Iranian Journal of Numerical Analysis and Optimization
Vol. 16, No. 4, 2026, pp 1591–1621
https://doi.org/10.22067/ijnao.2026.96253.1768
https://ijnao.um.ac.ir/

Research Article

Queuing analysis of a VoIP call center under retrial policy with re-service, server maintenance and service disruptions

Preena R. and Sumitha D.

Abstract

This paper develops a retrial queuing framework for a voice over internet protocol call center server with two obligatory service phases. Incoming voice packets arrive according to a Poisson process. The model captures caller behavior involving potential re-service requests based on satisfaction after the first phase, introducing dynamic retrial effects. It integrates operational uncertainties such as service disruptions due to malware attacks and immediate repair processes. Additionally, server maintenance activities and post-maintenance buffer searches are integrated to reflect realistic system operations. The supplementary variable technique is employed to derive steady-state solutions for performance evaluation. Numerical analysis highlights the influence of key parameters on system performance. To enhance the study, cost optimization is performed using the particle swarm optimization technique, offering effective strategies to balance service quality and operational cost.

AMS subject classifications (2020): 60K25; 60K30; 68M20; 90B22

Received 30 October 2025; revised 18 February 2026; accepted 2 April 2026

Preena R.
Department of Mathematics, Research Scholar, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore,Tamil Nadu, India. e-mail: preenaravikumar24@gmail.com

Sumitha D.
Department of Science and Humanities, Associate Professor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore,Tamil Nadu, India. e-mail: drsumithad@gmail.com

How to cite this article
Preena R. and Sumitha D., Queuing analysis of a VoIP call center under retrial policy with re-service, server maintenance and service disruptions. *Iran. J. Numer. Anal. Optim.*, 2026; 16(4): 1591-1621. https://doi.org/10.22067/ijnao.2026.96253.1768

Keywords: Retrial; Queue; Repair; Server; Re-service

1 Introduction

The retrial queue is defined as follows: If the arriving customers find the server busy, then they join the retrial group of customers in the orbit to repeat the attempt for their service after a random time. If arriving customers find the server free, then one of the customers joins the service area and the others leave for the orbit. Retrial queuing systems have been extensively applied in telecommunications. Works on general retrial queuing models are found in [12, 3, 10].

In the last two decades, researchers have worked on feedback (re-service) retrial queues to enhance customer satisfaction and the quality of service, where the customers can opt for re-service. Various performance measures of the M/G/1 retrial queue with Bernoulli feedback and single vacation were studied in [21] and the system size distribution at random points was derived using probability generating functions. The work quantifies the impact of feedback on the expected orbit length. Moreover, the inclusion of feedback, starting failures and vacation mechanisms provides a comprehensive framework for analyzing the behavior of retrial customers. Studies in [23, 7, 26] have examined retrial queuing systems with customer feedback and its impact on the performance of the system. Recently, feedback retrial queues in telecommunication systems were discussed [1], along with a sensitivity analysis of performance measures in a telecommunication system. The primary focus of the work is on modeling and analyzing a complex queuing system that reflects real-world scenarios where a batch of customers requiring different types of service rejoin the orbit if unsatisfied with the initial service. The feedback mechanism is a critical component that is embedded to understand the impact of feedback rates on the overall system's efficiency. A retrial queuing system was modeled in [25], analyzing how feedback probability and other system parameters influence server state probabilities and customer satisfaction. The analysis provides insights into how an increase in the feedback probability impacts the system and emphasizes the need to tune system parameters to keep congestion manageable.

In a retrial queuing system, a Bernoulli vacation (server maintenance) is a single vacation where the server is under maintenance for a certain period of time after service completion. During the vacation period, the server may also perform any alternative task. An extensive analysis of a two-phase batch arrival retrial queuing system with Bernoulli vacation schedule was carried out in [8] and the steady-state distribution of the server state and number of customers in the retrial group was obtained. Some important performance measures of the system were derived in [16] by introducing supplementary variables and employing the embedded Markov chain technique in a batch arrival retrial queuing system for essential and optional services with server breakdown and Bernoulli vacation. A mathematical model of a retrial queuing system featuring two types of customer services and a Bernoulli vacation policy was presented in [2]. The inclusion of the vacation mechanism aims to optimize system performance. A comprehensive literature review of retrial queuing systems dealing with Bernoulli vacation policies was provided in [20], offering a detailed summary of key aspects of the server state probabilities, analytical techniques and practical applications.

The phenomenon of orbital search (buffer search) in retrial queues is incorporated to effectively handle the customers waiting for service in the orbit or buffer. The joint distribution of the server state and orbit length in the steady-state of an M/G/1 retrial queue with non-persistent customers and orbital search is analyzed in [18]. The model is studied with a distinguishing and proactive feature, which is the orbital search mechanism, where the server actively seeks out customers in the orbit to provide service. The work is mathematically modeled to show the effects of orbital search on the steady-state behavior of the queue and the distribution of customers in the orbit. A two-phase retrial queue incorporating Bernoulli feedback and orbital search was investigated in [6], along with analytical derivations for key performance measures such as the expected number of customers in the orbit and the server state probabilities. An M/G/1 feedback retrial queuing system incorporating orbital search was examined in [5]. The authors derived explicit steady-state solutions for the number of customers in the system and in the orbit using probability generating functions and validated the analytical results with numerical examples.

Recently, retrial queues with negative customers (malware attack) such as viruses, have attracted growing interest from many researchers. The arrival of the negative customers is termed as G-queue and it was first introduced in [13]. A survey of negative customer arrivals in queuing systems and their applications were discussed in [27]. The steady-state analysis of an unreliable batch arrival feedback retrial queue with k-stages of service and Bernoulli vacation policy, where the server is subjected to breakdown due to the arrival of negative customers, is analyzed in [22]. A single-server queue with negative arrivals under a retrial policy and a Bernoulli vacation schedule was studied in [19]. The effect of vacation on overall system efficiency is analyzed and important performance indicators such as mean queue length and mean waiting time are calculated. Sensitivity analysis is carried out to illustrate the effects of model parameters on the performance of the system. A complex retrial queuing model integrating both the feedback process and the arrival of the negative customers was structured in [4]. The study provided a stochastic analysis along with explicit expressions for key performance measures of the system and demonstrated how negative arrivals influence the behavior of the system.

In the recent past, there has been a growing interest in the applications of retrial queues in call centers. A comprehensive empirical investigation into the factors driving customers to make repeated calls to call centers was conducted in [15], emphasizing the role of service preferences and perceived service quality. Their analytical framework helps to quantify these preferences and supports the development of more effective service designs. A queuing network model specifically designed for call centers managing customer retrials was developed and analyzed in [24]. A closed exponential queuing network was proposed as a stochastic framework for modeling call processing in modern call center environments.

Addressing a gap in the existing literature, this paper studies a retrial queuing model for a voice over internet protocol (VoIP) call center server operating under two obligatory service phases. The model incorporates customer-driven re-service requests contingent on satisfaction after the first phase and explicitly accounts for server disruptions caused by malware attacks, followed by immediate repairs. These features are analyzed together with server maintenance and buffer search mechanisms to examine their impact on system performance.

The structure of this paper is outlined in the following sections: Section 2 describes the system model in detail. Section 3 provides the stability condition for the model under study. Section 4 formulates the steady-state balance equations for the proposed model. Section 5 derives the steady-state results by solving these equations using the Supplementary Variable Technique. Section 6 presents the performance measures of the system. Section 7 focuses on the derivation of the mean buffer size and overall system size. Sections 8, 9, and 10 respectively provide the calculation of the reliability indices, mean busy period and mean busy cycle. Section 11 conducts numerical analysis to illustrate the system behavior under varying parameters. Section 12 offers a comprehensive cost analysis using the particle swarm optimization (PSO) technique. Finally, Section 13 concludes the paper with summery remarks and future research.

2 Model description

Multiple incoming voice packets arrive in accordance with a Poisson process with rate λ^t. Let the random variable Y be the packet size with $P\{Y = k\} = C_k$, $k = 1, 2, 3, \ldots, \infty$, and let $C(z)$ be the probability generating function with first two moments m_1 and m_2. The server renders service in two phases: Initial inquiry phase (IIP) and sales assistance phase (SAP).

If the incoming callers find the server busy or inaccessible, then the callers join the buffer to retry for service later with retrial rate η. The retrial time is generally distributed with distribution function $A(x)$, density function $a(x)$, Laplace–Stieltjes transform $A^*(x)$ and retrial completion rate $\eta(x)$.

If the callers find the server available, then one of the callers from the packet gets connected to the IIP and other callers join the buffer to retry for service after some random time. After completion of the IIP, depending upon the quality of the service, the caller either opts for re-service with probability δ or opts for SAP with the complementary probability $\bar{\delta} = (1-\delta)$. After completion of the SAP, the caller ends the call.

The service time during both phases is generally distributed with distribution function $B_i(x)$, density function $b_i(x)$, Laplace–Stieltjes transform $B_i^*(s)$, and service completion rate $\zeta_i(x)$ with moments ζ_{i1} and ζ_{i2}, where $i = 1, 2$.

After service completion, the server goes for server maintenance with probability ω or remains accessible with the complementary probability $\bar{\omega} = (1-\omega)$. The maintenance time of the server is generally distributed with distribution function $V(x)$, density function $v(x)$, Laplace–Stieltjes transform $V^*(s)$, and maintenance completion rate $g(x)$ with moments ϑ_1 and ϑ_2.

On returning from the maintenance activity, the server searches for the caller in the buffer with probability θ or remains accessible with the complementary probability $\bar{\theta} = (1-\theta)$. The malware attack interrupts the service with negative arrival rate λ^- during both phases of service which ultimately leads to call dropouts.

The repair for the failed server begins immediately. After the repair is completed, the server becomes accessible. The repair time during both phases is generally distributed with distribution function $R_i(x)$, density function $r_i(x)$, Laplace–Stieltjes transform $R_i^*(s)$ and repair completion rate $\beta_i(x)$ with moments β_{i1} and β_{i2}, where $i = 1, 2$. The structure of the described retrial queuing model is illustrated in Figure 1.

