wvb, wvi$, represents the energy consumption of the server in the working vacation period when the server is busy and idle, respectively;
- C_s represents the energy consumption when the server is in setup state;
- c_h represents the holding cost for each customers present in the system (including the customers in the orbit);
- c_{sys} represents the cost per unit energy consumption of the server;
- $c_m(\alpha)$ represents the cost function of the server being maintained during the working vacation period. Here, $c_m(\alpha)$ is assumed to be a monotonically increasing bounded function of α .

To ensure that the definitions for the energy and expected cost functions are realistic, we assume that the following relationship is satisfied:

$$C_{rwb} > C_{wvb} > C_{rwi} > C_{wvi} > C_s, \quad c_h > c_{sys}.$$

Then, by denoting the energy consumption and expected cost of the $sys.i$, which were defined at the beginning of this subsection, as $EC_i, SC_i, i = 0, 1, 2, 3$, respectively, we deduce

$$\begin{cases} EC_0 = C_{rwi}P_0 + C_{wvb}P_1 + C_{wvi}P_2 + C_{rwb}P_3, \\ EC_1 = C_{rwi}P_0^{(1)} + C_{wvb}P_1^{(1)} + C_{wvi}P_2^{(1)} + C_{rwb}P_3^{(1)}, \\ EC_2 = C_{rwi}P_0^{(2)} + C_sP_1^{(2)} + C_{rwb}P_3^{(2)}, \\ EC_3 = C_{rwi}P_0^{(3)} + C_{rwb}P_1^{(3)}, \\ SC_0 = c_{sys}EC_0 + c_hE_0[K] + c_m(\alpha), \\ SC_1 = c_{sys}EC_1 + c_hE_1[K], \\ SC_2 = c_{sys}EC_2 + c_hE_2[K], \\ SC_3 = c_{sys}EC_3 + c_hE_3[K]. \end{cases}$$

4.3. Optimal queuing strategies investigation Based on the definition provided before, we investigate optimal queuing strategies in this subsection. Before delving into this, we analyse the performance of the systems under different actual arrival rates λ to demonstrate the advantages of the M/G/1 retrial queue system with RWV and ISEV

policies. Specifically, we analyse the energy consumption and service efficiency of the system with different λ , while keeping all other system parameters fixed.

THEOREM 4.1. *For the M/G/1 retrial queue system with RWV and ISEV policy (2.1), there exists a unique $\lambda^* \in (0, 1/\beta'_1)$. If $\lambda^* < k' = \min\{1/\beta'_1, \alpha\}$, then λ^* is called the equilibrium arrival rate, and for any $\lambda \in (\lambda^*, k')$, the energy consumption and the mean number of customers in sys.0, sys.1, sys.2 and sys.3 have the following properties:*

$$\begin{cases} EC_0 < EC_2 < EC_1 < EC_3, \\ E_0[K] < E_3[K] < E_1[K] < E_2[K], \end{cases} \tag{4.8}$$

which means that for any $\lambda \in (\lambda^*, k')$, the M/G/1 retrial queue system with RWV and ISEV policies can not only consume the least energy, but also provide the highest service efficiency than other systems.

PROOF. According to (4.4)–(4.7), we can derive that for any $\lambda < 1/\beta'_1$, $E_3[K] < E_1[K] < E_2[K]$ and $EC_2 < EC_1 < EC_3$. Let $G_1(\lambda) = E_0[K] - E_3[K]$ and $G_2(\lambda) = EC_0 - EC_2$. Note that the real value functions $G_1(\lambda)$ and $G_2(\lambda)$ on λ are continuous functions defined in $(0, 1/\beta'_1)$. Thus, by performing careful calculations, we can verify that the first-order derivatives of $G_1(\lambda)$ and $G_2(\lambda)$ are negative for any $\lambda \in (0, 1/\beta'_1)$. Moreover, we deduce that

$$\begin{cases} \lim_{\lambda \rightarrow 0^+} G_1(\lambda) = \frac{\beta\alpha + \beta\mu_1(1 - \beta_1)}{\alpha(\alpha\beta + \beta\mu_1)} + \frac{\beta + \alpha\beta\beta_1}{\alpha\theta} > 0, \\ \lim_{\lambda \rightarrow \{1/\beta'_1\}^-} G_1(\lambda) = -\infty, \\ \lim_{\lambda \rightarrow 0^+} G_2(\lambda) = C_{wvi} - C_{rvi} > 0, \\ \lim_{\lambda \rightarrow \{1/\beta'_1\}^-} G_2(\lambda) = (EC_0 |_{\lambda=1/\beta'_1} - 1)C_{rvb} < 0. \end{cases}$$

Hence, $G_1(\lambda)$ and $G_2(\lambda)$ are decreasing in $(0, 1/\beta'_1)$, and there exists a unique solution λ_1 such that $G(\lambda_1) = 0$, and a unique solution λ_2 such that $G(\lambda_2) = 0$. Let $\lambda = \max\{\lambda_1, \lambda_2\}$; if $\lambda < k'$, then all $\lambda \in (\lambda, k')$ can ensure the stability of the system model and satisfy the inequalities $E_0[K] - E_2[K] < 0$ and $EC_0 - EC_1 < 0$. Thus, inequality (4.8) holds and λ is called the equilibrium arrival rate. However, if $\lambda \geq k$, although equations $G(\lambda_1) = 0$ and $G(\lambda_2) = 0$ have a unique solution, the system is not necessarily stable when $\lambda \in (k', \lambda)$. □

Theorem 4.1 establishes that given that the system is stable, the M/G/1 retrial queue system with RWV and ISEV policies outperform other systems in terms of both energy consumption and service efficiency, provided that the actual arrival rate $\lambda > \lambda^*$.

While Theorem 4.1 highlights the advantages of the system proposed in this paper, the focus of management design is typically on maximizing service efficiency or

minimizing expected cost. To achieve this, we need to investigate the optimal queuing strategies for the system.

THEOREM 4.2. *For the M/G/1 retrial queue system with RWV and ISEV policy (2.1), there exists an optimal vacation interruption rate α_1 such that*

$$E_0[K] |_{\alpha=\alpha_1} = \min_{\alpha \in Z} E_0[K],$$

and there exists an optimal vacation interruption rate α_2 :

$$SC_0 |_{\alpha=\alpha_2} = \min_{\alpha \in Z} SC_0,$$

where $Z = \{\alpha \mid \lambda \leq \alpha, \lambda\beta_1 < 1\}$.

