RAIRO-Oper. Res. 55 (2021) S1229–S1248
https://doi.org/10.1051/ro/2020103

RAIRO Operations Research
www.rairo-ro.org

ON THE RETRIAL QUEUE WITH IMPERFECT COVERAGE AND DELAY REBOOT

Tzu-Hsin Liu[1], Jau-Chuan Ke[2,*], Ching-Chang Kuo[2] and Fu-Min Chang[1]

Abstract. Retrial systems have been used extensively to model many practical problems in call center, data center, cloud service computing center and computer network system. This paper deals with a multi-server retrial system with the features of imperfect coverage and delay reboot. In the investigated system, arrivals may not be detected because of some fault issues. When this situation happened, the system is cleared by a reboot operation. Once arrivals are detected and located, they are attended to when a server is available; otherwise, they join a retrial orbit and generate repeated attempts till a free server is found. We analyze the presented model as a quasi-birth-and-death process and develop various performance indices. The optimal number of servers and optimal service rate are searched by constructing an average cost function. A heuristic search technique is employed to obtain the optimization approximate solution at a minimum cost. Numerical illustrations are given to demonstrate the optimization procedure and the effects of varying parameters on performance indices. We also present an application example to demonstrate the applicability of investigated model.

Mathematics Subject Classification. 90B22, 60K25.

Received February 17, 2020. Accepted September 18, 2020.

1. Introduction

In the practical applications of many fault-tolerant systems such as call centers, web access, telecommunication networks and computer systems, the customer retrial phenomenon, in which arriving customers may be immediately unable to enter the service area, is a common problem. Fault tolerance has to ensure availability as well as continued service even when some parts of a system fail. In many service systems, an extensive loss of customers as well as cost occurs due to the slow response of some customer complaints if not tackled properly with the help of suitable mechanism. But in some practical situations, the customer complaint device may prove inadequate to recover a customer complaint perfectly. These types of situations are called as imperfect coverage. When this situation happened, the missed customer complaint needs time to be found and cleared. And then the customer complaint can be handled continuously. These types of situations are called as reboot operation. Most of the studies focus on the machine repair problems with imperfect coverage in fault tolerance systems; however, the research of the customer's behaviour gets less attention on fault tolerance systems. As a multimedia contact center (MCC), the system focuses on the customer's behaviours. This prompts us to study

Keywords. Imperfect coverage, delay reboot, retrial queue, optimization, sensitivity analysis.

[1] Department of Finance, Chaoyang University of Technology, Taichung, Taiwan, ROC.

[2] Department of Applied Statistics, National Taichung University of Science and Technology, Taichung, Taiwan, ROC.

[*]Corresponding author: jauchuan@nutc.edu.tw

a multi-server retrial system with the features of imperfect coverage and delay reboot. In addition, the discussed system has a potential application in a fault tolerant incoming call center system which provides an arriving call service and is switched to route calls by the ACD (Automatic Call Distributor). ACD is a system that automatically detects incoming calls and then distributes them to relevant consultants, who provide the type of information or service required. Those consultants can be treated as servers. Sometimes arriving calls are not detected and located due to fault issues. It needs time to be found and cleared as such a situation occurred. For such an application, the administrator is interesting to search the optimum number of consultants and the optimum mean service rate at a minimum cost. The details of application example are introduced in Section 7.

Many eminent researchers in recent years have investigated many retrial queueing systems with variations. First, Gómez-Corral [12] provided a bibliographical guide for the study of queues with repeated attempts. Artalejo *et al.* [6] considered a multiserver type discrete-time queue with exponential retrials and geometric repeated attempts. Artalejo and Pla [4] investigated a multiserver system with customer impatience in telecommunication systems, and provided two truncation methods to analyze the retrial queue. Phung-Duc *et al.* [24–26] presented some analytical approach to stationary behaviour for multiserver retrial queues, which can be calculated by a numerically stabilizing process. Artalejo [2] further gave a detailed bibliography on retrial queues during 2000–2009. Some other results about a retrial queue can be found in [7–9, 11, 19, 23, 27, 32, 36].

Another interesting aspect is queuing systems with imperfect coverage. The probabilities of successful detection, location and recovery are included in the coverage factor [28]. Many studies on repairable systems with imperfect coverage have been conducted, such as [1, 13–17, 20, 31, 33, 34]. Recently, Yen and Wang [37] compared reliability and availability measures for three different systems with imperfect coverage and spare switching failures. Ke *et al.* [18] studied a machine repairing problem with an unreliable repairman and imperfect switchover of standbys. They obtained the stationary probability distribution by the method of supplementary variable. To the best of our knowledge, there is no works on a multi-server retrial queue with imperfect coverage and reboot delay, and such a system is not easily to be constructed and analyzed.

Following sections are organized as: Section 2 gives the model description in detail. For such model, the stability condition is derived. Furthermore, the stationary probability vectors are obtained by using matrix-geometric approach in Section 3. Section 4 shows the various performance indices in terms of matrix form. In Section 5, we develop a structure of average cost. A heuristic search technique is implemented to optimize an approximate solution at a minimum cost. In Section 6, some numerical illustrations are provided to examine the effect of system parameters on the performance indices and average cost. In the final, we illustrate a potential application and make a conclusion.

2. Retrial system model

In this system, arriving customers are assumed to follow a Poisson streams with rate λ. There are c identical servers in the system. The service time has an exponential distribution with rate μ. When an arriving customer is coming into the system, it may be detected and located with a probability θ (coverage probability, see [30]). Once the arriving customer is detected successfully by detection device, it will be transferred to one available server for serving. When all of the servers are occupied, the arriving customer will stay in the orbit. The access from the orbit is governed by the classical retrial policy. That is, each arriving customer in orbit repeats its request after an exponentially time of rate σ, so the retrial rate given that j arrivals in the orbit is $\sigma_j = j\sigma$, which is the action of retrial in the system. On the other hand, when an arriving customer or retrying customer is not detected successfully by detection device cause of the software or hardware errors, the system will go into an unsafe state. To clear the unsafe state, the non-detective customer needs to be located. To do that, the supervisor will issue a reboot operation to reboot the detection device and then locate the non-detective customer. After the non-detective customer being located, it will be treated as normal customer and continue his arriving process, and the system returns to the normal state. The reboot delay is exponentially distributed with rate β. During reboot delay, arriving customers and retrials are ignored, and servers stop their work. As we know, closed analytical solutions of multiserver retrial queues and direct algorithmic computations of these

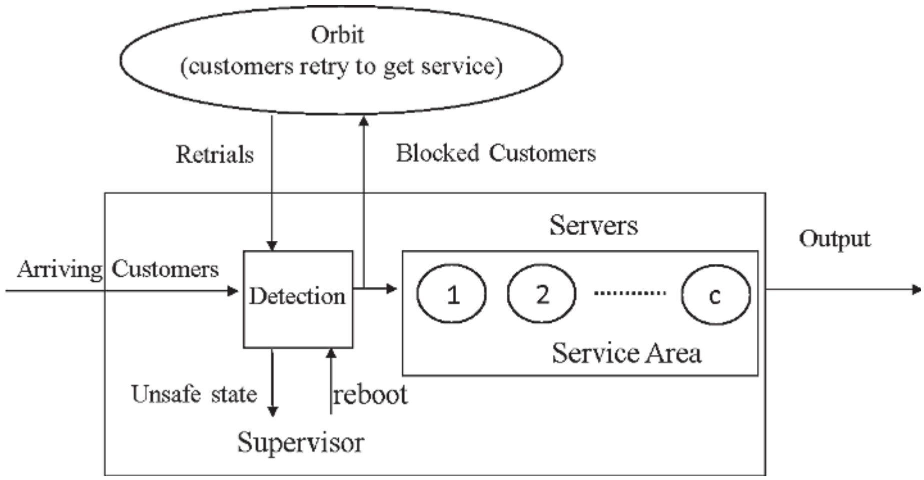

FIGURE 1. The general structure of investigated retrial queueing model.

limiting probabilities are still unavailable (see [5]). Therefore, we use the Neuts and Rao truncation technique to approximate the stationary results in this manuscript. That is, considering the specifications of system, the number of arrivals in the orbit allowed to make conduct retrials is limited up to a predetermined threshold N. This means that the number of arrivals in orbit is greater than the specified value N, the retrial rates do not depend on the length of orbit. Thus, the retrial rate becomes $\sigma_j = j\sigma$ if $j \leq N$ and $\sigma_j = N\sigma$ if $j > N$. About discussion for the choice of N, one can see the book by Artalejo and Gómez-Corral [3] on retrial queues. The general structure of investigated retrial queueing model is shown in Figure 1.