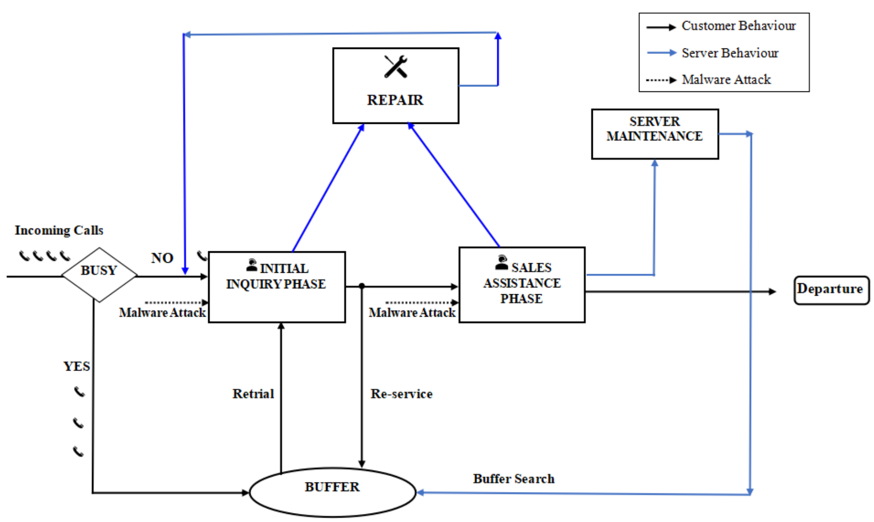


Figure 1: Structure of the described retrial queuing model in VoIP call center

The Markov process for the model can be defined as

$$\{M(t) : t \geq 0\} = \{X(t), C(t) : t \geq 0\},$$

where $X(t)$ is the number of callers in the buffer, $C(t)$ is the server states; that is, 0, 1, 2, 3, 4, 5 depending on whether the server is available, server is busy rendering IIP, server is busy rendering SAP, server is under repair during IIP, server is under repair during SAP, and the server is under maintenance, respectively.

2.1 Steady-state probabilities

- E_0 is defined as the server is in the available state with zero callers in the buffer.
- $E_n(x,t)dx,\ n \geq 1$, is defined as the server is in the available state with n callers in the buffer at time and $\chi_0(t)$ is the elapsed retrial time between x and $x+dx$.
- $W_n(x,t)dx,\ n \geq 0$, is defined as the server is busy serving in IIP with n callers in the buffer and $\chi_1(t)$ is the elapsed IIP service time between x and $x+dx$.
- $P_n(x,t)dx,\ n \geq 0$, is defined as the server is busy serving in SAP with n callers in the buffer and $\chi_2(t)$ is the elapsed SAP service time between x and $x+dx$.
- $R_{1,n}(x,t)dx,\ n \geq 0$, is defined as the server is under repair during IIP due to the negative arrival with n callers in the buffer and $\chi_3(t)$ is the elapsed repair time during IIP between x and $x+dx$.
- $R_{2,n}(x,t)dx,\ n \geq 0$, is defined as the server is under repair during SAP due to the negative arrival with n callers in the buffer and $\chi_4(t)$ is the elapsed repair time during SAP between x and $x+dx$.
- $V_n(x,t)dx,\ n \geq 0$, is defined as the server is on maintenance with n callers in the buffer and $\chi_5(t)$ is the elapsed maintenance time between x and $x+dx$.

3 Stability condition

Let N_n be the orbit length at the time of the nth customer departure, $n \geq 1$. Then $\{N_n, n \geq 1\}$ is ergodic if and only if

$$\begin{aligned}&\delta B_1^*(\lambda^-) + \lambda^+ m_1\big[1-\delta B_1^*(\lambda^-) - \bar{\delta} B_1^*(\lambda^-) B_2^*(\lambda^-)\big]/\lambda^- \\ &+ \lambda^+ m_1\big[\beta_{11}\big(1-B_1^*(\lambda^-)\big) + \bar{\delta}\beta_{21} B_1^*(\lambda^-)\big(1-B_2^*(\lambda^-)\big) + \omega\bar{\delta}\vartheta_1 B_1^*(\lambda^-)\end{aligned}$$

$$B_2^*(\lambda^-)\big] + m_1\big(1 - A^*(\lambda^+)\big)\big[1 - \omega\theta\bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)\big] < 1.$$

It can be proved along similar lines as in [14].

Since the arrival stream is a Poisson process, it can be shown from Burke's theorem [9] that the steady-state probabilities of
$\{X(t), C(t) : t \geq 0\}$ exist and are positive if and only if

$$\begin{aligned}
&\delta B_1^*(\lambda^-) + \lambda^+ m_1\big[1 - \delta B_1^*(\lambda^-) - \bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)\big]/\lambda^- \\
&+ \lambda^+ m_1\big[\beta_{11}\big(1 - B_1^*(\lambda^-)\big) + \bar{\delta}\beta_{21}B_1^*(\lambda^-)\big(1 - B_2^*(\lambda^-)\big) + \omega\bar{\delta}\vartheta_1 B_1^*(\lambda^-) \\
&B_2^*(\lambda^-)\big] + m_1\big(1 - A^*(\lambda^+)\big)\big[1 - \omega\theta\bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)\big] < 1.
\end{aligned}$$

From the mean drift

$$\begin{aligned}
\varrho_j =& \delta B_1^*(\lambda^-) + \lambda^+ m_1\big[1 - \delta B_1^*(\lambda^-) - \bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)\big]/\lambda^- \\
&+ \lambda^+ m_1\big[\beta_{11}\big(1 - B_1^*(\lambda^-)\big) + \bar{\delta}\beta_{21}B_1^*(\lambda^-)\big(1 - B_2^*(\lambda^-)\big) \\
&+ \omega\bar{\delta}\vartheta_1 B_1^*(\lambda^-)B_2^*(\lambda^-)\big] \\
&+ m_1\big(1 - A^*(\lambda^+)\big)\big[1 - \omega\theta\bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)\big] - 1, \qquad \text{for all } j \geq 1.
\end{aligned}$$

We have a reasonable conclusion that the term $\lambda^+ m_1\big[1 - \delta B_1^*(\lambda^-) - \bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)\big]/\lambda^-$ represents the number of arriving customers during the service time. The term $\lambda^+ m_1\big[\beta_{11}\big(1 - B_1^*(\lambda^-)\big) + \bar{\delta}\beta_{21}B_1^*(\lambda^-)\big(1 - B_2^*(\lambda^-)\big) + \omega\bar{\delta}\vartheta_1 B_1^*(\lambda^-)B_2^*(\lambda^-)\big]$ represents the number of arriving customers during the repair time and maintenance time of the server.

The term $\delta B_1^*(\lambda^-)$ corresponds to the number of customers opting for re-service. The term $m_1\big(1 - A^*(\lambda^+)\big)\big[1 - \omega\theta\bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)\big] - 1$ refers to the contribution to the buffer size due to a batch arrival during the retrial time excluding the arbitrary customer of the arriving batch whose service commences so that he no longer belongs to the buffer.

Similar interpretation can be provided for $j = 0$. The condition $j < 0$ ensures that the buffer size does not grow indefinitely in course of time.

Accordingly the stability condition is

$$\begin{aligned}
&\delta B_1^*(\lambda^-) + \lambda^+ m_1\big[1 - \delta B_1^*(\lambda^-) - \bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)\big]/\lambda^- \\
&+ \lambda^+ m_1\big[\beta_{11}\big(1 - B_1^*(\lambda^-)\big) + \bar{\delta}\beta_{21}B_1^*(\lambda^-)\big(1 - B_2^*(\lambda^-)\big) \\
&+ \omega\bar{\delta}\vartheta_1 B_1^*(\lambda^-)B_2^*(\lambda^-)\big] + m_1\big(1 - A^*(\lambda^+)\big)\big[1 - \omega\theta\bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)\big] < 1.
\end{aligned}$$

4 Steady-state equations

The steady-state equations are framed for

$$\lambda^+ E_0 = \bar{\omega}\int_0^{\infty} P_0(x)\zeta_2(x)dx + \int_0^{\infty} V_0(x)\vartheta(x)dx + \int_0^{\infty} R_1(x)\beta_1(x)dx + \int_0^{\infty} R_2(x)\beta_2(x)dx, \quad (1)$$

$$\frac{d}{dx}E_n(x) = -(\lambda^+ + \eta(x))E_n(x), \quad n \geq 1, \quad (2)$$

$$\frac{d}{dx}W_n(x) = -(\lambda^+ + \lambda^- + \zeta_1(x))W_n(x) + \lambda^+ \sum_{k=1}^{n} C_k W_{n-k}(x), \quad n \geq 0, \quad (3)$$

$$\frac{d}{dx}P_n(x) = -(\lambda^+ + \lambda^- + \zeta_2(x))P_n(x) + \lambda^+ \sum_{k=1}^{n} C_k P_{n-k}(x), \quad n \geq 0, \quad (4)$$

$$\frac{d}{dx}R_{1,n}(x) = -(\lambda^+ + \beta_1(x))R_{1,n}(x) + \lambda^+ \sum_{k=1}^{n} C_k R_{1,n-k}(x), \quad n \geq 0, \quad (5)$$

$$\frac{d}{dx}R_{2,n}(x) = -(\lambda^+ + \beta_2(x))R_{2,n}(x) + \lambda^+ \sum_{k=1}^{n} C_k R_{2,n-k}(x), \quad n \geq 0, \quad (6)$$

$$\frac{d}{dx}V_n(x) = -(\lambda^+ + \vartheta(x))V_n(x) + \lambda^+ \sum_{k=1}^{n} C_k V_{n-k}(x), \quad n \geq 0, \quad (7)$$

with boundary conditions

$$E_n(0) = \delta\int_0^{\infty} W_{n-1}(x)\zeta_1(x)\,dx + \bar{\omega}\int_0^{\infty} P_n(x)\zeta_2(x)\,dx + \bar{\theta}\int_0^{\infty} V_n(x)\vartheta(x)\,dx + \int_0^{\infty} R_{1,n}(x)\beta_1(x)\,dx + \int_0^{\infty} R_{2,n}(x)\beta_2(x)\,dx, \quad n \geq 1, \quad (8)$$

$$W_0(0) = \lambda^+ C_1 E_0 + \int_0^{\infty} E_1(x)\eta(x)\,dx + \theta\int_0^{\infty} V_1(x)\vartheta(x)\,dx, \quad (9)$$

$$
\begin{aligned}
W_n(0) = \lambda^+ C_{n+1} E_0 &+ \int_0^\infty E_{n+1}(x)\eta(x)\,dx \\
&+ \lambda^+ \sum_{k=1}^{n} C_k \int_0^\infty E_{n-k+1}(x)\,dx \\
&+ \theta \int_0^\infty V_{n+1}(x)\vartheta(x)\,dx, \quad n \geq 1,
\end{aligned} \tag{10}
$$

$$
P_n(0) = \bar{\delta} \int_0^\infty W_n(x)\zeta_1(x)\,dx, \quad n \geq 0, \tag{11}
$$

$$
R_{1,n}(0) = \lambda^- \int_0^\infty W_n(x)\,dx, \quad n \geq 0, \tag{12}
$$

$$
R_{2,n}(0) = \lambda^- \int_0^\infty P_n(x)\,dx, \quad n \geq 0, \tag{13}
$$

$$
V_n(0) = \omega \int_0^\infty P_n(x)\zeta_2(x)\,dx, \quad n \geq 0. \tag{14}
$$