PROOF. We analyse the range of α to discuss the existence of the optimal queuing strategies. According to Section 3, the equilibrium threshold condition of the system to be stable is $\lambda \leq \alpha$ and

$$\lambda\beta_1 = \lambda \int_0^\infty e^{-\int_0^x (1+f(\alpha))\mu_2^-(\xi) d\xi} dx < 1.$$

For any fixed arrival rate λ , if

$$\lambda\beta'_1 = \lambda \int_0^\infty e^{-\int_0^x \mu_2^-(\xi) d\xi} dx < 1,$$

then for any $\alpha \geq 0, \lambda\beta_1 < 1$ and the system (2.2) is stable, which means that $Z = [\lambda, \infty)$. However, if $\lambda\beta'_1 \geq 1$, by taking ISEV policy, we can reasonably adjust the vacation length so that $\lambda\beta_1 < 1$. Since $f(\alpha)$ is a monotonically decreasing function with respect to $\alpha \geq 0$, there exists a unique solution α' of equation $\lambda\beta = 1$. Thus, the range of α for the system to be stable is $Z = [\lambda, \alpha')$ if $\alpha' > \lambda$. Otherwise, $Z = \emptyset$.

Here, we analyse the situation of $\lambda\beta'_1 \geq 1$; the situation of $\lambda\beta'_1 < 1$ can be derived in the same way. We study the optimal queuing strategies of the system on the service efficiency in $Z = [\lambda, \alpha')$. The first-order derivative of $E_0[K]$ with α is

$$\frac{d}{d\alpha} E_0[K] = \frac{I_1 + I_2 + I_3}{J^2},$$

where

$$\begin{aligned} I_1 &= \alpha^2[S(1 - \lambda\beta_1)(S + Q)[(1 - 2\lambda + \lambda\beta_1)\beta'_1 + \lambda\beta'_2] \\ &\quad + \lambda^2S(S + Q)\beta_2 + (\lambda S\beta'_1 + (\lambda + \beta)(1 - \lambda\beta_1))OQ]\theta^2, \\ I_2 &= [-\lambda\alpha\beta'_1(1 - \lambda\beta_1)(S + Q - \lambda\beta)Q - (1 - \lambda\beta_1)^2[(S + Q)^2 - \lambda\alpha(S - \lambda\beta)]]\theta^2, \\ I_3 &= [(1 - \lambda\beta_1)(\alpha(\lambda\alpha + \alpha\beta)\beta'_1 - \beta) + \lambda\alpha\beta'_1R](S + Q)^2\theta, \\ J &= \alpha(1 - \lambda\beta_1)(S + Q)\theta, \end{aligned}$$

and

$$\begin{aligned}
 S &= \lambda\alpha + \alpha\beta + \lambda\beta, & Q &= (1 - \lambda\beta_1)\beta\mu_1, \\
 O &= \beta_1(1 - \lambda\beta_1) + \lambda\beta_2, & R &= \beta + \alpha\beta\beta_1 + \lambda\alpha\beta_1.
 \end{aligned}$$

To derive the unique optimal vacation interruption rate α_1^* on the service efficiency, the solution of equation $(d/d\alpha)E_0[K] = 0$ should be considered. By taking careful calculation, we can verify that the first-order derivatives of the real value function $g(\alpha)$ are monotonically increasing in $(0, \alpha')$. Then, we can verify that

$$\begin{cases} \lim_{\alpha \rightarrow 0^+} \frac{d}{d\alpha} E_0[K] = -\infty, \\ \lim_{\alpha \rightarrow \alpha'^-} \frac{d}{d\alpha} E_0[K] = \lim_{\lambda\beta_1 \rightarrow 1} \frac{d}{d\alpha} E_0[K] = +\infty, \end{cases}$$

which indicates that equation $(d/d\alpha)E_0[K] = 0$ has unique solution $\alpha_1 \in (0, \alpha')$. Thus, if $\alpha_1^0 > \lambda$, it means that the mean sojourn time of the customer in the orbit $E_0[K]$ will decrease first in (λ, α_1^0) and then increase in (α_1^0, α') . Thus, the optimal vacation interruption rate to maximize the service efficiency is $\alpha_1^* = \alpha_1^0$. However, if $\alpha_1^0 < \lambda$, then $E_0[K]$ will monotonically increase in $[\lambda, \alpha')$, and the optimal vacation interruption rate to maximize the service efficiency is $\alpha_1^* = \lambda$.

Therefore, by taking into account the above two situations, we can demonstrate that there exists an optimal vacation interruption rate α_1^* such that

$$E_0[K] |_{\alpha=\alpha_1^*} = \min_{\alpha \in Z} E_0[K].$$

However, we can investigate the optimal queuing strategies on the expected cost of the system (2.2) by using the similar method. Therefore, we also deduce that there exists an optimal vacation interruption rate α_2^* such that

$$SC_0 |_{\alpha=\alpha_2^*} = \min_{\alpha \in Z} SC_0. \quad \square$$

According to Theorem 4.2, there exist optimal queuing strategies that can improve the efficiency and reduce the expected cost. Furthermore, it is demonstrated that adjusting the vacation interruption rate α can help identify the optimal strategy.

5. Numerical experiments

In this section, we present two numerical experiments to illustrate the theoretical results from the perspectives of energy consumption, service efficiency and expected cost.