The transition diagram of this system is provided in Figure 2 for the case of $c = 3$. The states of the system at time t can be characterized by the process $\{(Q_1(t), Q_2(t), S(t)); t \geq 0\}$ where $Q_1(t)$ represents the number of busy consultants, $Q_2(t)$ indicates the number of calls in orbit, and $S(t)$ is defined as:

$$S(t) = \begin{cases} 0, \text{ the system is in safe state} \\ 1, \text{ the system is in unsafe state.} \end{cases}$$

Apparently, the process $\{(Q_1(t), Q_2(t), S(t)); t \geq 0\}$ composes a continuous-time Markov chain on the state space $\Delta = \{(i, n, k) | i = 0, 1, \ldots, c-k, n = 0, 1, \ldots, k = 0, 1\}$. Denote $P^k_{n,i}$, $(i, n, k) \in \Delta$, by the stationary distribution of the Markov chain $\{(Q_1(t), Q_2(t), S(t)); t \geq 0\}$. Referring to the state-transition diagram, the steady-state equations for $P^k_{n,i}$ are

$$\begin{aligned}
\lambda P^0_{0,0} &= \mu P^0_{0,1}, \\
(\lambda + i\mu) P^0_{0,i} &= \lambda\theta P^0_{0,i-1} + (i+1)\mu P^0_{0,i+1} + \beta P^1_{0,i-1} + \sigma_1\theta P^0_{1,i-1}, \quad 1 \leq i \leq c-1, \\
(\lambda + c\mu) P^0_{0,c} &= \lambda\theta P^0_{0,c-1} + \beta P^1_{0,c-1} + \sigma_1\theta P^0_{1,c-1}, \\
\beta P^1_{0,i} &= \lambda(1-\theta) P^0_{0,i} + \sigma_1(1-\theta) P^0_{1,i}, \quad 0 \leq i \leq c-1, \\
(\lambda + \sigma_n) P^0_{n,0} &= \mu P^0_{n,1}, \quad n \geq 1, \\
(\lambda + i\mu + \sigma_n) P^0_{n,i} &= \lambda\theta P^0_{n,i-1} + (i+1)\mu P^0_{n,i+1} + \beta P^1_{n,i-1} + \sigma_{n+1}\theta P^0_{n+1,i-1}, \quad 1 \leq i \leq c-1, \quad n \geq 1, \\
(\lambda + c\mu) P^0_{n,c} &= \lambda P^0_{n-1,c} + \lambda\theta P^0_{n,c-1} + \beta P^1_{n,c-1} + \sigma_{n+1}\theta P^0_{n+1,c-1}, \quad n \geq 1, \\
\beta P^1_{n,i} &= \lambda(1-\theta) P^0_{n,i} + \sigma_{n+1}(1-\theta) P^0_{n+1,i}, \quad 0 \leq i \leq c-1, \quad n \geq 1.
\end{aligned}$$

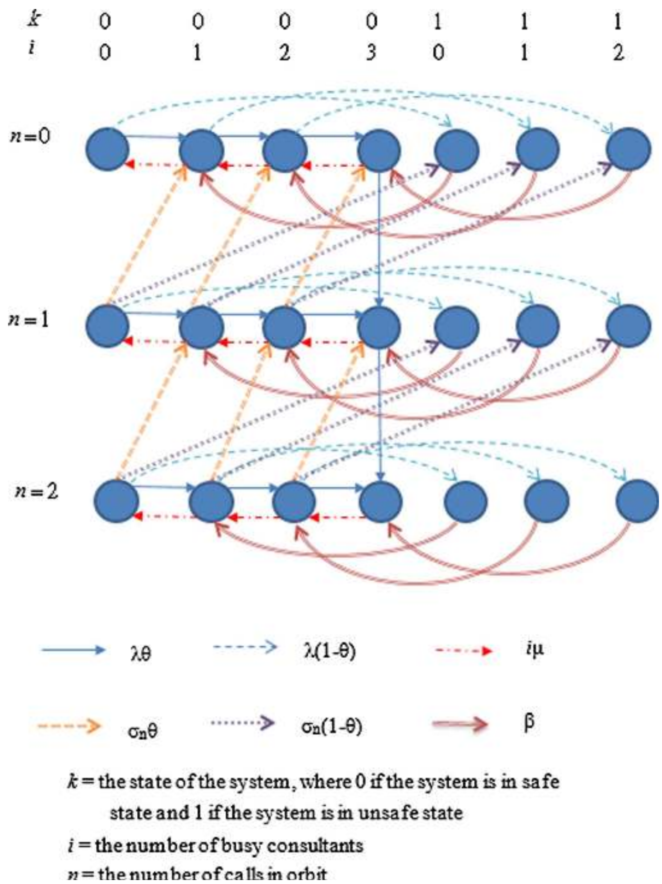

FIGURE 2. The transition diagram of this system for the case of $c = 3$.

To analyze the resulting system of the above linear equations, a matrix-geometric approach is used. The infinitesimal generator $\mathbf{Q}$ of this Markov chain is of the form (see [21]):

$$\mathbf{Q} = \begin{bmatrix} \mathbf{D}_0 & \mathbf{U} & & & & & & \\ \mathbf{L}_1 & \mathbf{D}_1 & \mathbf{U} & & & & & \\ & \ddots & \ddots & \ddots & & & & \\ & & \mathbf{L}_{N-1} & \mathbf{D}_{N-1} & \mathbf{U} & & & \\ & & & \mathbf{L}_N & \mathbf{D}_N & \mathbf{U} & & \\ & & & & \mathbf{L}_N & \mathbf{D}_N & \mathbf{U} & \\ & & & & & \ddots & \ddots & \ddots \end{bmatrix} \tag{2.1}$$

in which each block is square matrix with dimension $(2c+1)\times(2c+1)$. The elements of $\mathbf{U}$, $\mathbf{L}_j(1\le j\le N)$ and $\mathbf{D}_j(0\le j\le N)$ are described as:

$$\mathbf{U}=\begin{matrix}1\\ \vdots\\ c\\ c+1\\ c+2\\ \vdots\\ 2c+1\end{matrix}\begin{bmatrix}0&\cdots&0&0&0&\cdots&0\\ \vdots&\ddots&\vdots&\vdots&\vdots&\ddots&\vdots\\ 0&\cdots&0&0&0&\vdots&0\\ 0&\cdots&0&\lambda&0&\cdots&0\\ 0&\cdots&0&0&0&\cdots&0\\ \vdots&\ddots&\vdots&\vdots&\vdots&\ddots&\vdots\\ 0&\cdots&0&0&0&\cdots&0\end{bmatrix}$$

$$\mathbf{L}_j=\begin{matrix}1\\ \vdots\\ c\\ c+1\\ \vdots\\ 2c+1\end{matrix}\begin{bmatrix}0&\sigma_j\theta&\cdots&0&\sigma_j(1-\theta)&\cdots&0\\ \vdots&\vdots&\ddots&\vdots&\vdots&\ddots&\vdots\\ 0&0&\cdots&\sigma_j\theta&0&\vdots&\sigma_j(1-\theta)\\ 0&\cdots&0&0&0&\cdots&0\\ \vdots&\vdots&\ddots&\vdots&\vdots&\ddots&\vdots\\ 0&0&\cdots&0&0&\cdots&0\end{bmatrix},\quad 1\le j\le N,$$

$$\mathbf{D}_j=\begin{bmatrix}\mathbf{A}_{11}&\mathbf{A}_{12}\\ \mathbf{A}_{21}&\mathbf{A}_{22}\end{bmatrix},\quad 0\le j\le N,$$

$$\mathbf{A}_{11}=\begin{bmatrix}-\alpha_0&\lambda\theta&\cdots&0&0\\ \mu&-\alpha_1&\cdots&0&0\\ \vdots&\vdots&\ddots&\vdots&\vdots\\ 0&0&\cdots&-\alpha_{c-1}&\lambda\theta\\ 0&0&\cdots&c\mu&-\alpha_c\end{bmatrix}_{(c+1)\times(c+1)}$$