The normalizing equation is

$$
\begin{aligned}
&E_0 + \sum_{n=1}^{\infty}\int_0^\infty E_n(x)\,dx + \sum_{n=0}^{\infty}\int_0^\infty W_n(x)\,dx + \sum_{n=0}^{\infty}\int_0^\infty P_n(x)\,dx \\
&+ \sum_{n=0}^{\infty}\int_0^\infty R_{1,n}(x)\,dx + \sum_{n=0}^{\infty}\int_0^\infty R_{2,n}(x)\,dx + \sum_{n=0}^{\infty}\int_0^\infty V_n(x)\,dx = 1.
\end{aligned} \tag{15}
$$

Define the probability generating functions

$$
\begin{aligned}
\mathrm{E}(x,z) &= \sum_{n=1}^{\infty} E_n(x)z^n, & \mathrm{R}_1(x,z) &= \sum_{n=0}^{\infty} R_{1,n}(x)z^n, \\
\mathrm{W}(x,z) &= \sum_{n=0}^{\infty} W_n(x)z^n, & \mathrm{R}_2(x,z) &= \sum_{n=0}^{\infty} R_{2,n}(x)z^n, \\
\mathrm{P}(x,z) &= \sum_{n=0}^{\infty} P_n(x)z^n, & \mathrm{V}(x,z) &= \sum_{n=0}^{\infty} V_n(x)z^n.
\end{aligned}
$$

By multiplying equation (2) by z^n and summing over n, we get

$$\left(\frac{d}{dx}+\lambda^{+}+\eta(x)\right)E(x,z)=0 \tag{16}$$

Solving the partial differential equation (16), we get

$$\begin{aligned}E(x,z)=&Ce^{-\int[\lambda^{+}+\eta(x)]\,dx}\\=&Ce^{-\lambda^{+}x}[1-A(x)].\end{aligned}$$

Eliminating C by taking $x=0$, we obtain

$$\therefore\ E(x,z)=E(0,z)e^{-\lambda^{+}x}[1-A(x)]. \tag{17}$$

Similarly, by multiplying (3)-(7) by z^n, summing over n and solving, we get

$$W(x,z)=W(0,z)e^{-\left[\lambda^{-}+\lambda^{+}-\lambda^{+}C(z)\right]x}\left[1-B_1(x)\right], \tag{18}$$

$$P(x,z)=P(0,z)e^{-\left[\lambda^{-}+\lambda^{+}-\lambda^{+}C(z)\right]x}\left[1-B_2(x)\right], \tag{19}$$

$$R_1(x,z)=R_1(0,z)e^{-\left[\lambda^{+}-\lambda^{+}C(z)\right]x}\left[1-R_1(x)\right], \tag{20}$$

$$R_2(x,z)=R_2(0,z)e^{-\left[\lambda^{+}-\lambda^{+}C(z)\right]x}\left[1-R_2(x)\right], \tag{21}$$

$$V(x,z)=V(0,z)e^{-\left[\lambda^{+}-\lambda^{+}C(z)\right]x}\left[1-V(x)\right]. \tag{22}$$

Multiplying (8) to (14) by z^n and summing over n, we get

$$\begin{aligned}E(0,z)=\delta z\int_0^\infty W(x,z)\zeta_1(x)\,dx+\bar{\omega}\int_0^\infty P(x,z)\zeta_2(x)\,dx\\+\bar{\theta}\int_0^\infty V(x,z)\vartheta_(x)\,dx+\int_0^\infty R_1(x,z)\beta_1(x)\,dx\\+\int_0^\infty R_2(x,z)\beta_2(x)\,dx-\lambda^{+}E_0,\end{aligned} \tag{23}$$

$$W(0,z)=\frac{\lambda^{+}C(z)}{z}E_0+\frac{1}{z}\int_0^\infty E(x,z)\eta(x)\,dx$$

$$+\frac{\lambda^{+}C(z)}{z}\int_{0}^{\infty}E(x,z)\,dx+\frac{\theta}{z}\int_{0}^{\infty}V(x,z)\vartheta(x)\,dx, \tag{24}$$

$$P(0,z)=\bar{\delta}\int_{0}^{\infty}W(x,z)\zeta_{1}(x)dx, \tag{25}$$

$$R_{1}(0,z)=\lambda^{-}\int_{0}^{\infty}W(x,z)dx, \tag{26}$$

$$R_{2}(0,z)=\lambda^{-}\int_{0}^{\infty}P(x,z)dx, \tag{27}$$

$$V(0,z)=\omega\int_{0}^{\infty}P(x,z)\zeta_{2}(x)dx. \tag{28}$$

Substituting (17)-(22) into (23)-(28), we get the following equations (29) to (34):

$$E(0,z)=\frac{\lambda^{+}E_{0}[[\theta\omega\bar{\delta}B_{1}^{*}(u_{1}(z))B_{2}^{*}(u_{1}(z))V^{*}(u_{2}(z))-z]u_{1}(z)+C(z)G_{2}(z)]}{K_{1}(z)}, \tag{29}$$

$$\mathrm{W}(0,z)=\frac{\lambda^{+}E_{0}A^{*}(\lambda^{+})(C(z)-1)}{K(z)}, \tag{30}$$

$$\mathrm{P}(0,z)=\frac{\lambda^{+}E_{0}A^{*}(\lambda^{+})\bar{\delta}(C(z)-1)B_{1}^{*}(u_{1}(z))}{K(z)}, \tag{31}$$

$$R_{1}(0,z)=\frac{\lambda^{+}\lambda^{-}E_{0}A^{*}(\lambda^{+})(C(z)-1)[1-B_{1}^{*}(u_{1}(z))]}{K_{1}(z)}, \tag{32}$$

$$R_{2}(0,z)=\frac{\lambda^{+}\lambda^{-}E_{0}A^{*}(\lambda^{+})\bar{\delta}(C(z)-1)B_{1}^{*}(u_{1}(z))[1-B_{2}^{*}(u_{1}(z))]}{K_{1}(z)}, \tag{33}$$

$$V(0,z)=\frac{\lambda^{+}E_{0}A^{*}(\lambda^{+})\omega\bar{\delta}(C(z)-1)B_{1}^{*}(u_{1}(z))B_{2}^{*}(u_{1}(z))}{K(z)}, \tag{34}$$

where

$$\begin{aligned}
u_1(z) &= \lambda^- + \lambda^+ - \lambda^+ C(z),\\
u_2(z) &= \lambda^+ - \lambda^+ C(z),\\
K(z) &= [[z - \theta\omega\bar{\delta}B_1^*(u_1(z))B_2^*(u_1(z))V^*(u_2(z))]u_1(z) - G_1(z)G_3(z)]/u_1(z),\\
K_1(z) &= [z - \theta\omega\bar{\delta}B_1^*(u_1(z))B_2^*(u_1(z))V^*(u_2(z))]u_1(z) - G_1(z)G_3(z),\\
G_1(z) &= A^*(\lambda^+) + C(z)(1 - A^*(\lambda^+)),\\
G_2(z) &= [\delta z B_1^*(u_1(z)) + \bar{\delta}B_1^*(u_1(z))B_2^*(u_1(z))(\bar{\omega} + \omega\bar{\theta}V^*(u_2(z))]u_1(z)\\
&\quad + \lambda^- R_1^*(u_2(z))[1 - B_1^*(u_1(z))]\\
&\quad + \lambda^-\bar{\delta}R_2^*(u_2(z))B_1^*(u_1(z))[1 - B_2^*(u_1(z))],\\
G_3(z) &= [\delta z B_1^*(u_1(z)) + \bar{\delta}B_1^*(u_1(z))\ B_2^*(u_1(z))(\bar{\omega} + \omega\bar{\theta})]u_1(z)\\
&\quad + \lambda^- R_1^*(u_2(z))[1 - B_1^*(u_1(z))]\\
&\quad + \lambda^-\bar{\delta}R_2^*(u_2(z))B_1^*(u_1(z))[1 - B_2^*(u_1(z))].
\end{aligned}$$