We assume that the service time of the server in the regular working period without ISEV policy follows the Erlang-2 distribution [11], with a probability density function $g(x) = \phi^2 x e^{-\phi x}$, where ϕ is a constant. Thus, we can deduce the service completion rate of the server under the ISEV policy as $\mu_2(x) = (1 + f(\alpha))(\phi^2 x / (\phi x + 1))$. Taking $\phi = 2$ and $f(\alpha) = 1/(\alpha^2 + 2)$, we can deduce $\beta_1' = 1$. Using the values of the system parameters:

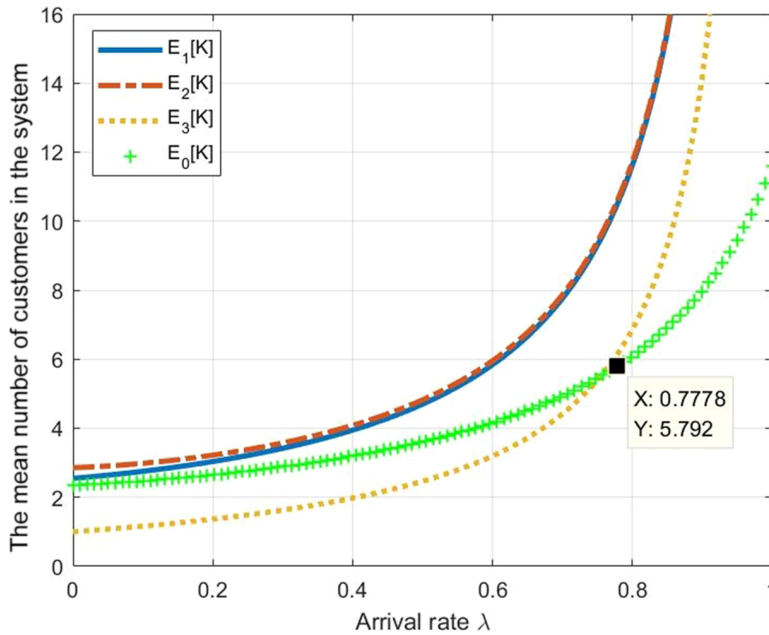


FIGURE 3. Mean number of customers in the system with different arrival rate λ .

$$C_{rwb} = 350, \quad C_{rwi} = 100, \quad C_{wvb} = 150, \quad C_{wvi} = 50,$$

$$C_s = 45, \quad c_h = 5, \quad c_{sys} = 1, \quad c_m(\alpha) = 180\left(\frac{1}{\alpha^2 + 2}\right),$$

we conduct the following numerical experiments.

EXAMPLE 5.1. Assuming $\beta = 0.6$, $\theta = 1.4$, $\alpha = 1$ and $\mu_1 = 0.45$, we can use the theoretical analysis in Section 3 to deduce that $\lambda < 1$ will ensure the asymptotical stability of $sys.i, i = 0, 1, 2, 3$. The service efficiency and energy consumption of the systems under different arrival rates λ are illustrated in Figures 3 and 4, respectively.

As shown in Figures 3 and 4, for any $\lambda < 1$, the RWV policy can improve service efficiency but consume more energy, as evidenced by $E_1[K] < E_2[K]$ and $EC_1 > EC_2$. Additionally, for any $\lambda > 0.7778$, the mean number of customers in $sys.0$ (the “+” line) is smaller than other systems, as illustrated in Figure 3. Moreover, for any $\lambda > 0.322$, the energy consumption of $sys.0$ (the “+” line) is lower than other systems, as shown in Figure 4. Thus, for any $\lambda > \max\{0.7778, 0.322\}$, the mean number of customers and energy consumption of $sys.0$ are both the lowest, in agreement with the conclusion of Theorem 4.1.

EXAMPLE 5.2. Figures 5 and 6 present the mean number of customers in the system and the expected cost of $sys.0$ on the vacation interruption rate α , when the actual

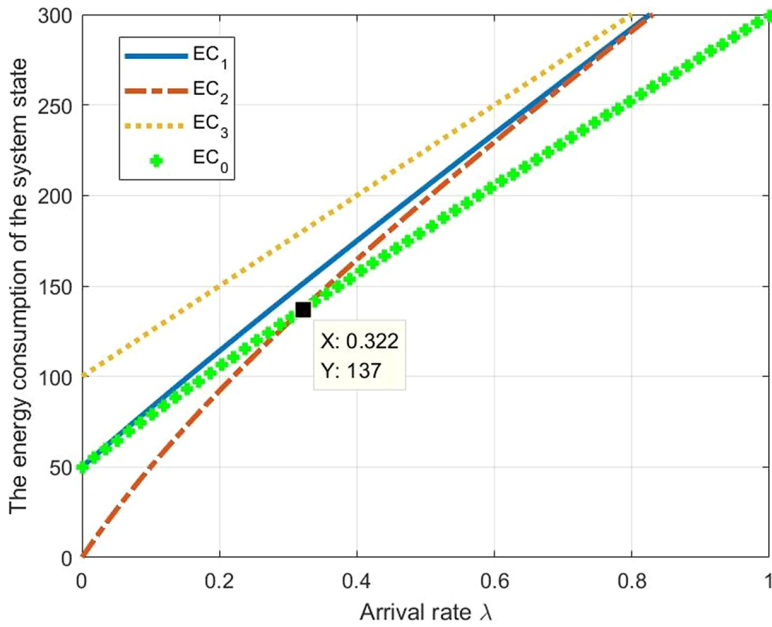


FIGURE 4. Consumption of systems state with different arrival rate λ .

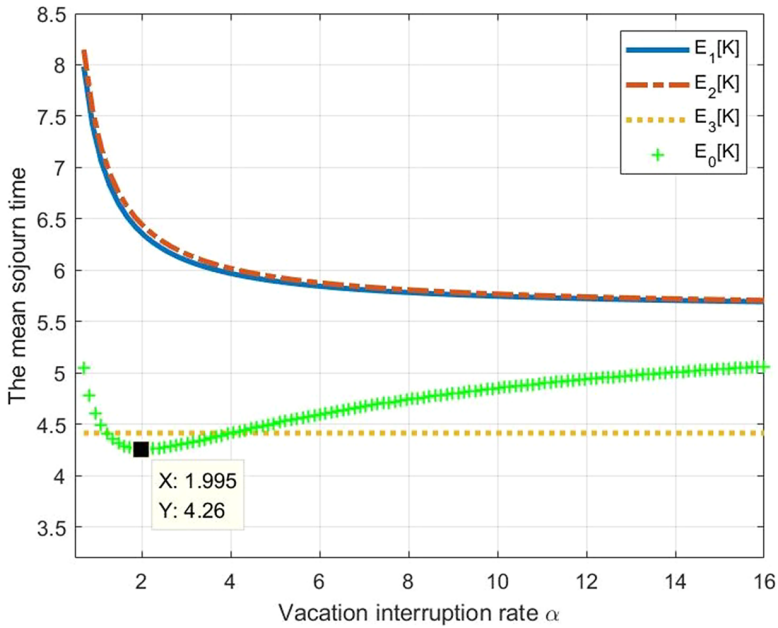


FIGURE 5. Mean number of customers in the system with different α .