$$\mathbf{A}_{12}=\begin{bmatrix}\lambda(1-\theta)&0&\cdots&0\\ 0&\lambda(1-\theta)&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\ 0&0&\cdots&\lambda(1-\theta)\\ 0&0&\cdots&0\end{bmatrix}_{(c+1)\times c}$$

$$\mathbf{A}_{21}=\begin{bmatrix}0&\beta&0&0\\ 0&0&\beta&0\\ \vdots&\vdots&\cdots&\vdots\\ 0&0&0&\beta\end{bmatrix}_{c\times(c+1)}$$

$$\mathbf{A}_{22} = \begin{bmatrix} -\beta & 0 & \cdots & 0 \\ 0 & -\beta & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\beta \end{bmatrix}_{c \times c}$$

where the elements $\alpha_i (i = 1, 2, \ldots, c)$ in $\mathbf{A}_{11}$ and $\mathbf{D}_j$ is given by

$$\alpha_i = \begin{cases} \lambda + i\mu + \sigma_j, & i = 0, 1, 2, \ldots, c-1, \quad 1 \le j \le N \\ \lambda + c\mu, & i = c, \quad 1 \le j \le N. \end{cases}$$

As is shown above, our model has the standard structure of a QBD (quasi-birth-death) process with infinitesimal generator $\mathbf{Q}$. Let $\mathbf{P} = [\mathbf{P}_0, \mathbf{P}_1, \ldots]$ represent the stationary probability vector of the generator $\mathbf{Q}$, where $\mathbf{P}_n = \left[P^0_{n,0}, \ldots, P^0_{n,c}, P^1_{n,0}, \ldots, P^1_{n,c-1}\right]$, $n = 0, 1, 2, \ldots$. Thus, we could use the matrix-geometric solution to derive the stationary probabilities.

3. Matrix-geometric analysis

Now we focus on the stationary analysis by using the matrix-geometric approach. Firstly, the stability condition is determined. If the stability condition is satisfied, then the retrial system can be analyzed in stationary state.

3.1. Stability condition

The stationary probability vector of the generator matrix, $\mathbf{H}$, is denoted by π; that is, $\pi\mathbf{H} = \mathbf{0}$, $\pi\mathbf{e} = 1$ where $\mathbf{e}$ represents a column vector with suitable size such that all components equal to one and

$$\mathbf{H} = \mathbf{U} + \mathbf{L}_N + \mathbf{D}_N = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{bmatrix},$$

$$\mathbf{H}_{11} = \begin{bmatrix} -\alpha_0 & \theta(\lambda+\sigma_N) & \cdots & 0 & 0 \\ \mu & -\alpha_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\alpha_{c-1} & \theta(\lambda+\sigma_N) \\ 0 & 0 & \cdots & c\mu & -\alpha_c \end{bmatrix}_{(c+1)\times(c+1)}$$

$$\mathbf{H}_{12} = \begin{bmatrix} (1-\theta)(\lambda+\sigma_N) & 0 & \cdots & 0 \\ 0 & (1-\theta)(\lambda+\sigma_N) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (1-\theta)(\lambda+\sigma_N) \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{(c+1)\times c}$$

$$\mathbf{H}_{21} = \begin{bmatrix} 0 & \beta & 0 & \cdots & 0 \\ 0 & 0 & \beta & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \beta \end{bmatrix}_{c\times(c+1)}$$

$$\mathbf{H}_{22} = \begin{bmatrix} -\beta & 0 & \cdots & 0 \\ 0 & -\beta & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\beta \end{bmatrix}_{c\times c}$$

where the elements $\alpha_i = \lambda + i\mu + \sigma_N$, $i = 0, 1, \ldots, c-1$.

The vector π partitioned as $\pi = \left[\pi_0^0, \ldots, \pi_c^0, \pi_0^1, \ldots, \pi_{c-1}^1\right]$ is yielded by solving the following equations:

$$\pi_i^0 = \frac{1}{i!}\left(\frac{\lambda+\sigma_N}{\mu}\right)^i \pi_0^0 = \frac{1}{i!}\rho^i\pi_0^0, \qquad i = 0, 1, \ldots, c,$$
$$\pi_i^1 = \frac{(1-\theta)(\lambda+\sigma_N)}{\beta}\pi_i^0 = \frac{1}{i!}(1-\theta)\xi\rho^i\pi_0^0, \quad i = 0, 1, \ldots, c-1$$

subject to $\sum_{i=0}^{c}\pi_i^0 + \sum_{i=0}^{c-1}\pi_i^1 = 1$, where $\rho = (\lambda+\sigma_N)/\mu$ and $\xi = (\lambda+\sigma_N)/\beta$.

Hence,

$$\pi_0^0 = \left\{\frac{e^\rho\Gamma(c+1,\rho)}{\Gamma(c+1)} + (1-\theta)\xi\frac{e^\rho\Gamma(c,\rho)}{\Gamma(c)}\right\}^{-1},$$

where $\Gamma(k,z) = \int_z^\infty e^{-t}t^{k-1}\mathrm{d}t$.

Neuts [22] had shown that the stationary probability vector $\mathbf{P} = [\mathbf{P}_0, \mathbf{P}_1, \ldots]$ exists if and only if the following holds good:

$$\pi\mathbf{U}\mathbf{e} < \pi\mathbf{L}_N\mathbf{e}.$$

It is easily verified that the stability condition of the system is given by (see Appendix A):

$$\lambda < c\sigma_N\rho^{-c}e^\rho\Gamma(c,\rho).$$

From the above formula, one can find that it is difficult to have intuitive explanations behind the stability condition. Hence, we only attempt to explain the case of a single server. The stability condition of a single server orbit queue with imperfect coverage is

$$\frac{\lambda}{\mu} < \frac{\sigma_N}{\lambda+\sigma_N}.$$

That is, the expected number of orbiting customers who enter service successfully should be greater than the traffic intensity of the traditional queueing system.

3.2. Computation of stationary probabilities

Let $\mathbf{P} = [\mathbf{P}_0, \mathbf{P}_1, \ldots]$ be the stationary probabilities of the generator $\mathbf{Q}$. Since $\mathbf{Q}$ is a level independent QBD process, its stationary probability vector is given by $\mathbf{P}_j = \mathbf{P}_N\mathbf{R}^{j-N}$, $j \geq N+1$, in which $\mathbf{R}$ is the minimal nonnegative solution of the matrix-quadratic equation $\mathbf{R}^2\mathbf{L}_N + \mathbf{R}\mathbf{D}_N + \mathbf{U} = \mathbf{0}$, whose spectral radius is less than one (see [22]). Using the Maple computer program and matrix algorithm, the explicit formula of the rate matrix **R** is developed as follows (the derivation is provided in Appendix B):

$$\mathbf{R} = \begin{array}{r} 1 \\ \vdots \\ c \\ c+1 \\ c+2 \\ \vdots \\ 2c+1 \end{array} \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ r_0^0 & r_1^0 & \cdots & r_c^0 & r_0^1 & \cdots & r_{c-1}^1 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

where

$$r_0^0 = \frac{1}{\rho} r_1^0,$$

$$r_i^0 = \frac{(i+1)\mu}{\lambda + i\mu + \sigma_N - (\lambda + r_c^0 \sigma_N)\phi} r_{i+1}^0, \quad i = 1, 2, \ldots, c-1,$$

$$\sigma_N \phi_{c-1} (r_c^0)^2 - (\lambda + c\mu - \lambda\phi_{c-1}) r_c^0 + \lambda = 0,$$

$$r_i^1 = \frac{(\lambda + r_c^0 \sigma_N)(1-\theta)}{\beta} r_i^0, \quad i = 1, 2, \ldots, c-1$$

$$\phi_i = \begin{cases} \frac{1}{\rho}, & i = 1 \\ \frac{(i+1)\mu}{\lambda + i\mu + \sigma_N - (\lambda + r_c^0 \sigma_N)\varphi_{i-1}}, & i = 2, 3, \ldots, c-1. \end{cases}$$

The r_i^k can be obtained from the equations listed above by recursive manners.