5 Steady-state results

i. The partial probability generating function of the buffer size when the server is available is

$$\begin{aligned}
E(z) &= \int_0^\infty E(x,z)\,dx\\
&= \int_0^\infty E(0,z)e^{-\lambda^+ x}[1 - A(x)]\,dx\\
&= \frac{E_0\,[1 - A^*(\lambda)]\Big\{C(z)G_z(z) - [z - \theta\omega\bar{\delta}B_1^*(u_1(z))B_2^*(u_1(z))V^*(u_2(z))]u_1(z)\Big\}}{K_1(z)}. \qquad (35)
\end{aligned}$$

ii. The partial probability generating function of the buffer size when the server is busy serving in IIP is

$$\begin{aligned}
W(z) &= \int_0^\infty W(x,z)\,dx\\
&= \int_0^\infty W(0,z)e^{-(u_1(z))x}\,[1 - B_1(x)]\ dx\\
&= \frac{\lambda^+ E_0 A^*(\lambda^+)\{[C(z) - 1][1 - B_1^*(u_1(z))]\}}{K_1(z)}. \qquad (36)
\end{aligned}$$

iii. The partial probability generating function of the buffer size when the server is busy serving in SAP is

$$\begin{aligned} P(z) &= \int_0^\infty P(x,z)\,dx \\ &= \int_0^\infty P(0,z)e^{-(u_1(z))x}\left[1-B_2(x)\right]\,dx \\ &= \frac{\lambda^+ E_0 A^*(\lambda^+)\{\bar{\delta}[C(z)-1]B_1^*(u_1(z))[1-B_2^*(u_1(z))]\}}{K_1(z)}. \end{aligned} \tag{37}$$

iv. The partial probability generating function of the buffer size when the server is under repair due to negative arrival during IIP is

$$\begin{aligned} R_1(z) &= \int_0^\infty R_1(x,z)\,dx \\ &= \int_0^\infty R_1(0,z)e^{-(u_2(z))x}\left[1-R_1(x)\right]\,dx \\ &= \frac{\lambda^- E_0 A^*(\lambda^+)\{[1-B_1^*(u_1(z))][R_1^*(u_2(z))-1]\}}{K_1(z)}. \end{aligned} \tag{38}$$

v. The partial probability generating function of the buffer size when the server is under repair due to negative arrival during SAP is

$$\begin{aligned} R_2(z) &= \int_0^\infty R_2(x,z)\,dx \\ &= \int_0^\infty R_2(0,z)e^{-(u_2(z))x}\left[1-R_2(x)\right]\,dx \\ &= \frac{\lambda^- E_0 A^*(\lambda^+)\{\bar{\delta}B_1^*(u_1(z))[1-B_2^*(u_1(z))][R_2^*(u_2(z))-1]\}}{K_1(z)}. \end{aligned} \tag{39}$$

vi. The partial probability generating function of the buffer size when the server is under maintenance is

$$\begin{aligned} V(z) &= \int_0^\infty V(x,z)\,dx \\ &= \int_0^\infty V(0,z)e^{-(u_2(z))x}\left[1-V(x)\right]\,dx \\ &= \frac{E_0 A^*(\lambda^+)\{\omega\bar{\delta}B_1^*(u_1(z))B_2^*(u_1(z))[V^*(u_2(z))-1]\}u_1(z)}{K_1(z)}. \end{aligned} \tag{40}$$

Using (35) to (40), we get the following result:

- The probability generating function of the number of callers in the buffer is

$$\begin{aligned}\phi_q(z) &= E_0 + E(z) + W(z) + P(z) + R_1(z) + R_2(z) + V(z)\\ &= \frac{\begin{array}{l}\{E_0 A^*(\lambda^+)\{[z - \delta z\, B_1^*(u_1(z)) - \bar{\delta}\, B_1^*(u_1(z))B_2^*(u_1(z))]\\ u_1(z) + \lambda^+(C(z)-1)[1 - B_1^*(u_1(z))[\delta + \bar{\delta}\, B_2^*(u_1(z))]]\\ -\lambda^-[1 - B_1^*(u_1(z))] - \lambda^- \bar{\delta}\, B_1^*(u_1(z))[1 - B_2^*(u_1(z))]\}\}\end{array}}{K_1(z)}.\end{aligned} \tag{41}$$

- The probability generating function of the number of callers in the system is

$$\begin{aligned}\phi_s(z) &= E_0 + E(z) + z(W(z) + P(z)) + R_1(z) + R_2(z) + V(z)\\ &= \frac{\begin{array}{l}\{E_0 A^*(\lambda^+)\{[z - \delta z\, B_1^*(u_1(z)) - \bar{\delta}\, B_1^*(u_1(z))B_2^*(u_1(z))]\\ u_1(z) + \lambda^+ z(C(z)-1)[1 - B_1^*(u_1(z))[\delta + \bar{\delta}\, B_2^*(u_1(z))]]\\ -\lambda^-[1 - B_1^*(u_1(z))] - \lambda^- \bar{\delta}\, B_1^*(u_1(z))[1 - B_2^*(u_1(z))]\}\}\end{array}}{K_1(z)},\end{aligned} \tag{42}$$

where

$$E_0 = \frac{K_2}{\lambda^- A^*(\lambda^+)[1 - \delta B_1^*(\lambda^-)]},$$

$$\begin{aligned}K_2 = \lambda^+\{&1 - \delta B_1^*(\lambda^-) + m_1(1 - A^*(\lambda^+))[\omega\theta\bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-) - 1]\\ &- S_5 - \lambda^+ m_1[1 - \delta B_1^*(\lambda^-) - \bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)]\}.\end{aligned}$$

6 Performance measures

Performance measures of the system are analyzed based on the following system characteristics:

6.1 Server state probabilities

i. The steady-state probability when the server is available

$$E=\lim_{z\to 1}E(z)=\frac{\begin{array}{c}\{E_0[1-A^*(\lambda^-)][\ \lambda^- m_1[1-\omega\theta\bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)]\\ -\lambda^-[1-\delta B_1^*(\lambda^-)]-S_5-\lambda^+ m_1[1-\delta B_1^*(\lambda^-)\\ -\bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)]]\}\end{array}}{K_2}. \tag{43}$$

ii. The steady-state probability when the server is busy in the IIP

$$W=\lim_{z\to 1}W(z)=\frac{\lambda^+ E_0A^*(\lambda^+)m_1[1-B_1^*(\lambda^-)]}{K_2}. \tag{44}$$

iii. The steady-state probability when the server is busy in the SAP

$$P=\lim_{z\to 1}P(z)=\frac{\lambda^+ E_0A^*(\lambda^+)\bar{\delta}m_1B_1^*(\lambda^-)[1-B_2^*(\lambda^-)]}{K_2}. \tag{45}$$

iv. The steady-state probability when the server is under repair during IIP

$$R_1=\lim_{z\to 1}R_1(z)=\frac{\lambda^-\lambda^+ E_0A^*(\lambda^+)m_1\beta_{11}[1-B_1^*(\lambda^-)]}{K_2}. \tag{46}$$

v. The steady-state probability when the server is under repair during SAP

$$R_2=\lim_{z\to 1}R_2(z)=\frac{\lambda^-\lambda^+ E_0A^*(\lambda^+)m_1\bar{\delta}\beta_{21}B_1^*(\lambda^-)[1-B_2^*(\lambda^-)]}{K_2}. \tag{47}$$

vi. The steady-state probability when the server is under maintenance

$$V=\lim_{z\to 1}V(z)=\frac{\lambda^-\lambda^+ E_0A^*(\lambda^+)\omega\bar{\delta}\vartheta_1 m_1B_1^*(\lambda^-)B_2^*(\lambda^-)}{K_2}. \tag{48}$$

6.2 Expected buffer size

The expected size of the buffer for the server states under steady-state conditions are calculated and given in Table 1.

In Table 1, we have

$$\begin{aligned}K_3=&-\lambda^+ m_2[1-\delta B_1^*(\lambda^-)-\bar{\delta}B_1^*(\lambda^-)(1-B_2^*(\lambda^-))]\\ &-\lambda^- m_2(1-A^*(\lambda^+))[1-\theta\omega\delta B_1^*(\lambda^-)B_2^*(\lambda^-)]\\ &-2\lambda^+ m_1[1-\delta B_1^*(\lambda^-)]+2(\lambda^+ m_1)^2S_1-2m_1(1-A^*(\lambda^+))\big\{\lambda^-\bar{\delta}B_1^*(\lambda^-)\end{aligned}$$

Table 1: Expected buffer size

S. No.	Server States	Expected Buffer Size
1.	Available	$M_E = \lim_{z\to 1} \frac{d}{dz}E(z) = \frac{H_2K_2 - H_1K_3}{2[K_2]^2}$
2.	Busy	$M_W = \lim_{z\to 1} \frac{d}{dz}W(z) = \frac{H_4K_2 - H_3K_3}{2[K_2]^2}$
		$M_P = \lim_{z\to 1} \frac{d}{dz}P(z) = \frac{H_6K_2 - H_5K_3}{2[K_2]^2}$
3.	Repair	$M_{R_1} = \lim_{z\to 1} \frac{d}{dz}R_1(z) = \frac{H_8K_2 - H_7K_3}{2[K_2]^2}$
		$M_{R_2} = \lim_{z\to 1} \frac{d}{dz}R_2(z) = \frac{H_{10}K_2 - H_9K_3}{2[K_2]^2}$
4.	Maintenance	$M_V = \lim_{z\to 1} \frac{d}{dz}V(z) = \frac{H_{12}K_2 - H_{11}K_3}{2[K_2]^2}$