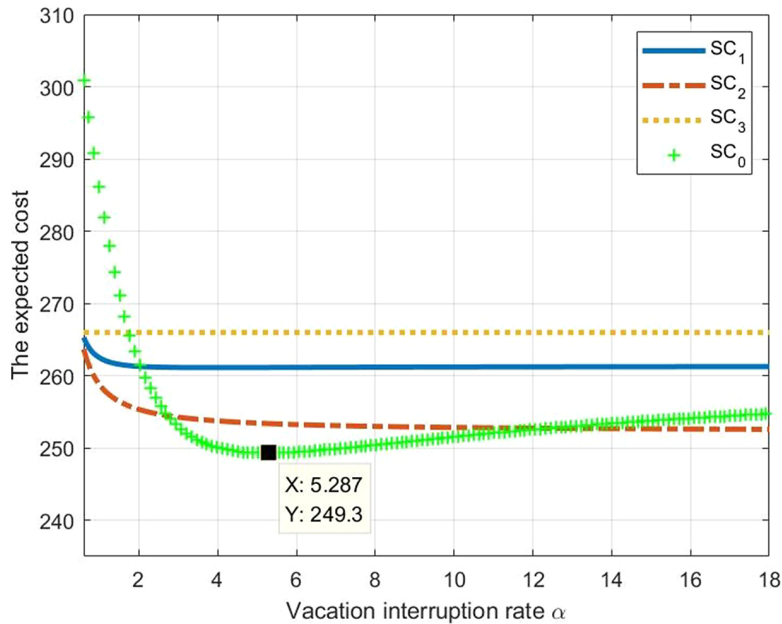


FIGURE 6. Expected cost of systems with different α .

arrival rate is $\lambda = 0.7$, $\beta = 0.5$, $\theta = 1.4$ and $\mu_1 = 0.45$. Figure 5 shows that the optimal vacation interruption rate $\alpha_1^* = 1.995$, which minimizes the mean number of customers in *sys.0* (the “+” line) with respect to service efficiency. However, Figure 6 shows that the optimal vacation interruption rate $\alpha_2^* = 5.287$, which minimizes the expected cost of *sys.0* (the “+” line). Example 5.2 directly follows from Theorem 4.2.

6. Conclusion

In this study, we proposed an M/G/1 retrial queue system with RWV and ISEV policies and investigated its optimal queuing strategies. First, we studied the dynamic behaviour of the system and derived the equilibrium threshold condition for the system to be stable. Then, we obtained some explicit expressions for the steady-state indices of the system and its special cases. Additionally, we obtained some important system performance measures, including energy consumption, service efficiency and expected cost. We also investigated the optimal queuing strategies of the system, and proved that the proposed system outperforms other systems by consuming the least energy while providing the highest service efficiency. Meanwhile, we demonstrated the existence of optimal strategies for improving efficiency and reducing expected costs, which can be obtained by adjusting the vacation interruption rate. Finally, we validated our analytical results through numerical experiments.

The proposed model is a generalized version of many existing queue models with vacations and vacation interruption policies, and has potential real-life applications in

ICT systems. Our investigation may provide valuable insights for system managers to design and optimize their systems for energy savings and efficiency improvement.

Appendix A. Proof of Step 2 in Theorem 3.2

First, we consider the equation $(A + B + C)\vec{P} = 0$. That is,

$$(\lambda + \beta + n\theta)P_{0,n} - \int_0^\infty \mu_2(x)P_{3,n}(x) dx = 0, \tag{A.1}$$

$$(\lambda + \alpha + \mu_1)P_{1,n} - \lambda P_{2,n} - \lambda P_{1,n-1} = 0, \tag{A.2}$$

$$\lambda P_{2,n} - \beta P_{0,n} - \mu_1 P_{1,n} = 0, \tag{A.3}$$

$$\frac{d}{dx}P_{3,n}(x) + (\lambda + \mu_2(x))P_{3,n}(x) - \lambda P_{3,n-1}(x) = 0, \tag{A.4}$$

where $n = 0, 1, 2, \dots$, and $P_{1,-1} = 0, P_{3,-1}(x) = 0$. The proof can be divided into two parts. First, we prove that $\{P_{0,n}, n = 0, 1, 2 \dots\}$ is a positive sequence. Second, we prove that $\sum_{n=0}^\infty P_{0,n} < \infty$ if

$$\lambda \int_0^\infty e^{-\int_0^x \mu_2(\xi) d\xi} dx < 1 \quad \text{and} \quad \lambda < \alpha.$$

For the first part, we obtain the following equations by careful calculation:

$$P_{1,n} = \frac{\lambda^n \beta}{(\lambda + \alpha)^{n+1}} P_{0,0} + \dots + \frac{\beta}{\lambda + \alpha} P_{0,n}, \tag{A.5}$$

$$P_{2,n} = \frac{\mu_1}{\lambda} \left[\frac{\lambda^n \beta}{(\lambda + \alpha)^{n+1}} P_{0,0} + \dots + \frac{\beta}{\lambda + \alpha} P_{0,n} \right] + \frac{\beta P_{0,n}}{\lambda}, \tag{A.6}$$

$$P_{3,n}(x) = P_{3,n}(0)g_0(x) + \int_0^x \lambda P_{3,n-1}(v)g_1(x, v) dv, \tag{A.7}$$

and

$$\begin{aligned} (n + 1)\theta P_{0,n+1} = & \frac{1}{\int_0^\infty \mu_2(x)g_0(x) dx} \int_0^\infty \mu_2(x) \left[\sum_{j=0}^n \sum_{i=0}^j \frac{(\lambda x)^{n-j}}{(n-j)!} \beta \left(\frac{\lambda}{\lambda + \alpha} \right)^{j-i+1} P_{0,i} \right. \\ & \left. + \sum_{i=0}^{n-1} \frac{(\lambda x)^{n-i}}{(n-i)!} (\lambda + \beta)P_{0,i} + \sum_{j=0}^n (\lambda + \beta + j\theta)R_j P_{0,j} \right] g_0(x) dx, \end{aligned} \tag{A.8}$$

where

$$\begin{aligned} g_0(x) = e^{-\int_0^x (\lambda + \mu_2(\xi)) d\xi}, \quad g_1(x, v) = e^{-\int_v^x (\lambda + \mu_2(\xi)) d\xi}, \\ R_0 = \sum_{j=n+1}^\infty \frac{(\lambda x)^j}{j!}, \quad R_i = \sum_{j=n+2-i}^\infty \frac{(\lambda x)^j}{j!}, \quad 0 < i \leq n. \end{aligned}$$

Equations (A.8) and (A.5)–(A.7) imply that for any fixed n , the solutions of (A.1)–(A.4) are positive if $P_{0,0}$ is positive. Obviously, $P_{0,0} > 0$ can be derived by taking $n = 0$ into (A.1)–(A.4).