The size of N will affect the rate matrix **R**, which dominates the stationary probabilities **P** and performance indices. To understand the effect of the truncated parameter N, a numerical experiment is performed to investigate the relationship between the truncated parameter N and the spectral radius of **R**. The default parameters are chosen as $\lambda = 2.5, \mu = 5, \beta = 10, \theta = 0.7, \sigma = 1$. Figure 3 presents the numerical results. It can be seen from this figure that when the truncated parameter N exceeds a certain point, the improvement of the spectral radius of **R** is not obvious. Moreover, when the number of servers is large, the influence of the truncated parameter N on the spectral radius of **R** is not significant.

The steady-state equations can be represented in matrix form as follows:

$$\mathbf{P}_0\mathbf{D}_0 + \mathbf{P}_1\mathbf{L}_1 = \mathbf{0}, \tag{3.1}$$

$$\mathbf{P}_{j-1}\mathbf{U} + \mathbf{P}_j\mathbf{D}_j + \mathbf{P}_{j+1}\mathbf{L}_{j+1} = \mathbf{0}, \quad 1 \le j \le N-1, \tag{3.2}$$

$$\mathbf{P}_{N-1}\mathbf{U} + \mathbf{P}_N\mathbf{D}_N + \mathbf{P}_N\mathbf{R}\mathbf{L}_N = \mathbf{0}, \tag{3.3}$$

$$\mathbf{P}_N\mathbf{R}^{j-N-1}\mathbf{U} + \mathbf{P}_N\mathbf{R}^{j-N}\mathbf{D}_N + \mathbf{P}_N\mathbf{R}^{j-N+1}\mathbf{L}_N = \mathbf{0}, \quad j \ge N+1. \tag{3.4}$$

Equations (3.1)–(3.3) can be rewritten in the following:

$$\mathbf{P}_{j-1} = \mathbf{P}_j \varphi_j, \quad 1 \le j \le N-1, \tag{3.5}$$

$$\mathbf{P}_N(\varphi_N\mathbf{U} + \mathbf{D}_N + \mathbf{R}\mathbf{L}_N) = \mathbf{0}, \tag{3.6}$$

where

$$\varphi_j = \begin{cases} \mathbf{L}_1(-\mathbf{D}_0)^{-1}, & j = 1 \\ \mathbf{L}_j\left[-(\varphi_{j-1}\mathbf{U} + \mathbf{D}_{j-1})\right]^{-1}, & 2 \le j \le N. \end{cases}$$

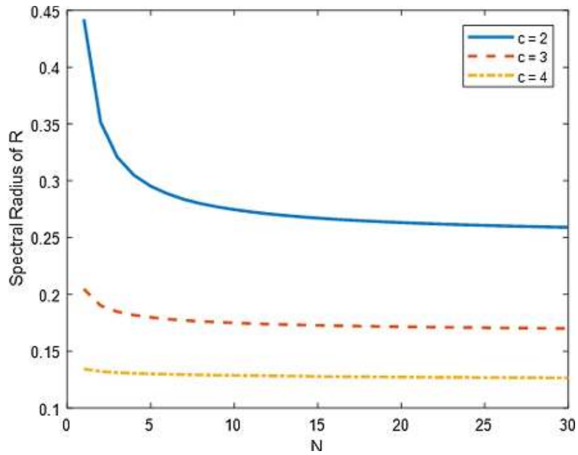

FIGURE 3. The spectral radius of $\mathbf{R}$ *vs.* truncated parameter N for $\lambda = 3$, $\mu = 5$, $\beta = 5$, $\theta = 0.7$, and $\sigma = 2.5$.

Consequently, the stationary probability vector $\mathbf{P}_N$ can be obtained by solving (3.6) and the following normalization condition:

$$\begin{aligned}\sum_{j=0}^{\infty} \mathbf{P}_j \mathbf{e} &= [\mathbf{P}_0 + \mathbf{P}_1 + \cdots + \mathbf{P}_{N-1} + \mathbf{P}_N + \mathbf{P}_{N+1} + \mathbf{P}_{N+2} + \cdots]\, \mathbf{e} \\ &= \left[\mathbf{P}_N \mathbf{\Phi}_1 + \mathbf{P}_N \mathbf{\Phi}_2 + \cdots + \mathbf{P}_N \mathbf{\Phi}_N + \mathbf{P}_N + \mathbf{P}_N \mathbf{R} + \mathbf{P}_N \mathbf{R}^2 + \cdots\right] \mathbf{e} \\ &= \mathbf{P}_N \left[\sum_{j=1}^{N} \mathbf{\Phi}_j + (\mathbf{I} - \mathbf{R})^{-1}\right] \mathbf{e} = 1, \end{aligned} \tag{3.7}$$

where $\mathbf{\Phi}_j = \prod_{k=j}^{N} \phi_k$. Once $\mathbf{P}_N$ is obtained, the prior state probabilities $[\mathbf{P}_0, \mathbf{P}_1, \ldots, \mathbf{P}_{N-1}]$ are computed from (3.5), and $[\mathbf{P}_{N+1}, \mathbf{P}_{N+2}, \ldots]$ are obtained by the formula $\mathbf{P}_j = \mathbf{P}_N \mathbf{R}^{j-N}$, $j \geq N+1$.

4. Performance indices

To evaluate the performance of the presented system, we derive the explicit expressions for different performance indices in the following:

– Expected number of failed machines in the orbit is:

$$L = \sum_{i=0}^{\infty} j \mathbf{\Pi}_j \mathbf{e} = \mathbf{P}_N \left[\sum_{j=1}^{N} j \mathbf{\Phi}_j + N(\mathbf{I} - \mathbf{R})^{-1} + \mathbf{R}(\mathbf{I} - \mathbf{R})^{-2}\right] \mathbf{e}.$$

– Expected number of busy consultants is:

$$E[B] = \mathbf{P}_N \left[\sum_{j=1}^{N} j \mathbf{\Phi}_j + (\mathbf{I} - \mathbf{R})^{-1}\right] [0, 1, \ldots, c, 0, 1, \ldots, c-1]^T.$$

– Expected number of idle consultants is:

$$E[I] = \mathbf{P}_N \left[\sum_{j=1}^{N} j\mathbf{\Phi}_j + (\mathbf{I}-\mathbf{R})^{-1} \right] [c, c-1, \ldots, 0, 1, c, \ldots, 1]^T .$$

– Probability that the system is in safe state is:

$$P_S = \mathbf{P}_N \left[\sum_{j=1}^{N} j\mathbf{\Phi}_j + (\mathbf{I}-\mathbf{R})^{-1} \right] \left[\overbrace{1, \ldots, 1}^{\#=c+1}, \overbrace{0, \ldots, 0}^{\#=c} \right]^T .$$

– Probability that the system is in unsafe state is:

$$P_{US} = \mathbf{P}_N \left[\sum_{j=1}^{N} j\mathbf{\Phi}_j + (\mathbf{I}-\mathbf{R})^{-1} \right] \left[\overbrace{0, \ldots, 0}^{\#=c+1}, \overbrace{1, \ldots, 1}^{\#=c} \right]^T .$$

5. Cost-minimum analysis

We develop an average cost function per unit time for a fault tolerant call center system with retrial queue and imperfect coverage, in which the number of servers (c) and the mean service rate (μ) are decision variables. Following the formation of the average cost function, our aim is to decide the optimal number of servers (c^*) and the optimal service rate (μ^*) so as to minimize the cost function. We consider the following cost elements associated with different activities:

$c_h \equiv$ holding cost per unit time per customer present in orbit;
$c_b \equiv$ cost per unit time per busy server;
$c_f \equiv$ cost per unit time of each available server;
$c_u \equiv$ cost incurred per unit time when the system is in unsafe state;
$c_s \equiv$ cost per customer served by a mean service rate μ.

Based on the above cost elements and the corresponding performance indices, the average cost function per unit time is constructed as below:

$$AC(c, \mu) = c_h L + c_b E[B] + c_u \beta P_{US} + c_f c + c_s \mu.$$

Due to the fact that this cost function is a nonlinear and highly complicated expression, it is arduous to obtain the optimal solution. To overcome this, a heuristic search technique is implemented to obtain an approximate solution at a minimum cost. We propose an algorithm for searching the optimum solution as below.