$$
\begin{aligned}
&- \lambda^+ m_1[\delta B_1^*(\lambda^-) + \bar{\delta}(1-\theta\omega)B_1^*(\lambda^-)B_2^*(\lambda^-)] + S_2\big\}\\
&- \lambda^-[S_3 + 2\lambda^+ m_1\delta\zeta_{11} + 2(\lambda^+ m_1)^2 S_4],\\
H_1 &= E_0(1-A^*(\lambda^+))\{\lambda^- m_1[1-\omega\theta\bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)] - \lambda^+ m_1[1-\delta B_1^*(\lambda^-)\\
&- \bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)] - S_5 - \lambda^-(1-\delta B_1^*(\lambda^-))\},\\
H_2 &= E_0(1-A^*(\lambda^+))\big\{\lambda^+ m_2[1-\delta B_1^*(\lambda^-) - \bar{\delta}B_1^*(\lambda^-)B_2^*(\lambda^-)]\\
&+ \lambda^- m_2[1-\theta\omega\delta B_1^*(\lambda^-)B_2^*(\lambda^-)] + 2\lambda^+ m_1[1-\delta B_1^*(\lambda^-)] - 2(\lambda^+ m_1)^2 S_1\\
&+ 2m_1\big\{\lambda^-\delta B_1^*(\lambda^-) - \lambda^+ m_1[\delta B_1^*(\lambda^-) + \bar{\delta}(1-\theta\omega)B_1^*(\lambda^-)B_2^*(\lambda^-)] + S_2\big\}\\
&+ \lambda^-[S_3 + 2\lambda^+ m_1\delta\zeta_{11} + 2(\lambda^+ m_1)^2 S_4]\big\},\\
H_3 &= \lambda^+ m_1 A^*(\lambda^+)E_0[1-B_1^*(\lambda^-)],\\
H_4 &= \lambda^+ A^*(\lambda^+)E_0\{[1-B_1^*(\lambda^-)] - 2\lambda^+ m_1^2\zeta_{11}\},\\
H_5 &= \lambda^+\bar{\delta}m_1 A^*(\lambda^+)E_0 B_1^*(\lambda^-)[1-B_2^*(\lambda^-)],\\
H_6 &= \lambda^+\bar{\delta}A^*(\lambda^+)E_0\big\{m_2 B_1^*(\lambda^-)[1-B_2^*(\lambda^-)] + 2m_1[\lambda^+ m_1\zeta_{11}(1-B_2^*(\lambda^-))\\
&- \lambda^+ m_1\zeta_{21}B_1^*(\lambda^-)]\big\},\\
H_7 &= \lambda^-\lambda^+ m_1 A^*(\lambda^+)E_0\beta_{11}[1-B_1^*(\lambda^-)],\\
H_8 &= -\lambda^- A^*(\lambda^+)E_0\big\{2(\lambda^+ m_1)^2\zeta_{11}\beta_{11} - [1-B_1^*(\lambda^-)]h_3\big\},\\
H_9 &= \lambda^-\lambda^+\bar{\delta}m_1 A^*(\lambda^+)E_0\beta_{21}B_1^*(\lambda^-)[1-B_2^*(\lambda^-)],
\end{aligned}
$$

$$
\begin{aligned}
H_{10} &= \lambda^- \bar{\delta} A^*(\lambda^+) E_0 \big\{ 2(\lambda^+ m_1)^2 \beta_{21} [\lambda^+ m_1 \zeta_{11} (1 - B_2^*(\lambda^-)) \\
&\quad - \lambda^+ m_1 \zeta_{21} B_1^*(\lambda^-)] + B_1^*(\lambda^-)[1 - B_2^*(\lambda^-)] h_4 \big\}, \\
H_{11} &= \lambda^- \lambda^+ \omega \bar{\delta} m_1 A^*(\lambda^+) E_0 \vartheta_1 B_1^*(\lambda^-) B_2^*(\lambda^-), \\
H_{12} &= \omega \bar{\delta} A^*(\lambda^+) E_0 \big\{ 2(\lambda^+ m_1)^2 \vartheta_1 [\lambda^- (\zeta_{11} B_2^*(\lambda^-) - \zeta_{21} B_1^*(\lambda^-)) \\
&\quad - B_1^*(\lambda^-) B_2^*(\lambda^-)] + \lambda^- B_1^*(\lambda^-) B_2^*(\lambda^-) h_5 \big\}, \\
h_1 &= \lambda^+ m_2 \zeta_{11} + (\lambda^+ m_1)^2 \zeta_{12}, \\
h_2 &= \lambda^+ m_2 \zeta_{21} + (\lambda^+ m_1)^2 \zeta_{22}, \\
h_3 &= \lambda^+ m_2 \beta_{11} + (\lambda^+ m_1)^2 \beta_{12}, \\
h_4 &= \lambda^+ m_2 \beta_{21} + (\lambda^+ m_1)^2 \beta_{22}, \\
h_5 &= \lambda^+ m_2 \vartheta_1 + (\lambda^+ m_1)^2 \vartheta_2, \\
S_1 &= \delta \zeta_{11} + \bar{\delta} [\zeta_{11} B_2^*(\lambda^-) + \zeta_{21} B_1^*(\lambda^-) + \omega \vartheta_1 B_1^*(\lambda^-) B_2^*(\lambda^-)], \\
S_2 &= \lambda^- \lambda^+ m_1 \big\{ \beta_{11} (1 - B_1^*(\lambda^-)) + \bar{\delta} \beta_{21} B_1^*(\lambda^-)(1 - B_2^*(\lambda^-)) \\
&\quad + \omega \theta \bar{\delta} [\zeta_{11} B_2^*(\lambda^-) + \zeta_{21} B_1^*(\lambda^-)] + \omega \bar{\theta} \bar{\delta} \vartheta_1 B_1^*(\lambda^-) B_2^*(\lambda^-) \big\}, \\
S_3 &= h_3 (1 - B_1^*(\lambda^-)) + \bar{\delta} h_4 B_1^*(\lambda^-)(1 - B_2^*(\lambda^-)) + (1 - \theta\delta) \omega h_5 B_1^*(\lambda^-) B_2^*(\lambda^-), \\
S_4 &= (1 - \theta\delta) \omega \vartheta_1 [\zeta_{11} B_2^*(\lambda^-) + \zeta_{21} B_1^*(\lambda^-)] - \zeta_{11} \beta_{11} - \zeta_{21} \beta_{21} B_1^*(\lambda^-) \\
&\quad + \bar{\delta} \zeta_{11} \beta_{21} (1 - B_2^*(\lambda^-)), \\
S_5 &= \lambda^- \lambda^+ m_1 \big\{ \beta_{11} (1 - B_1^*(\lambda^-)) + \bar{\delta} \beta_{21} B_1^*(\lambda^-)(1 - B_2^*(\lambda^-)) \\
&\quad + \omega \bar{\delta} \vartheta_1 B_1^*(\lambda^-) B_2^*(\lambda^-) \big\}.
\end{aligned}
$$

7 Mean buffer size and system size

- The mean number of callers in the buffer is given by

$$L_q = \lim_{z \to 1} \frac{d}{dz} \phi_q(z).$$

Let $\mathrm{Nr}(\phi_q(z))$ and $\mathrm{Dr}(\phi_q(z))$ be the numerator and denominator of $\phi_q(z)$.

As $\mathrm{Nr}(\phi_q(1)) = \mathrm{Dr}(\phi_q(1)) = 0$, using L Hospital's rule, we have

$$L_q = \frac{\mathrm{Dr}'(\phi_q(1))\,\mathrm{Nr}''(\phi_q(1)) - \mathrm{Nr}'(\phi_q(1))\,\mathrm{Dr}''(\phi_q(1))}{2\left[\mathrm{Dr}'(\phi_q(1))\right]^2}, \tag{49}$$

where

$$\mathrm{Nr}'(\phi_q(1)) = \lambda^- E_0 A^*(\lambda^+)[1 - \delta B_1^*(\lambda^-)],$$

$$\mathrm{Nr}''(\phi_q(1)) = -2\lambda^+ m_1 E_0 A^*(\lambda^+)[1 - \delta B_1^*(\lambda^-) + \lambda^- \delta\mu_{11}],$$

$$\mathrm{Dr}'(\phi_q(1)) = K_2,$$

$$\mathrm{Dr}''(\phi_q(1)) = K_3.$$

- The mean number of callers in the system is given by

$$L_s = \lim_{z\to 1} \frac{d}{dz}\phi_s(z).$$

Let $\mathrm{Nr}(\phi_s(z))$ and $\mathrm{Dr}(\phi_s(z))$ be the numerator and denominator of $\phi_s(z)$.

As $\mathrm{Nr}(\phi_s(1)) = \mathrm{Dr}(\phi_s(1)) = 0$, using L Hospital's rule, we have

$$L_s = \frac{\mathrm{Dr}'(\phi_s(1))\,\mathrm{Nr}''(\phi_s(1)) - \mathrm{Nr}'(\phi_s(1))\,\mathrm{Dr}''(\phi_s(1))}{2\left[\mathrm{Dr}'(\phi_s(1))\right]^2}, \tag{50}$$

where

$$\begin{aligned}\mathrm{Nr}'(\phi_s(1)) = E_0 A^*(\lambda^+)\big\{&\lambda^-[1 - \delta B_1^*(\lambda)] \\ &- \lambda^+ m_1[1 - \delta B_1^*(\lambda^-) - \bar{\delta} B_1^*(\lambda^-) B_2^*(\lambda^-)]\big\},\end{aligned}$$

$$\begin{aligned}\mathrm{Nr}''(\phi_s(1)) =& E_0 A^*(\lambda^+)\big\{\lambda^-(1 - \delta h_1) \\ &+ - 2\lambda^+ m_1[\lambda^- \bar{\delta}\mu_{11}\bar{\delta} B_1^*(\lambda^-) B_2^*(\lambda^-)]\big\},\end{aligned}$$

$$\mathrm{Dr}'(\phi_s(1)) = K_2,$$

$$\mathrm{Dr}''(\phi_s(1)) = K_3.$$

8 Reliability indices

8.1 Availability of the server

The availability of the server under steady-state condition is

$$\begin{aligned} J(t) &= E_0 + \lim_{z\to 1} E(z) + \lim_{z\to 1} W(z) + \lim_{z\to 1} P(z) \\ &= \frac{\lambda^-[1-\delta B_1^*(\lambda^-)] - S_5}{\lambda^-[1-\delta B_1^*(\lambda^-)]}. \end{aligned}$$

8.2 Failure frequency

The failure frequency of the server under steady-state condition is

$$\begin{aligned} U(t) &= \lambda^- \big[\lim_{z\to 1} W(z) + \lim_{z\to 1} P(z)\big] \\ &= \frac{\lambda^+ m_1[1-\delta B_1^*(\lambda^-) - \bar{\delta} B_1^*(\lambda^-) B_2^*(\lambda^-)]}{1-\delta B_1^*(\lambda^-)}. \end{aligned}$$

9 Mean busy period

Mean busy period is the average time of the server when it is busy serving the callers from the arrival time of the first caller until the server is available for the next caller.

$$\begin{aligned} \mathrm{E}[S_B] &= \frac{1}{\lambda^+ m_1}\left(\frac{1}{E_0} - 1\right) \\ &= \frac{\begin{array}{c}\lambda^-[1-A^*(\lambda^+)][\delta B_1^*(\lambda^-) - 1 - m_1[\omega\theta\bar{\delta} B_1^*(\lambda^-) B_2^*(\lambda^-) - 1]] \\ + \lambda^+ m_1[1-\delta B_1^*(\lambda^-) - \bar{\delta} B_1^*(\lambda^-) B_2^*(\lambda^-)] + S_5\end{array}}{\lambda^+ m_1 K_2}. \end{aligned}$$

10 Mean busy cycle

Mean busy cycle is the average time of a complete cycle which includes both idle period and busy period of the server.

$$\mathrm{E}[S_C] = \frac{1}{\lambda^+ E_0} = \frac{\lambda^- A^*(\lambda^+)[1-\delta B_1^*(\lambda^-)]}{\lambda^+ K_2}.$$

11 Numerical analysis

Numerical analysis of the system is carried out using MATLAB by assuming that the retrial time, IIP service time, SAP service time, repair time during both phases and vacation time follow exponential distribution with respective rates $\eta, \zeta_1, \zeta_2, \beta_1, \beta_2$ and ϑ. The default values of the parameters are set as $\lambda^+ = 0.8$, $\lambda^- = 0.4$, $\delta = 0.8$, $\omega = 0.3$, $\theta = 0.7$, $\eta = 15$, $\zeta_1 = 10$, $\zeta_2 = 8$, $\beta_1 = 10$, $\beta_2 = 5$, $\vartheta = 25$, $C_1 = 0.5$ and $C_2 = 0.5$[17].