For the second part, we provide the following notation for convenience. Let

$$b_n = \int_0^\infty \frac{1}{(n-1)!} (\lambda x)^{n-1} \mu_2(x) e^{-\int_0^x (\lambda + \mu_2(\xi)) d\xi} dx.$$

According to Zheng et al. [32, Theorem 2.1], we obtain the following estimates:

$$\begin{aligned} \sum_{i=1}^\infty b_i &= 1, \quad b_1 \geq \frac{\mu}{\lambda + \mu}, \quad \sum_{i=n}^\infty b_i \leq \left(\frac{\lambda}{\lambda + \mu}\right)^{n-1}, \\ \frac{\sum_{i=n}^\infty b_i}{b_1} &\leq \frac{\lambda}{\mu} \left(\frac{\lambda}{\lambda + \mu}\right)^{n-2}, \quad n \geq 2, \end{aligned} \tag{A.9}$$

where $0 \leq (1/x) \int_0^x \mu_2(\tau) d\tau \leq \mu$ for any $x \in [0, \infty)$. According to (A.8) and (A.9),

$$\begin{aligned} P_{0,n+1} &< \frac{1}{n+1} \left[\left(\frac{\beta}{\theta} \frac{\lambda}{\lambda + \alpha} + \frac{\lambda + \beta \lambda}{\theta \mu} + n \frac{\lambda}{\mu} \frac{\lambda}{\lambda + \mu} \right) P_{0,n} \right. \\ &+ \left(\left(\frac{\beta}{\theta} \left(\frac{\lambda}{\lambda + \alpha} \right)^2 \left(1 + \frac{\lambda}{\lambda + \mu} \right) + \frac{\lambda + \beta \lambda}{\theta \mu} \frac{\lambda}{\lambda + \mu} + (n-1) \frac{\lambda}{\mu} \left(\frac{\lambda}{\lambda + \mu} \right)^2 \right) P_{0,n-1} \right. \\ &+ \dots \\ &\left. + \left(\frac{\beta}{\theta} \left(\frac{\lambda}{\lambda + \alpha} \right)^{n+1} \left(1 + \frac{\lambda}{\lambda + \mu} + \dots + \left(\frac{\lambda}{\lambda + \mu} \right)^n \right) + \frac{\lambda + \beta \lambda}{\theta \mu} \left(\frac{\lambda}{\lambda + \mu} \right)^n \right) P_{0,0} \right]. \end{aligned}$$

For convenience, let

$$\begin{aligned} \varphi_i^{(n+1)} &= \frac{\beta}{\theta} \left(\frac{\lambda}{\lambda + \alpha} \right)^i \left(1 + \frac{\lambda}{\lambda + \mu} + \dots + \left(\frac{\lambda}{\lambda + \mu} \right)^{i-1} \right) \\ &+ \frac{\lambda + \beta \lambda}{\theta \mu} \left(\frac{\lambda}{\lambda + \mu} \right)^{i-1} + (n+1-i) \frac{\lambda}{\mu} \left(\frac{\lambda}{\lambda + \mu} \right)^i, \end{aligned}$$

and denote $k = \min\{\mu, \alpha\}$, $K = \max\{\mu, \alpha\}$.

For any fixed n , let

$$\delta_1^{(n+1)} = \frac{\lambda}{\lambda + k} + f_{\max}^{(n+1)},$$

where $f_{\max}^{(n+1)}$ is the maximum of the continuous function

$$f^{(n+1)}(i) = \frac{1}{n+1-i} \frac{\beta K}{\theta \lambda} \left(\frac{\lambda}{\lambda + \alpha} \right)^{i+1}$$

in the interval $[1, n+1]$. It is not difficult to verify that $\{f_{\max}^{(n+1)}\}$ is a positive sequence and that it monotonically decreases toward 0. Thus, we can deduce that for any fixed n

and $i \in [1, n]$, the following inequality holds:

$$\begin{aligned} \delta_1^{(n+1)} \varphi_i^{(n+1)} - \varphi_{i+1}^{(n+1)} &= (n + 1 - i) \frac{\lambda}{\mu} \left(\frac{\lambda}{\lambda + \mu} \right)^{i+1} [f_{\max}^{(n+1)} - f^{(n+1)}(i)] \\ &+ \frac{\beta}{\theta} \frac{\lambda}{\mu} \left(\frac{\lambda + \mu}{\lambda + k} - \frac{\lambda + \mu}{\lambda + \alpha} \right) \left(\frac{\lambda}{\lambda + \alpha} \right)^i + \frac{\lambda + \beta}{\theta} \frac{\lambda}{\mu} \left(\frac{\lambda}{\lambda + k} - \frac{\lambda}{\lambda + \mu} \right) \\ &+ (n - i) \frac{\lambda}{\mu} \left(\frac{\lambda}{\lambda + k} - \frac{\lambda}{\lambda + \mu} \right) \left(\frac{\lambda}{\lambda + \mu} \right)^i + \frac{\lambda}{\mu} \frac{\lambda}{\lambda + k} \left(\frac{\lambda}{\lambda + \mu} \right)^i \\ &+ f_{\max}^{(n+1)} \left(\frac{\beta}{\theta} \frac{\lambda + \mu}{\mu} \left(\frac{\lambda}{\lambda + \alpha} \right)^i \left(1 - \left(\frac{\lambda}{\lambda + \mu} \right)^i \right) + \frac{\lambda + \beta}{\theta} \frac{\lambda}{\mu} \left(\frac{\lambda}{\lambda + \mu} \right)^{i-1} \right) > 0. \end{aligned}$$