Procedure for searching the optimum solution

Input: λ, β, θ, σ, N.
Step 1. Set $c = 1$ and $AC(0, \mu_0^* = 0) = \infty$.
Step 2. Employ Quasi-Newton method to obtain the solution μ_{c-1}^*, μ_c^*, μ_{c+1}^* which minimizes the average cost function, respectively.
Step 3. Compute $AC(c-1, \mu_{c-1}^*)$, $AC(c, \mu_c^*)$ and $AC(c+1, \mu_{c+1}^*)$.
Step 4. If $AC(c-1, \mu_{c-1}^*) \geq AC(c, \mu_c^*)$ and $AC(c, \mu_c^*) < AC(c+1, \mu_{c+1}^*)$, then $c^* = c$, $\mu^* = \mu_c^*$, $AC(c^*, \mu^*) = AC(c, \mu_c^*)$ and go to Output.
Otherwise, make $c = c + 1$ and go to Step 2.
Output: c^*, μ^*, $AC(c^*, \mu^*)$.

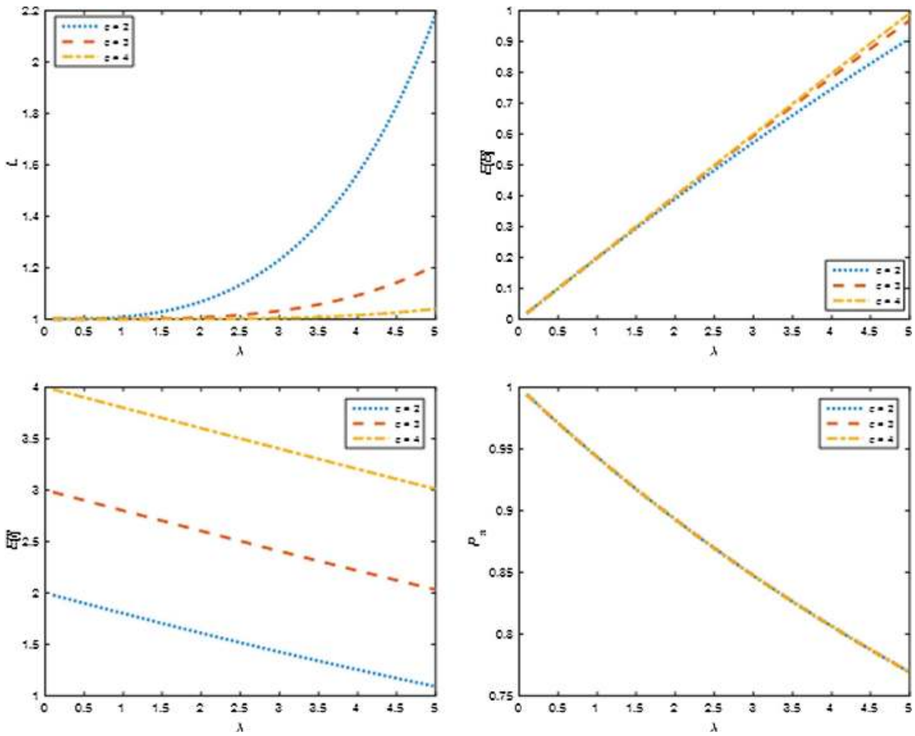

FIGURE 4. Performance indices *vs.* λ for $\mu = 5$, $\beta = 5$, $\theta = 0.7$, and $\sigma = 2.5$.

6. Numerical illustrations

Some numerical experiments are used to illustrate the influence of different parameters on the performance indices and average cost function.

6.1. Sensitivity analysis of the performance indices

Some figures are given to explore the influence of the varying parameters on the main performance indices of the presented system, under steady-state. The threshold $N = 20$ is chosen, and the default parameters are considered as follows:

Case 1: $\mu = 5, \beta = 10, \theta = 0.7, \sigma = 1$ with varying values of λ.
Case 2: $\lambda = 2.5, \beta = 10, \theta = 0.7, \sigma = 1$ with varying values of μ.
Case 3: $\lambda = 2.5, \mu = 3, \theta = 0.7, \sigma = 1$ with varying values of β.
Case 4: $\lambda = 2.5, \mu = 3, \beta = 10, \sigma = 1$ with varying values of θ.

Figure 4 shows the effect of λ on the performance indices. It is observed that L and $E[B]$ increase with increasing trend of λ, but $E[I]$ and P_S decrease. As the value of λ increases, it means that the mean arrival time decreases. From the moment the server is idle, the arriving customer and the retrial customer compete for access to the server. Moreover, the shorter the mean arrival time, the more likely the server is busy, which increases L and decreases P_S. In Figure 5, with the vary of service rate μ, the curves of the performance indices are provided. We find that L and $E[B]$ decrease with increasing trend of μ but $E[I]$ increases. Moreover, the

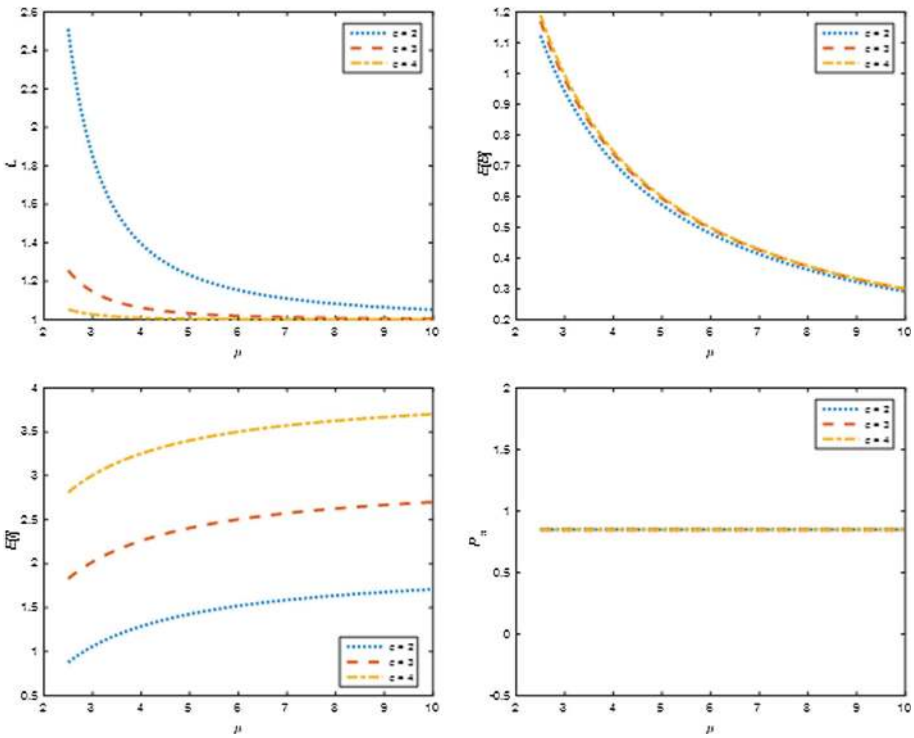

FIGURE 5. Performance indices *vs.* μ for $\lambda = 3$, $\beta = 5$, $\theta = 0.7$, and $\sigma = 2.5$.

effect of μ on P_S is not obvious. Figures 6 and 7 show the effect of β and θ on the performance indices. They indicate that the greater the value of β and θ is, the bigger the value of $E[B]$ and P_S. We also find that the effects of β and θ on L and $E[I]$ are not obvious. Furthermore, P_S is insensitive to change in c. As expected that a call is immediately detected on the call center and does not directly relevant of consultants.

6.2. Optimal analysis of the average cost function

The effect of the parameters on the average cost function is presented, in this sub-section, through some numerical results. The default values of various cost elements considered are $c_h = \$10$, $c_b = \$75$, $c_f = \$80$, $c_u = \$60$ and $c_S = \$8$. Figures 8 and 9 show the effects of some parameters such as μ and c on the average cost function. The other parameters for Figures 8 and 9 are set as follows:

Figure 8: $\lambda = 2.5, \beta = 10, \theta = 0.8, \sigma = 1, c = 3$ with varying values of μ.

Figure 9: $\lambda = 2.5, \beta = 10, \theta = 0.8, \sigma = 1, \mu = 3$ with varying values of c.