The effect of the positive arrival rate λ^+ on all the server state probabilities when the server is available in the empty state (E_0), when the server is available in the non-empty system (E), when the server is busy serving the caller in two phases of service (W & P), when the server is under repair during two phases of service (R_1 & R_2) and when the server is on maintenance activity (V) is analyzed and given in Table 2.

Table 2: Effect of λ^+ on the performance measures

λ^+	E_0	E	W	P	R_1	R_2	V	Z
0.1	0.8820	0.0355	0.0625	0.0149	0.0025	0.0012	0.0014	1.0000
0.2	0.7628	0.0722	0.1250	0.0298	0.0050	0.0024	0.0029	1.0000
0.3	0.6425	0.1100	0.1875	0.0446	0.0075	0.0036	0.0043	1.0000
0.4	0.5212	0.1488	0.2500	0.0595	0.0100	0.0048	0.0057	1.0000
0.5	0.3987	0.1887	0.3125	0.0744	0.0125	0.0060	0.0071	1.0000
0.6	0.2752	0.2298	0.3750	0.0893	0.0150	0.0071	0.0086	1.0000
0.7	0.1506	0.2719	0.4375	0.1042	0.0175	0.0083	0.0100	1.0000
0.8	0.0248	0.3152	0.5000	0.1190	0.0200	0.0095	0.0114	1.0000

The results indicate that as the arrival rate increases, the probability of the server being available in the empty state decreases, while the remaining server state probabilities increase. Several of these measures exhibit near-linear trends over the selected arrival-rate range, which is expected due to their analytical dependence on the arrival rate when other system parameters are held fixed and the system operates under moderate traffic conditions. Furthermore, the normalization condition given in (15) is satisfied, that is,

$$Z = E_0 + E + W + P + R_1 + R_2 + V = 1.$$

The effects of various parameters on the server state probability when the server is available in the empty state (E_0) and the mean number of callers in the system (L_S) are shown in the Figures 2 to 8.

It is evident from the analysis that

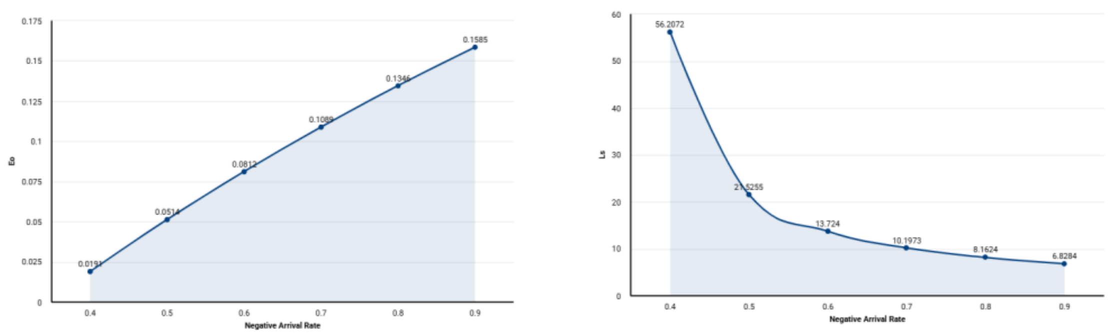
- Figure 2 shows that, with the increase of the negative arrival rate, E_0 rises and L_S reduces.
- In Figure 3, as the retrial rate rises, there is an escalation in E_0 and reduction in L_S.

- In Figure 4, as the re-service probability increases, E_0 reduces and L_S rises.

- In Figure 5, as the buffer search probability rises, there is an escalation in E_0 and a reduction in L_S.

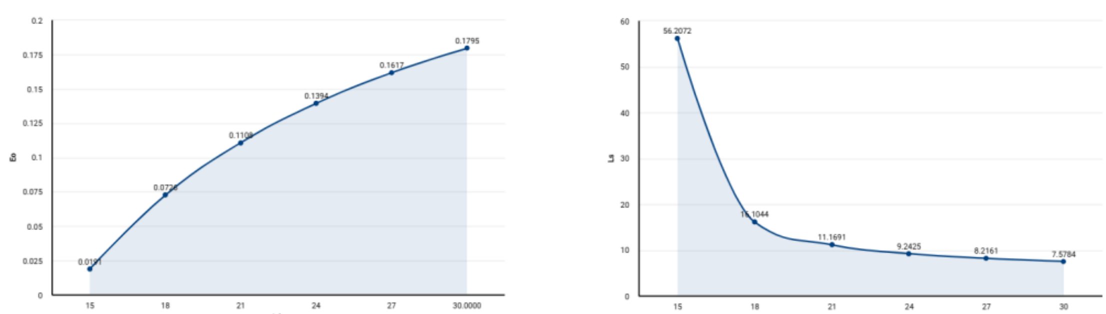
- Figures 6 and 7 show that, with increases of the both IIP and SAP service completion rate, there is an escalation in E_0 and reduction in L_S.

- In Figure 8, as the server maintenance completion rate increases, E_0 rises and L_S reduces.



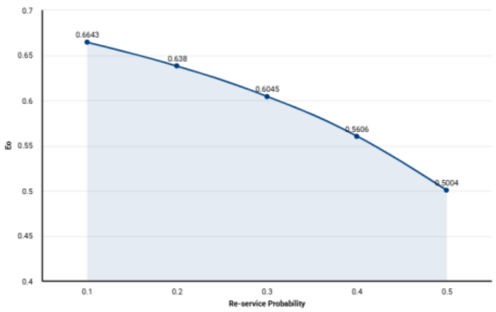
(a) Negative Arrival Rate vs E_0 (b) Negative Arrival Rate vs Ls