Therefore, we can derive that

$$\begin{aligned} P_{0,n+1} &< \frac{1}{n + 1} (\varphi_1^{(n+1)} P_{0,n} + \delta_1^{(n+1)} \varphi_1^{(n+1)} P_{0,n-1} + \delta_1^{(n+1)} \varphi_2^{(n+1)} P_{0,n-1} \\ &+ \dots + \delta_1^{(n+1)} \varphi_n^{(n+1)} P_{0,0}) \\ &< \frac{\varphi_1^{(n+1)}}{n + 1} (P_{0,n} + \delta_1^{(n+1)} P_{0,n-1} + (\delta_1^{(n+1)})^2 P_{0,n-2} + \dots + (\delta_1^{(n+1)})^{n+1} P_{0,0}) \\ &< \frac{\varphi_1^{(n+1)}}{n + 1} \left(\delta_1^{(n+1)} + \frac{\varphi_1^{(n)}}{n} \right) \left(\delta_1^{(n)} + \frac{\varphi_1^{(n-1)}}{n - 1} \right) \dots (\delta_1^{(2)} + \varphi_1^{(1)}) P_{0,0}. \end{aligned}$$

Because the positive sequence $\{\delta_1^{(n+1)} + \varphi_1^{(n)}/n\}$ is monotonically decreasing, and that conditions $\lambda \int_0^\infty e^{-\int_0^x \mu_2(\xi) d\xi} dx < 1$ and $\lambda \leq \alpha$ hold,

$$\lim_{n \rightarrow \infty} \left(\delta_1^{(n+1)} + \frac{\varphi_1^{(n)}}{n} \right) = \frac{\lambda}{\lambda + k} + \frac{\lambda}{\mu} \frac{\lambda}{\lambda + \mu} < 1,$$

which implies that there must exist an $N > 0$ such that $\delta_1^{(n+1)} + \varphi_1^{(n)}/n < 1$ for any $n > N$. Therefore, it can be deduced that for any $n > N$,

$$\begin{aligned} P_{0,n+1} &< \frac{\varphi_1^{(n+1)}}{n + 1} \left(\delta_1^{(n+1)} + \frac{\varphi_1^{(n)}}{n} \right) \left(\delta_1^{(n)} + \frac{\varphi_1^{(n-1)}}{n - 1} \right) \dots (\delta_1^{(1)} + \varphi_1^{(0)}) P_{0,0} \\ &< \frac{\varphi_1^{(n+1)}}{n + 1} \left(\delta_1^{(N+2)} + \frac{\varphi_1^{(N+1)}}{N + 1} \right)^{n+1-N} \prod_{i=1}^N \left(\delta_1^{(i+1)} + \frac{\varphi_1^{(i)}}{i} \right) P_{0,0} \\ &< \left(\delta_1^{(N+2)} + \frac{\varphi_1^{(N+1)}}{N + 1} \right)^{n+1-N} M, \end{aligned}$$

where

$$M = \left(\frac{\beta}{\theta} \frac{\lambda}{\lambda + \alpha} + \frac{\lambda + \beta}{\theta} \frac{\lambda}{\mu} + \frac{\lambda}{\mu} \frac{\lambda}{\lambda + \mu} \right) \prod_{i=1}^N \left(\delta_1^{(i+1)} + \frac{\varphi_1^{(i)}}{i} \right) P_{0,0}.$$

Thus,

$$\begin{aligned} \sum_{n=0}^{\infty} P_{0,n} &= \sum_{n=0}^N P_{0,n} + \sum_{n=N+1}^{\infty} P_{0,n} < \sum_{n=0}^N P_{0,n} + M \sum_{n=N+1}^{\infty} q^{n-N} \\ &= \sum_{n=0}^N P_{0,n} + M \lim_{n \rightarrow \infty} \frac{q}{1-q} (1 - q^n) < \infty. \end{aligned}$$

According to the expressions of $P_{1,n}$, $P_{2,n}$ and $P_{3,n}(x)$, in (A.5), (A.6) and (A.7), we can derive

$$\|P\|_X = \sum_{n=0}^{\infty} (|P_{0,n}| + |P_{1,n}| + |P_{2,n}| + \|P_{3,n}(x)\|_{L^1}) < \infty.$$

This implies that equation $(A + B + C)\vec{P} = 0$ has a unique nonzero solution and 0 is an eigenvalue of the system operator $A + B + C$, and the eigenvector of 0 is positive if $\lambda \int_0^{\infty} e^{-\int_0^x \mu_2(\xi) d\xi} dx < 1$ and $\lambda \leq \alpha$.

Appendix B. Proof of Step 3 in Theorem 3.2

First, we recall the definition of algebraic multiplicity of an eigenvalue from the work of Gupur [15, page 5].

DEFINITION B.1. Let T be a linear operator whose domain $D(T)$ and range $R(T)$ both lie in the same complex normed linear space X . Assume $\gamma \in \sigma_p(T)$, then the dimension of

$$\{x \in D(T) \mid Tx = \gamma x\}$$

is called the geometric multiplicity of γ . We call the smallest integer k such that

$$\{f \in D(T^k) \mid (\gamma I - T)^k f = 0\} = \{f \in D(T^{k+1}) \mid (\gamma I - T)^{k+1} f = 0\}$$

is the algebraic index of γ . The dimension of the spectral subspace

$$\text{Rang}^{\mathbb{P}} = \text{Rang} \left\{ \frac{1}{2\pi i} \int_{\Gamma} (\gamma I - T)^{-1} d\gamma \right\}$$

is called an algebraic multiplicity of γ , where Γ is the circle with centre γ and sufficiently small radius such that there are no other eigenvalues of T except for γ .

DEFINITION B.2. Let T be a linear operator whose domain $D(T)$ and range $R(T)$ both lie in the same complex normed linear space X . Assume $\gamma \in \sigma_p(T)$. If the geometric multiplicity of γ and the algebraic index of γ are equal to 1 simultaneously, then the algebraic multiplicity of γ is also equal to 1 and called a simple eigenvalue of T .