From Figures 8 and 9, the average cost function seems to be convex with respect to both μ and c. We also can experience that cost to the system is increased by more number of servers. Furthermore, Table 1 displays the optimal values of μ and c and their respective minimum cost. The default parameters are chosen as $\beta = 10$, $\theta = 0.7$, $\sigma = 1$. As expected intuitively, the average cost function increases with an increased λ. As λ becomes larger, the optimal mean service rate μ^* and the optimal number of the servers c^* also increase to maintain the service quality and the acceptable total cost.

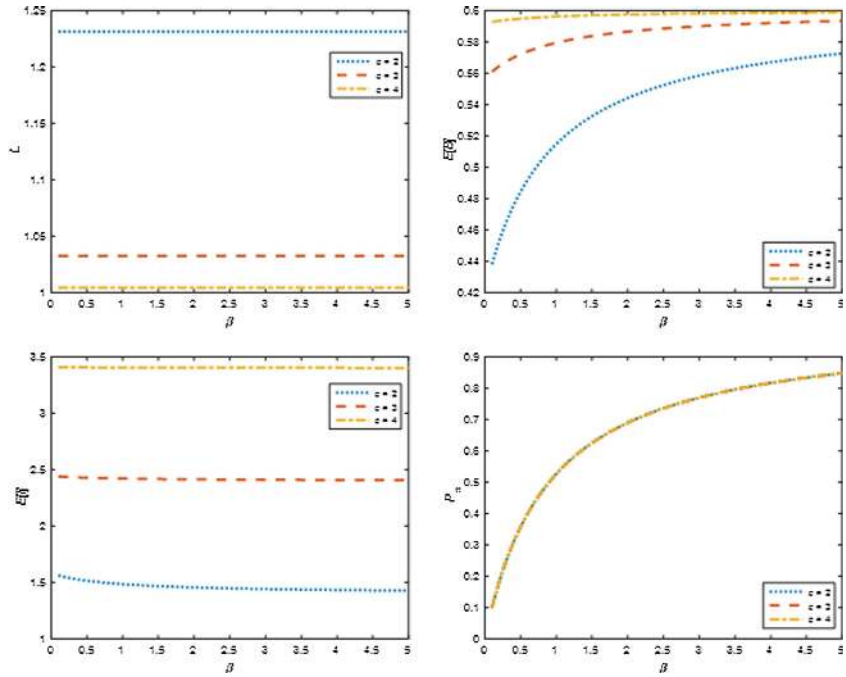

FIGURE 6. Performance indices *vs.* β for $\lambda = 3$, $\mu = 5$, $\theta = 0.7$, and $\sigma = 2.5$.

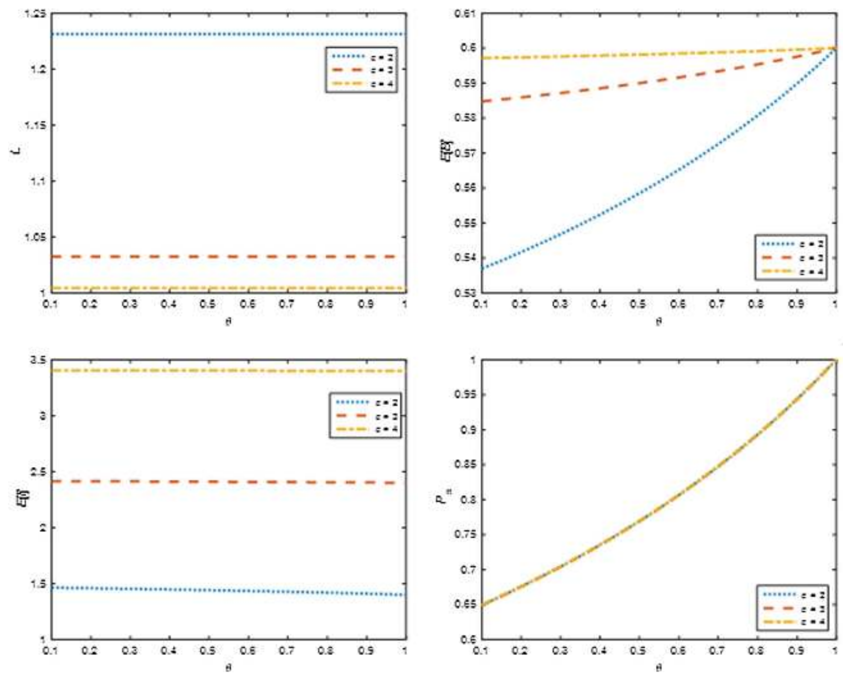

FIGURE 7. Performance indices *vs.* θ for $\lambda = 3$, $\mu = 5$, $\beta = 5$, and $\sigma = 2.5$.

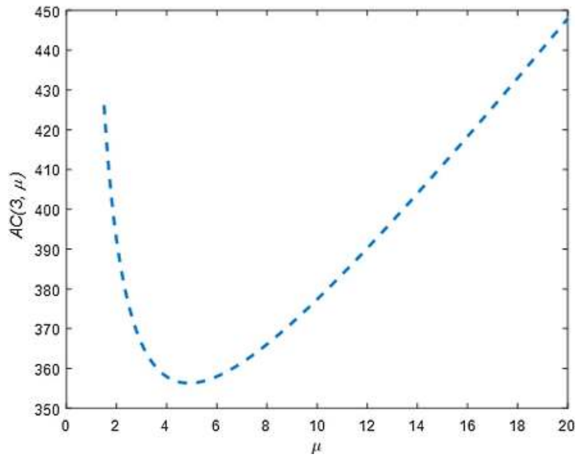

FIGURE 8. Plot of the average cost function $AC(3, \mu)$ *versus* the mean service rate.

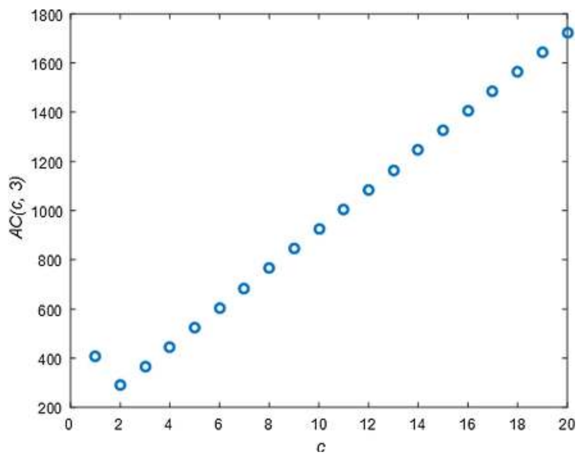

FIGURE 9. Plot of the average cost function $AC(c, 3)$ *versus* the mean service rate.

TABLE 1. Average cost function for different λ with $\beta = 10$, $\theta = 0.7$, and $\sigma = 1$.

λ	$AC(c^*, \mu^*)$	μ^*	c^*
0.5	135.32	2.35	1
1	161.41	3.56	1
1.5	184.87	4.67	1
2	274.61	4.52	2
2.5	291.40	5.12	2
3	307.32	5.70	2

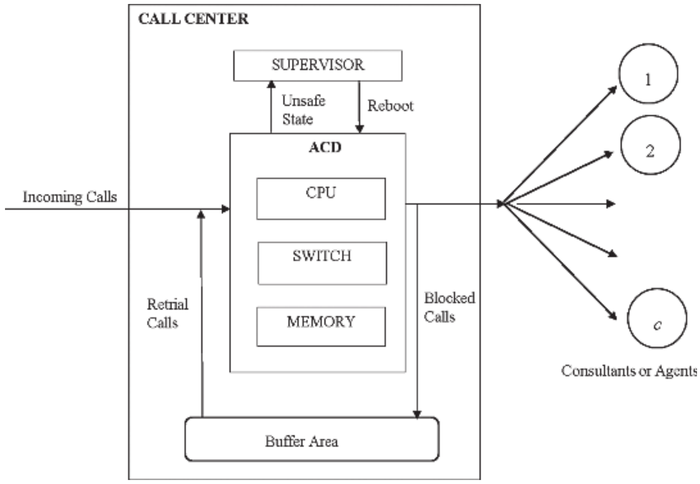

FIGURE 10. Incoming call center diagram.