Figure 2: Effect of the negative arrival rate on the performance measures

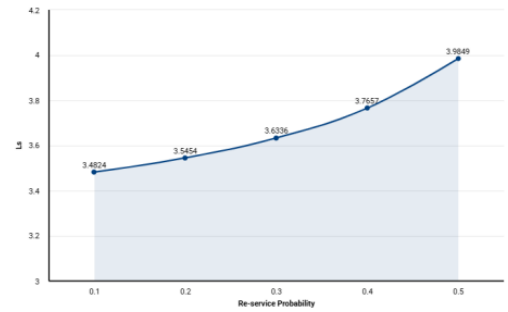


(c) Retrial Rate vs E_0 (d) Retrial Rate vs Ls

Figure 3: Effect of the retrial rate on the performance measures

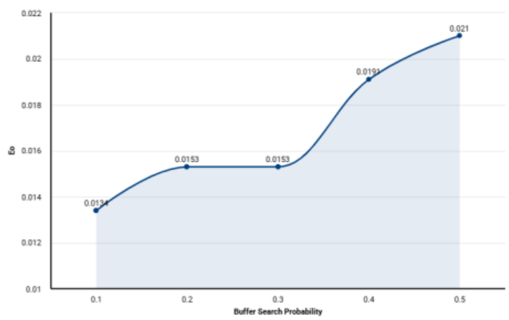


(e) Re-service Probability vs E_0

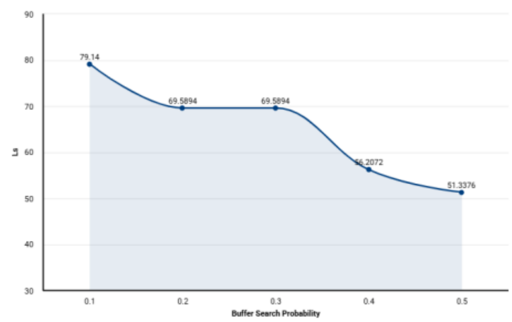


(f) Re-service Probability vs Ls

Figure 4: Effect of the re-service probability on the performance measures

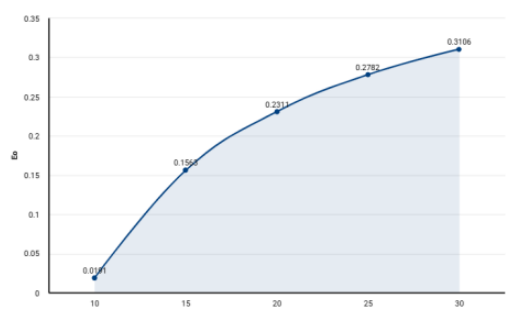


(g) Buffer Search Probability vs E_0

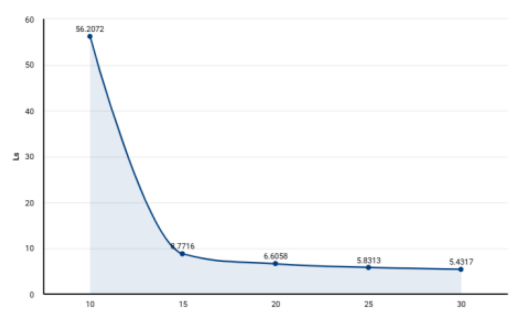


(h) Buffer Search Probability vs Ls

Figure 5: Effect of the buffer search probability on the performance measures

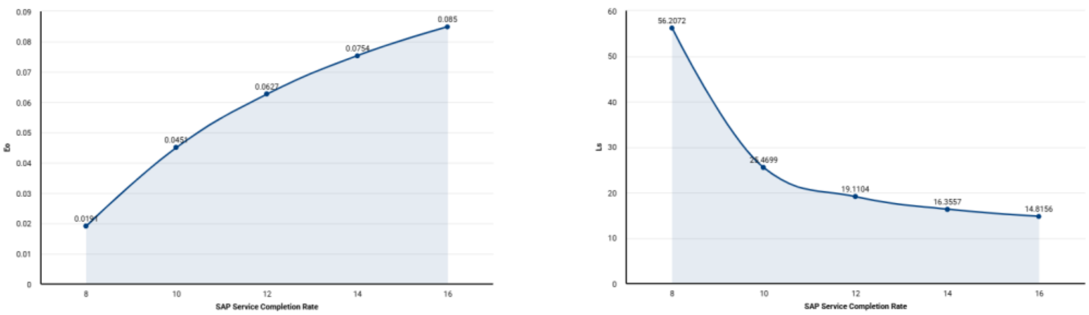


(i) IIP Service Completion Rate vs E_0



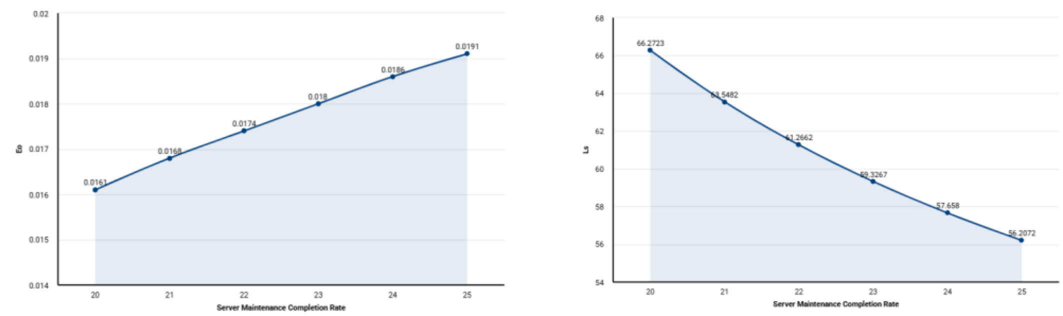
(j) IIP Service Completion Rate vs Ls

Figure 6: Effect of the IIP service completion rate on the performance measures



(k) SAP Service Completion Rate vs E_0 (l) SAP Service Completion Rate vs Ls

Figure 7: Effect of the SAP service completion rate on the performance measures



(m) Server Maintenance Completion Rate vs E_0 (n) Server Maintenance Completion Rate vs Ls

Figure 8: Effect of the server maintenance completion rate on the performance measures

12 Cost optimization

PSO is a computational technique that draws inspiration from the collective behavior of bird flocks [11]. PSO is an optimization algorithm that is population-based in nature, where a group of particles will iteratively try to find an optimal solution to a specific problem. PSO is applied for the optimization of retrial queuing systems by formulating a cost function to enhance the system performance, which is a critical problem in VoIP call centers where resource utilization and efficiency of cost are significant.

The expected total cost (ETC) for the system under study has been calculated by selecting parameter values based on standard practices reported in the literature [17] are set as $\lambda^+ = 0.8$, $\lambda^- = 0.4$, $\delta = 0.8$, $\omega = 0.3$, $\theta = 0.7$, $\eta = 15$, $\zeta_1 = 10$, $\zeta_2 = 8$, $\beta_1 = 10$, $\beta_2 = 5$, $\vartheta = 25$, $C_1 = 0.5$, $C_2 = 0.5$ and the cost values as

- Server Cost: Server cost is often calculated based on the resources used and the bandwidth consumption for each call, $Cs = \$0.20$.
- Holding Cost: Holding cost is associated with the time that a caller is on hold, $C_H = \$0.80$
- Service Cost: Service cost includes the cost for providing two phases of service, $C_E = \$0.20$.
- Repair Cost: Repair cost includes the cost for providing repair to the failed server per unit time, $C_R = \$0.20$.
- Maintenance Cost: Maintenance cost per unit time, $C_V = \$0.10$.

Using the above cost components and the steady-state performance measures derived from the retrial queuing model, the ETC function is formulated as

$$ETC = C_S + C_H L_S + C_E(W + P) + C_R(R_1 + R_2) + C_V V.$$

Objective: The objective of the optimization is to minimize the ETC by optimally selecting system parameters that significantly influence the performance of the retrial queuing system. These parameters include the incoming call rate (λ^+), re-service probability (δ), server maintenance probability (ω) and buffer search probability (θ) by varying the input learning parameters retrial rate η and IIP service completion rate $\zeta_1 = 10$.

To minimize the total cost, ETC function is formulated as

$$ETC\,(\eta^*, \zeta_1^*) = \min ETC\,(\eta, \zeta_1)$$

12.1 PSO algorithm

Step 1

Input parameters: The input learning parameters (η, ζ_1) are given.

Step 2

Initialization: Randomly initialize a swarm of particles with random values for the decision variables (η, ζ_1) within the bounds; lower bound = [10, 10] and upper bound = [40, 40]

- Number of particles = 25
- Maximum number of iterations = 50

- Inertia weight $= 0.9$

- Cognitive coefficient $= 1.5$ and Social coefficient $= 1.5$. The inertia weight close to unity promotes broader search capability in the initial stages, while equal cognitive and social coefficients ensure balanced influence between a particle's personal best position and the global best position of the swarm.

Step 3

Evaluation: Evaluate the cost function at each particle's current position.

Step 4

Iteration: Each particle iteratively updates its position based on its own best-known position (personal best) and the best-known position of the entire swarm (global best).

Step 5

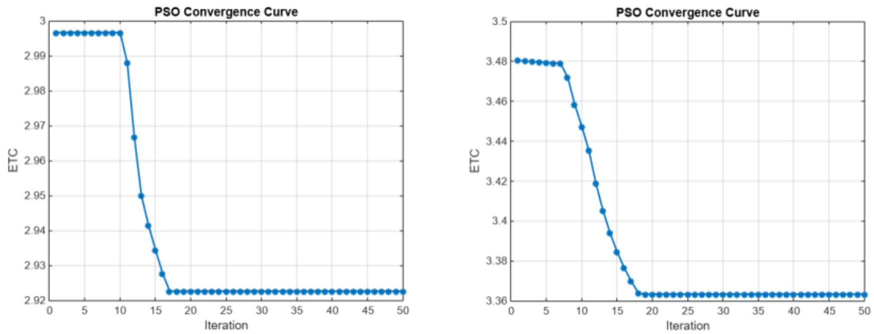
Position Update: The personal best and global best positions are updated based on the cost evaluations.

Step 6

Optimal Solution: The global best position at the end of the optimization process represents the optimal set of parameters of the retrial queuing system, minimizing the cost.

Output: The minimum cost $ETC\,(\eta^*, \zeta_1^*)$ is obtained.

Results and Discussion: The global best position obtained at the end of the PSO process represents the optimal set of system parameters that minimize the ETC of the retrial queuing system. The convergence curves presented in Figures 9–12 show smooth and stable behavior, with the optimal ETC values reached within a small number of iterations.
To examine the sensitivity of the optimization results, key system parameters were varied within reasonable operational ranges. It was observed that the optimal ETC values changed gradually and remained within a narrow range, indicating that the obtained solutions are not highly sensitive to moderate parameter variations. This behavior confirms the robustness and reliability of the PSO-based optimization for the proposed retrial queuing model.

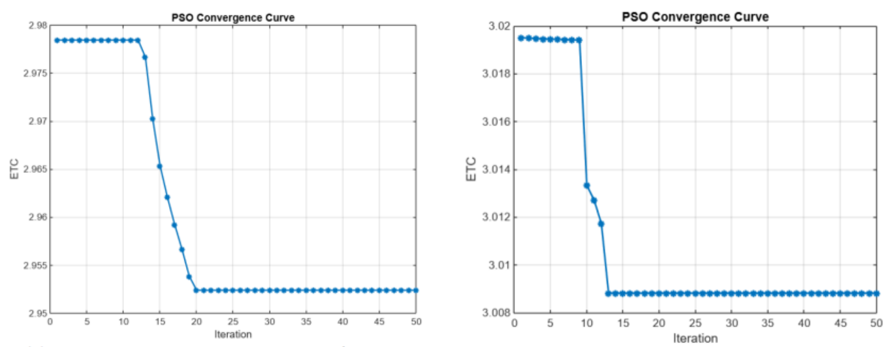


(a) PSO Convergence Curve vs Iteration on $\lambda^+ = 0.7$

(b) PSO Convergence Curve vs Iteration on $\lambda^+ = 0.8$

Figure 9: PSO convergence curve vs iteration on positive arrival rate

- Figure 9 illustrates that, at a lower positive arrival rate $\lambda^+ = 0.7$, the optimal value of the ETC is \$2.9226, achieved in the 17^{th} iteration. In contrast, at a higher arrival rate $\lambda^+ = 0.8$, the optimal ETC increases to \$3.3632, obtained in the 19^{th} iteration. This indicates that as the arrival rate increases, the optimal ETC also rises, reflecting that higher call volumes lead to increased operational costs.