According to Definitions B.1 and B.2, to prove that 0 is a simple eigenvalue of operator $(A + B + C)^*$ with algebraic multiplicity one is equivalent to verify that the

geometric multiplicity of 0 and the algebraic index of 0 are equal to 1 simultaneously. Therefore, we show that as follows. Let

$$X^* = \{\vec{Q} = (Q_0, Q_1, Q_2, \dots)^T \mid \|\vec{Q}\| = \sup\{\|Q_n\|, n = 0, 1, 2, \dots\} < \infty\}$$

be the adjoint space of X , where $Q_n = (Q_{0,n}, Q_{1,n}, Q_{2,n}, Q_{3,n}(x))^T \in R^3 \times L^\infty(R^+)$ and $\|Q_n\| = \sup\{|Q_{0,n}|, |Q_{1,n}|, |Q_{2,n}|, \|Q_{3,n}(x)\|_{L^\infty}\}, n = 0, 1, 2, \dots$. Let $(A + B + C)^*$ be the adjoint operator of $(A + B + C)$, then for all $\vec{Q} \in X^*$, we can verify that

$$(A + B + C)^* \vec{Q} = \begin{pmatrix} \begin{pmatrix} E_0 Q_{0,0} + \beta Q_{2,0} + \lambda Q_{3,0}(0) \\ HQ_{1,0} + \lambda Q_{1,1} + \alpha Q_{3,0}(0) + \mu_1 Q_{2,0} \\ -\lambda Q_{2,0} + \lambda Q_{1,0} \\ \mu(x) Q_{0,0} + FQ_{3,0}(x) + \lambda Q_{3,1}(x) \end{pmatrix} \\ \vdots \\ \begin{pmatrix} E_n Q_{0,n} + \beta Q_{2,n} + n\theta Q_{3,n-1}(0) + \lambda Q_{3,n}(0) \\ HQ_{1,n} + \lambda Q_{1,n+1} + \alpha Q_{3,n}(0) + \mu_1 Q_{2,n} \\ -\lambda Q_{2,n} + \lambda Q_{1,n} \\ \mu(x) Q_{0,n} + FQ_{3,n}(x) + \lambda Q_{3,n+1}(x) \end{pmatrix} \\ \vdots \end{pmatrix}$$

with the domain

$$D((A + B + C)^*) = \{\vec{Q} = (Q_0, Q_1, Q_2, \dots)^T \in X^* \mid Q_n \in R^3 \times L^\infty(R^+), Q_{3,n}(x) \text{ are continuous functions}\},$$

where

$$E_n = -\lambda - \beta - n\theta, \quad F\phi(x) = \left(-\frac{d}{dx} - \lambda - \mu(x)\right)\phi(x), \quad H = -\lambda - \alpha - \mu_1.$$

Obviously, operator equation $(A + B + C)^* \vec{Q} = 0$ has the nonzero solution

$$\vec{Q}^{(k)} = k((1, 1, 1, 1(x))^T, (1, 1, 1, 1(x))^T, \dots)^T, \quad k \neq 0.$$

Thus, 0 is the eigenvalue of $(A + B + C)^*$ and the geometric multiplicity of 0 is 1.

Now, we show that the algebraic index of 0 is also equal to 1. Let $\vec{Q}^{(k')}$ be the eigenvector of 0 and let $\vec{P}' \in \mathcal{N}(A + B + C)$ which satisfies $\langle \vec{P}', \vec{Q}^{(k')} \rangle \neq 0$. We assume that there exists $\vec{Q}^* \in X^*$ such that $(A + B + C)^* \vec{Q}^* = \vec{Q}^{(k')}$, then we have

$$0 = \langle (A + B + C)\vec{P}', \vec{Q}^* \rangle = \langle \vec{P}', (A + B + C)^* \vec{Q}^* \rangle = \langle \vec{P}', \vec{Q}^{(k')} \rangle \neq 0,$$

which indicates a contradiction. Therefore, there is no $\vec{Q}^* \in X^*$ such that $(A + B + C)^* \vec{Q}^* = \vec{Q}^{(k')}$. According to Definition B.1, the algebraic index of 0 is equal to 1, and thus 0 is a simple eigenvalue of the operator $(A + B + C)^*$ with algebraic multiplicity one.

Acknowledgements

The authors are thankful to the referees for their constructive suggestions which helped to improve the presentation of this paper.