7. APPLICATION EXAMPLE

The investigated system has a potential application in a fault tolerant call center system. Fault tolerances are of utmost importance in call center systems design and operation. We investigate an incoming call center which provides an arriving call service and is switched to route calls by the ACD (Automatic Call Distributor). ACD is a system that automatically detects the incoming calls and then distributes them to relevant consultants, who provide the type of information or service required. Generally, the incoming calls (arriving calls) arrive at the incoming call center with a Poisson stream with a rate $\lambda = 2.5$ calls/s and are placed by customers calling in to a center. Basically, an incoming call may be detected by ACD and located with a probability $\theta = 0.7$. Once the call is detected successfully by ACD, ACD will transfer the call to the consultant or agent. The incoming calls may not be detected and located by ACD because of some fault issues, such as fault detection and call-failure scenarios. When such a situation occurred, the system will go into an unsafe state. To clear the unsafe state, the non-detective call needs to be located. To do that, the supervisor will issue a reboot operation to reboot the ACD and then locate the non-detective call. After the non-detective call being located, it will be treated as normal call and continue his process, and the system returns to the normal state. The reboot delay is exponentially distributed with rate $\beta = 10$. Additionally, when all of the consultants or agents are occupied, the incoming call will stay in the buffer area repeats its call after an exponentially distributed time with rate $\sigma = 1$ calls/s, which is the action of retrial in the system. To lower the processing overhead resulted from a large number of requests, the administrator may adopt threshold-based policy to restrict the number of calls allowed to make retrial. With this situation, the number of calls in the buffer area allowed to make retrials is limited to $N = 10$. A diagram of this incoming call center is shown in Figure 10.

The administrator is interesting to search the optimum number of consultants and the optimum mean service rate at a minimum cost. The used cost elements are defined and set in the following: the holding cost per unit time per call present in buffer area is \$10; the cost per unit time per busy consultant is \$75; the cost per unit time of each available consultant is \$80; the cost incurred per unit time when the system is in unsafe state is \$60; the cost per call served by a consultant with mean service rate is \$8. With these cost elements, we implement the program by MATLAB software. The program output gives us that the optimum number of consultants is 2 and the optimum mean service rate is 5.12 calls/s. The related minimum expected cost is \$291.40.

8. Conclusions and future work

Utilizing matrix-geometric methods, this study analyzed a multi-server retrial system with the features of imperfect coverage and delay reboot, and examined the model in steady-state and established the stability condition and stationary distribution of the system. In addition, various performance indices were expressed in the matrix form. An average cost function per unit time is optimized for finding the number of consultants and mean service rate. Numerical performances were also provided. By demonstrating an example of the application, we provided several managerial insights to assist system analysts for decision making. Based on the listed above results, there arise some interesting extensions of the model which we will investigate in the near future. One possible extension is to study fault tolerant call center systems where the servers are subject to breakdown. Another way to generalize the fault tolerant call center system is to study the system with more practical behavior of incoming calls (such as feedback, balking, etc.).

Appendix A. Prove that the stability condition of the system

The vector π partitioned as $\pi = \left[\pi_0^0, \ldots, \pi_c^0, \pi_0^1, \ldots, \pi_{c-1}^1\right]$ is yielded by solving the following equations:

$$\pi_i^0 = \frac{1}{i!}\left(\frac{\lambda+\sigma_N}{\mu}\right)^i \pi_0^0 = \frac{1}{i!}\rho^i \pi_0^0, \quad i = 0, 1, \ldots, c,$$
$$\pi_i^1 = \frac{(1-\theta)(\lambda+\sigma_N)}{\beta}\pi_i^0 = \frac{1}{i!}(1-\theta)\xi\rho^i\pi_0^0, \quad i = 0, 1, \ldots, c-1$$

subject to $\sum_{i=0}^{c}\pi_i^0 + \sum_{i=0}^{c-1}\pi_i^1 = 1$, where $\rho = (\lambda+\sigma_N)/\mu$ and $\xi = (\lambda+\sigma_N)/\beta$. Hence,

$$\pi_0^0 = \left\{\frac{e^\rho\Gamma(c+1,\rho)}{\Gamma(c+1)} + (1-\theta)\xi\frac{e^\rho\Gamma(c,\rho)}{\Gamma(c)}\right\}^{-1},$$

where $\Gamma(k,z) = \int_z^\infty e^{-t}t^{k-1}\mathrm{d}t$.

Neuts [22] had shown that the stationary probability vector $\mathbf{P} = [\mathbf{P}_0, \mathbf{P}_1, \ldots]$ exists if and only if the following holds good:

$$\pi\mathbf{U}\mathbf{e} < \pi\mathbf{L}_N\mathbf{e}.$$

Substituting the matrices $\mathbf{L}_N$, $\mathbf{U}$ and π into the above inequality, we have

$$\pi\mathbf{U}\mathbf{e} = \left[\pi_0^0, \ldots, \pi_c^0, \pi_0^1, \ldots, \pi_{c-1}^1\right] \begin{matrix} 1 \\ \vdots \\ c \\ c+1 \\ c+2 \\ \vdots \\ 2c+1 \end{matrix} \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \vdots & 0 \\ 0 & \cdots & 0 & \lambda & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \lambda\pi_c^0$$

$$\pi\mathbf{L}_N\mathbf{e} = \left[\pi_0^0, \ldots, \pi_c^0, \pi_0^1, \ldots, \pi_{c-1}^1\right] \begin{bmatrix} 0 & \sigma_N\theta & \cdots & 0 & \sigma_N(1-\theta) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N\theta & 0 & \vdots & \sigma_N(1-\theta) \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \sigma_N \sum_{i=0}^{c-1} \pi_i^0.$$

By $\pi\mathbf{U}\mathbf{e} < \pi\mathbf{L}_N\mathbf{e}$, we have the following inequality:

$$\frac{1}{c!}\lambda\rho^c\pi_0^0 < \sigma_N \sum_{i=1}^{c-1} \frac{1}{i!}\rho^i\pi_0^0.$$

It implies that

$$\frac{1}{c!}\lambda\rho^c < \sigma_N \frac{e^\rho\Gamma(c,\rho)}{(c-1)!}.$$

Hence, the stability condition of the system is given by

$$\lambda < c\sigma_N\rho^{-c}e^\rho\Gamma(c,\rho).$$

Appendix B. The derivation of rate matrix R

Due to the structure of $\mathbf{U}$, $\mathbf{L}_N$ and $\mathbf{D}_N$ matrices, $\mathbf{R}$ is a matrix with the form

$$\mathbf{R} = \begin{matrix} 1 \\ \vdots \\ c \\ c+1 \\ c+2 \\ \vdots \\ 2c+1 \end{matrix} \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ r_0^0 & r_1^0 & \cdots & r_c^0 & r_0^1 & \cdots & r_{c-1}^1 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Substituting this $\mathbf{R}$ above into the matrix-quadratic equation $\mathbf{R}^2\mathbf{L}_N + \mathbf{R}\mathbf{D}_N + \mathbf{U} = \mathbf{0}$, we have

$$\mathbf{R}^2\mathbf{L}_N = \begin{matrix} 1 \\ \vdots \\ c \\ c+1 \\ c+2 \\ \vdots \\ 2c+1 \end{matrix} \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & r_0^0 r_c^0\sigma_N\theta & \cdots & r_{c-1}^0 r_c^0\sigma_N\theta & r_0^0 r_c^0\sigma_N(1-\theta) & \cdots & r_{c-1}^0 r_c^0\sigma_N(1-\theta) \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$
\mathbf{RD}_N = \begin{array}{c} 1 \\ \vdots \\ c \\ c+1 \\ c+2 \\ \vdots \\ 2c+1 \end{array}
\begin{bmatrix}
0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
a_0 & a_1 & \cdots & a_c & b_0 & \cdots & b_{c-1} \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0
\end{bmatrix}
$$

$$
\mathbf{U} = \begin{array}{c} 1 \\ \vdots \\ c \\ c+1 \\ c+2 \\ \vdots \\ 2c+1 \end{array}
\begin{bmatrix}
0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & 0 & \vdots & 0 \\
0 & \cdots & 0 & \lambda & 0 & \cdots & 0 \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0
\end{bmatrix}
$$

where

$$
a_i = \begin{cases}
-(\lambda+\sigma_N)r_0^0 + \mu r_1^0, & i = 0 \\
\lambda\theta r_{i-1}^0 - (\lambda + i\mu + \sigma_N)r_i^0 + (i+1)\mu r_{i+1}^0 + \beta r_{i-1}^1, & i = 1, 2, \ldots, c-1 \\
\lambda\theta r_{c-1}^0 - (\lambda + c\mu)r_c^0 + \beta r_{c-1}^1, & i = c
\end{cases}
$$