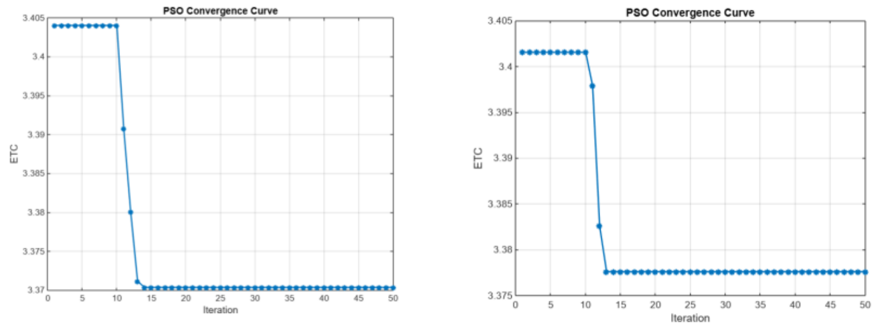


(c) PSO Convergence Curve vs Iteration on $\delta = 0.5$

(d) PSO Convergence Curve vs Iteration on $\delta = 0.6$

Figure 10: PSO convergence curve vs iteration on re-service probability

- Figure 10 shows that, at a re-service probability $\delta = 0.5$, the optimal ETC is \$2.9524, achieved in the 20^{th} iteration. As the re-service probability increases to $\delta = 0.6$, the optimal ETC also rises to \$3.0088, reached in the 13^{th} iteration. This suggests that when more customers require additional service after their IIP, the overall system cost increases. This highlight the impact of re-service on operational efficiency and cost management in call centers.

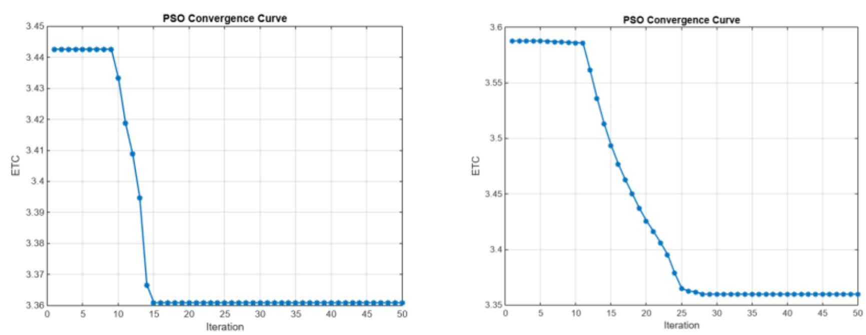


(e) PSO Convergence Curve vs Iteration on $\omega = 0.4$

(f) PSO Convergence Curve vs Iteration on $\omega = 0.5$

Figure 11: PSO convergence curve vs iteration on server maintenance probability

- Figure 11 shows that, at the server maintenance probability $\omega = 0.4$, the optimal value of the ETC, \$3.3704, is obtained in the 14^{th} iteration. When the server maintenance probability increases to $\omega = 0.5$, the optimal ETC slightly rises to \$3.3776, reached in the 13^{th} iteration. This marginal increase in cost shows that more frequent maintenance activities can lead to minor operational expenses in the call center.



(g) PSO Convergence Curve vs Iteration on $\theta = 0.6$

(h) PSO Convergence Curve vs Iteration on $\theta = 0.7$

Figure 12: PSO convergence curve vs iteration on the buffer search probability

- Figure 12 shows that, at the buffer search probability $\theta = 0.6$, the optimal value of the ETC, \$3.3608 is obtained in the 15^{th} iteration. As the buffer search probability increases to $\theta = 0.7$, the optimal ETC slightly decreases to \$3.3596, obtained in the 28^{th} iteration. This slight reduction in cost suggests that the buffer search mechanism reduces the system's overall cost.

13 Conclusion and future work

In this study, a comprehensive analytical model for a single-server VoIP call center with two obligatory service phases has been developed and analyzed. The model incorporates critical real-world factors such as re-service based on caller satisfaction, service interruptions caused by malware attacks, maintenance and buffer search activities. Key performance metrics, including server state probabilities, expected buffer size, mean number of callers in the system and buffer, reliability indices, mean busy period and mean busy cycle have been rigorously derived. The verification of the stochastic decomposition property validates the structural soundness of the model. Numerical analysis reveals the significant impact of system parameters on the performance measures. Furthermore, the PSO technique has been successfully employed to optimize the ETC of operating the proposed system. This research provides valuable retrial queuing insights in VoIP call center systems under practical constraints.

Future research could extend the model by incorporating a standby server operating at a reduced service rate during malware-induced interruptions, thereby enhancing service continuity. Exponential distributions are assumed for all time variables to preserve analytical solvability; extending to general distributions would increase model complexity and offer a direction for future research. Multi-server configurations can be investigated further, which would involve tracking multiple server states and require multi-dimensional Markov chains or matrix-analytic techniques to capture interactions among servers. Incorporating advanced customer behavior, such as impatience, alongside exploring alternative optimization techniques, would enrich the model's applicability.

Acknowledgements

The authors are grateful to their anonymous referees and editor for their constructive comments.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Abdollahi, S. and Salehi Rad, M.R. *Analysis of a batch arrival retrial queue with two-phase services, feedback and admission*, *Bull. Iran. Math. Soc.* (2022), 1–14.

[2] Arivudainambi, D. and Gowsalya, M. *A single server non-Markovian retrial queue with two types of service and Bernoulli vacation*, *Int. J. Oper. Res.* 33(1) (2018), 55–81.

[3] Artalejo, J.R. *Accessible bibliography on retrial queues: Progress in 2000–2009*, *Math. Comput. Model.* 51(9–10) (2010), 1071–1081.

[4] Boualem, M., Bouchentouf, A.A., Bareche, A. and Cherfaoui, M. *Stochastic interpretation for a single server retrial queue with Bernoulli feedback and negative customers*, *Appl. Math. J. Chin. Univ.* 40(1) (2025), 1–19.

[5] Boussaha, Z., Oukid, N., Zeghdoudi, H., Soualhi, S. and Djellab, N. *On the M/G/1 feedback retrial queuing with orbital search of customers*, *Adv. Math. Sci. J.* 11(8) (2022), 723–739.

[6] Chandrika, K.U. and Sumitha, D. *Performance measures of feedback batch arrival retrial queue with optional second service and orbital search*, *Adv. Appl. Res.* 3(2) (2011), 179–186.

[7] Chang, F.M., Liu, T.H. and Ke, J.C. *On an unreliable-server retrial queue with customer feedback and impatience*, *Appl. Math. Model.* 55 (2018), 171–182.

[8] Choudhury, G. *A two phase batch arrival retrial queuing system with Bernoulli vacation schedule*, *Appl. Math. Comput.* 188(2) (2007), 1455–1466.

[9] Cooper, R.B. *Introduction to Queueing Theory*, Elsevier (North-Holland), Amsterdam, 1981.

[10] Dimitriou, I. *A single server retrial queue with event-dependent arrival rates*, *Ann. Oper. Res.* 331(2) (2023), 1053–1088.

[11] Eberhart, R. and Kennedy, J. *Particle swarm optimization*, *Proc. IEEE Int. Conf. Neural Netw.* 4 (1995), 1942–1948.

[12] Falin, G.I. *A survey of retrial queues*, *Queueing Syst.* 7 (1990), 127–167.

[13] Gelenbe, E. *Random neural networks with negative and positive signals and product form solution*, *Neural Comput.* 1(4) (1989), 502–510.

[14] Gómez-Corral, A. *Stochastic analysis of a single server retrial queue with general retrial times*, *Naval Res. Logist.* 46(5) (1999), 561–581.

[15] Hu, K., Allon, G. and Bassamboo, A. *Understanding customer retrials in call centers: Preferences for service quality and service speed*, *Manuf. Serv. Oper. Manag.* 24(2) (2022), 1002–1020.

[16] Jain, M., Sharma, G.C. and Sharma, R. *A batch arrival retrial queuing system for essential and optional services with server breakdown and Bernoulli vacation*, *Int. J. Internet Enterp. Manag.* 8 (2012), 16–45.

[17] Keerthiga, S. and Indhira, K. *Analysis of single server feedback retrial queue with Bernoulli working vacation and starting failure*, *Reliab. Theory Appl.* 19(3(79)) (2024), 310–326.

[18] Krishnamoorthy, A., Deepak, T.G. and Joshua, V.C. *An M/G/1 retrial queue with non-persistent customers and orbital search*, *Stoch. Anal. Appl.* 23(5) (2005), 975–997.

[19] Kumar, B. *A retrial queuing system with negative arrival under Bernoulli vacation schedule and optional repair*, (2024) PREPRINT (Version 1) available at Research Square https://doi.org/10.21203/rs.3.rs-4591933/v1

[20] Micheal, M.N.G. and Indhira, K. *A literature review on retrial queuing system with Bernoulli vacation*, *Yugosl. J. Oper. Res.* 34(1) (2024), 109–134.

[21] Mokaddis, G.S., Metwally, S.A. and Zaki, B.M. *A feedback retrial queuing system with starting failures and single vacation*, *J. Appl. Sci. Eng.* 10(3) (2007), 183–192.

[22] Radha, J., Indhira, K. and Chandrasekaran, V.M. *A repairable M[X]/Gk/1 feedback retrial G-queue, balking and reneging, search of customers with vacation*, *Int. J. Knowl. Manag. Tour. Hosp.* 1(4) (2017), 457–476.

[23] Rajadurai, P., Saravanarajan, M.C. and Chandrasekaran, V.M. *Analysis of an M/G/1 feedback retrial queue with unreliable server, non-persistent customers, single working vacation and vacation interruption*, *Int. J. Serv. Oper. Manag.* 24(2) (2016), 235–266.

[24] Rusilko, T.V. and Pankov, A.V. *Queuing network model of a call center with customer retrials and impatient customers*, *Izv. Saratov Univ. Math. Mech. Inform.* 24(2) (2024), 287–297.

[25] Seenivasan, M., Ramesh, R., Chandralekha, S. and Pandisaranya, S. *Single server retrial queuing system with reneging, feedback and balking customers*, *Proc. 1st Int. Conf. Innov. Commun. Electr. Comput. Eng. (ICICEC)* (2024), 1–7.

[26] Singla, N. and Kaur, H. *A two-state retrial queuing model with feedback and balking*, *Reliab. Theory Appl.* 16(SI 2(64)) (2021), 142–155.

[27] Van Do, T. *Bibliography on G-networks, negative customers and applications*, *Math. Comput. Model.* 53(1–2) (2011), 205–212.