References

- [1] S. Aguir, F. Karaesmen, O. Z. Akşın and F. Chauvet, “The impact of retrials on call center performance”, *OR Spectrum* **26** (2004) 353–376; doi:[10.1007/S00291-004-0165-7](https://doi.org/10.1007/S00291-004-0165-7).
- [2] J. R. Artalejo, “A queueing system with returning customers and waiting line”, *Oper. Res. Lett.* **17** (1995) 191–199; doi:[10.1016/0167-6377\(95\)00017-E](https://doi.org/10.1016/0167-6377(95)00017-E).
- [3] J. R. Artalejo and A. Gómez-Corral, “Steady state solution of a single-server queue with linear repeated requests”, *J. Appl. Probab.* **34** (1997) 223–233; doi:[10.1017/S0021900200100841](https://doi.org/10.1017/S0021900200100841).
- [4] J. R. Artalejo and A. Gómez-Corral, “Analysis of a stochastic clearing system with repeated attempts”, *Comm. Statist. Stocha. Model.* **14** (1998) 623–645; doi:[10.1080/15326349808807492](https://doi.org/10.1080/15326349808807492).
- [5] P. J. Burke, “The output of a queueing system”, *Oper. Res.* **6** (1956) 699–704; doi:[10.1287/opre.4.6.699](https://doi.org/10.1287/opre.4.6.699).
- [6] A. Burnetas and A. Economou, “Equilibrium customer strategies in a single server Markovian queue with setup times”, *Queueing Syst.* **56** (2007) 213–228; doi:[10.1007/s11134-007-9036-7](https://doi.org/10.1007/s11134-007-9036-7).
- [7] T. V. Do, “M/M/1 retrial queue with working vacations”, *Acta Inform.* **47** (2010) 67–75; doi:[10.1007/s00236-009-0110-y](https://doi.org/10.1007/s00236-009-0110-y).
- [8] H. Dong and G. Xu, “The well-posedness of a retrial queue system with Bernoulli feedback and starting failure”, *Evol. Inst. Econ. Rev.* **12** (2015) 115–117; doi:[10.1007/10.1007/s40844-015-0009-y](https://doi.org/10.1007/10.1007/s40844-015-0009-y).
- [9] B. T. Doshi, “Queueing systems with vacations—A survey”, *Queueing Syst.* **1** (1986) 29–66; doi:[10.1007/BF01149327](https://doi.org/10.1007/BF01149327).
- [10] K. J. Engel, R. Nagel and S. Brendle, *One-parameter semigroups for linear evolution equations* (Springer, New York, 2012).
- [11] M. Escobar, A. R. Odoni and E. Roth, “Approximate solution for multi-server queueing systems with Erlangian service times,” *Comput. Oper. Res.* **29** (2002) 1353–1374; doi:[10.1016/S0305-0548\(01\)00036-3](https://doi.org/10.1016/S0305-0548(01)00036-3).
- [12] G. Falin and J. Templeton, *Retrial queues* (Chapman & Hall, London, 1997).
- [13] S. Gao and Z. Liu, “An M/G/1 queue with single working vacation and vacation interruption under Bernoulli schedule”, *Appl. Math. Model.* **37** (2013) 1564–1579; doi:[10.1016/j.apm.2012.04.045](https://doi.org/10.1016/j.apm.2012.04.045).
- [14] S. Gao, J. Wang and W. W. Li, “An M/G/1 retrial queue with general retrial times, working vacations and vacation interruption”, *Asia-Pac. J. Oper. Res.* **31** (2014) 1440006; doi:[10.1142/S0217595914400065](https://doi.org/10.1142/S0217595914400065).
- [15] G. Gupur, *Functional analysis methods for reliability models* (Springer, Germany, 2011).
- [16] J. D. Kim, D. W. Choi and K. C. Chae, “Analysis of queue-length distribution of the M/G/1 queue with working vacations (M/G/1/WV)”, *Proc. Hawaii Inter. Conf. Stat. Relat. Field.* 2003, Honolulu, HI, USA, June 5–8, 2003, 1191–1200; <http://hdl.handle.net/10203/146317>.
- [17] V. Kovtun, I. Izonin and M. Gregus, “Mathematical models of the information interaction process in 5G-IoT ecosystem: different functional scenarios”, *ICT Express* **9** (2023) 264–269; doi:[10.1016/j.ict.2021.11.008](https://doi.org/10.1016/j.ict.2021.11.008).
- [18] D. Y. Kuznetsov and A. A. Nazarov, “Analysis of non-Markovian models of communication networks with adaptive protocols of multiple random access”, *Autom. Remote Control* **62** (2001) 124–146; doi:[10.1023/A:1010231008994](https://doi.org/10.1023/A:1010231008994).
- [19] Z. Liu and Y. Song, “Geo/Geo/1 retrial queue with non-persistent customers and working vacations”, *J. Appl. Math. Comput.* **42** (2013) 103–115; doi:[10.1007/s12190-012-0623-3](https://doi.org/10.1007/s12190-012-0623-3).
- [20] A. A. Nazarov and S. A. Tsoj, “A common approach to studies of Markov models for data transmission networks controlled by the static random multiple access protocols”, *Autom. Contr. Comput. Sci.* **4** (2004) 73–85.

- [21] A. Pazy, *Semigroups of linear operators and applications to partial differential equations* (Springer, New York, 2012).
- [22] L. D. Servi and S. G. Finn, "M/M/1 queues with working vacations (M/M/1/WV)", *Perform. Eval.* **50** (2002) 41–52; doi:[10.1016/S0166-5316\(02\)00057-3](https://doi.org/10.1016/S0166-5316(02)00057-3).
- [23] K. Son, H. Kim, Y. Yi and B. Krishnamachari, "Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks", *IEEE J. Sel. Area. Comm.* **29** (2011) 1525–1536; doi:[10.1109/JSAC.2011.110903](https://doi.org/10.1109/JSAC.2011.110903).
- [24] H. Takagi, *Queueing analysis: a foundation of performance evaluation, Volume 1: vacation and priority systems, Part 1* (Elsevier Science, Amsterdam, The Netherlands, 1991) 1525–1536.
- [25] W. Wang, "The well-posedness and regularity of an Mⁿ/G/1 queue with feedback and optional server vacations based on a single vacation policy", *Int. J. Inform. Manage. Sci.* **20** (2009) 205–216; <https://api.semanticscholar.org/CorpusID:125378017>.
- [26] W. Wang and G. Q. Xu, "The well-posedness of an M/G/1 queue with second optional service and server breakdown", *Comput. Math. Appl.* **57** (2009) 729–739; doi:[10.1016/j.camwa.2008.09.038](https://doi.org/10.1016/j.camwa.2008.09.038).
- [27] D. A. Wu and H. Takagi, "M/G/1 queue with multiple working vacations", *Performance Evaluat.* **63** (2006) 654–681; doi:[10.1016/J.PEVA.2005.05.005](https://doi.org/10.1016/J.PEVA.2005.05.005).
- [28] J. Wu, Y. Wu, S. Zhou and Z. Niu, "Traffic-aware power adaptation and base station sleep control for energy-delay tradeoffs in green cellular networks", *2012 IEEE Global Communications Conf. (GLOBECOM)* (Anaheim, CA, USA, 2012), 3171–3176; doi: [10.1109/GLOCOM.2012.6503602](https://doi.org/10.1109/GLOCOM.2012.6503602).
- [29] M. Zhang and Q. Liu, "An M/G/1 G-queue with server breakdown, working vacations and vacation interruption", *Opsearch* **52** (2015) 256–270; doi:[10.1007/s12597-014-0183-4](https://doi.org/10.1007/s12597-014-0183-4).
- [30] S. Zhang, Y. Wu, S. Zhou and Z. Niu, "Traffic-aware network planning and green operation with BS sleeping and cell zooming", *IEICE Trans. Commun.* **97** (2014) 2337–2346; doi:[10.1587/transcom.E97.B.2337](https://doi.org/10.1587/transcom.E97.B.2337).
- [31] Y. Zhang and J. T. Wang, "Equilibrium pricing in an M/G/1 retrial queue with reserved idle time and setup time", *Appl. Math. Model.* **49** (2017) 514–530; doi:[10.1016/j.apm.2017.05.017](https://doi.org/10.1016/j.apm.2017.05.017).
- [32] F. Zheng, C. Gao and G. T. Zhu, "Asymptotic stability of the M/G/1 queueing system", *Acta Anal. Func. Appl.* **13** (2011) 218–224; doi:[10.3724/SP.J.1160.2011.00218](https://doi.org/10.3724/SP.J.1160.2011.00218).