$$
b_i = \lambda(1-\theta)r_i^0 - \beta r_i^1, i = 0, 1, \ldots, c-1.
$$

Hence, we have the following set of equations:

$$
\begin{aligned}
-(\lambda+\sigma_N)r_0^0 + \mu r_1^0 &= 0, \\
\lambda\theta r_{i-1}^0 - (\lambda + i\mu + \sigma_N)r_i^0 + (i+1)\mu r_{i+1}^0 + \beta r_{i-1}^1 + r_c^0 r_{i-1}^0 \sigma_N \theta &= 0, \quad i = 1, 2, \ldots, c-1, \\
\lambda\theta r_{c-1}^0 - (\lambda + c\mu)r_c^0 + \beta r_{c-1}^1 + r_c^0 r_{c-1}^0 \sigma_N \theta + \lambda &= 0, \\
\lambda(1-\theta)r_i^0 - \beta r_i^1 + r_c^0 r)i^0 \sigma_N (1-\theta) &= 0, \quad i = 1, 2, \ldots, c-1.
\end{aligned}
$$

Solve the above set of equations,

$$
\begin{aligned}
r_0^0 &= \frac{1}{\rho} r_1^0, \\
r_i^0 &= \frac{(i+1)\mu}{\lambda + i\mu + \sigma_N - (\lambda + r_c^0 \sigma_N)\phi} r_{i+1}^0, \quad i = 1, 2, \ldots, c-1, \\
&\sigma_N \phi_{c-1} (r_c^0)^2 - (\lambda + c\mu - \lambda\phi_{c-1}) r_c^0 + \lambda = 0, \\
r_i^1 &= \frac{(\lambda + r_c^0 \sigma_N)(1-\theta)}{\beta} r_i^0, \qquad i = 1, 2, \ldots, c-1 \\
\phi_i &= \begin{cases}
\frac{1}{\rho}, & i = 1 \\
\frac{(i+1)\mu}{\lambda + i\mu + \sigma_N - (\lambda + r_c^0 \sigma_N)\varphi_{i-1}}, & i = 2, 3, \ldots, c-1.
\end{cases}
\end{aligned}
$$

Acknowledgements. The authors gratefully acknowledge the constructive comments of editors and the anonymous reviewers. This research was partially supported by the Ministry of Science and Technology of Taiwan under grant MOST 108-2221-E-025-003-MY3.

References

[1] S.V. Amari, A.F. Myers, A. Rauzy and K.S. Trivedi, Imperfect coverage models: status and trends. In: *Handbook of Performability Engineering.* Springer, Berlin-Heidelberg (2008) 321–348.

[2] J.R. Artalejo, Accessible bibliography on retrial queues: progress in 2000–2009. *Math. Comput. Model.* **51** (2010) 1071–1108.

[3] J.R. Artalejo and A. Gómez-Corral, Retrial Queueing Systems: A Computational Approach. Springer, Berlin-Heidelberg (2008).

[4] J.R. Artalejo and V. Pla, On the impact of customer balking, impatience and retrials in telecommunication systems. *Comput. Math. App.* **57** (2009) 217–229.

[5] J.R. Artalejo and M. Pozo, Numerical calculation of the stationary distribution of the main multiserver retrial queue. *Ann. Oper. Res.* **116** (2002) 41–56.

[6] J.R. Artalejo, A. Economou and A. Gómez-Corral, Algorithmic analysis of the Geo/Geo/c retrial. *Eur. J. Oper. Res.* **189** (2008) 1042–1056.

[7] J. Chang and J. Wang, Unreliable *M/M/1/1* retrial queues with set-up time. *Qual. Technol. Quantit. Manage.* **15** (2017) 589–601.

[8] C.J. Chang, J.C. Ke and F.M. Chang, Unreliable retrial queue with loss and feedback under threshold-based policy. *Int. J. Ind. Syst. Eng.* **30** (2018) 1–20.

[9] G. Choudhury and J.C. Ke, An unreliable retrial queue with delaying repair and general retrial times under Bernoulli vacation schedule. *Appl. Math. Comput.* **230** (2014) 436–450.

[10] S. Dudin and O. Dudina, Retrial multi-server queuing system with PHF service time distribution as a model of a channel with unreliable transmission of information. *Appl. Math. Model.* **65** (2019) 676–695.

[11] S. Gao, J. Wang and T.V. Do, Analysis of a discrete-time repairable queue with disasters and working breakdowns. *RAIRO:OR* **53** (2019) 1197–1216.

[12] A. Gómez-Corral, A bibliographical guide to the analysis of retrial queues through matrix analytic techniques. *Ann. Oper. Res.* **141** (2006) 163–191.

[13] M. Jain and R. Gupta, Optimal replacement policy for a repairable system with multiple vacations and imperfect fault coverage. *Comput. Ind. Eng.* **66** (2013) 710–719.

[14] J.C. Ke and T.H. Li, A repairable system with imperfect coverage and reboot. *Appl. Math. Comput.* **246** (2014) 148–158.

[15] J.C. Ke, H.I. Huang and C.H. Lin, A redundant repairable system with imperfect coverage and fuzzy parameters. *Appl. Math. Model.* **32** (2008) 2839–2850.

[16] J.C. Ke, S.L. Lee and Y.L. Hsu, Bayesian analysis for a redundant repairable system with imperfect coverage. *Commun. Stat. Simul. Comput.* **37** (2008) 993–1004.

[17] J.C. Ke, S.L. Lee and M.Y. Ko, Two-unit redundant system with detection delay and imperfect coverage: confidence interval estimation. *Qual. Technol. Quantit. Manage.* **8** (2011) 1–14.

[18] J.C. Ke, T.H. Liu and D.Y. Yang, Modeling of machine interference problem with unreliable repairman and standbys imperfect switchover. *Reliab. Eng. Syst. Saf.* **174** (2018) 12–18.

[19] S.P. Madheswari, B.K. Kumar and P. Suganthi, Analysis of *M/G/1* retrial queues with second optional service and customer balking under two types of Bernoulli vacation schedule. *RAIRO:OR* **53** (2019) 415–443.

[20] E. Morozov and K. Zhukova, A large deviation analysis of retrial models with constant and classic retrial rates. *Perform. Eval.* **135** (2019) 102021.

[21] A.F. Myers, K-out-of-n: G system reliability with imperfect fault coverage. *IEEE Trans. Reliab.* **56** (2007) 464–473.

[22] M.F. Neuts, Matrix Geometric Solutions in Stochastic Models: An Algorithmic Approach. The John Hopkins University Press, Baltimore (1981).

[23] T. Phung-Duc and K. Kawanishi, Performance analysis of call centers with abandonment, retrial and after-call work. *Perform. Eval.* **80** (2014) 43–62.

[24] T. Phung-Duc, H. Masuyama, S. Kasahara and Y. Takahashi, *M/M/3/3* and *M/M/4/4* retrial queues. *J. Ind. Manage. Optim.* **5** (2009) 431–451.

[25] T. Phung-Duc, H. Masuyama, D. Kasahara and Y. Takahashi, State-dependent *M/M/c/c+r* retrial queue with Bernoulli abandonment. *J. Ind. Manage. Optim.* **6** (2010) 517–540.

[26] T. Phung-Duc, H. Masuyama, S. Kasahara and Y. Takahashi, A matrix continued fraction approach to multi-server retrial queues. *Ann. Oper. Res.* **202** (2013) 161–183.

[27] P. Rajadurai, A study on *M/G/1* retrial queueing system with three different types of customers under working vacation policy. *Int. J. Math. Model. Numer. Optim.* **8** (2018) 393–417.

[28] S.S. Sanga and M. Jain, *FM/FM/1* double orbit retrial queue with customers' joining strategy: a parametric nonlinear programming approach. *Appl. Math. Comput.* **362** (2019) 124542.

[29] S.S. Sanga and M. Jain, Cost optimization and ANFIS computing for admission control of *M/M/1/K* queue with general retrial times and discouragement. *Appl. Math. Comput.* **363** (2019) 124624.

[30] K.S. Trivedi, Probability and Statistics with Reliability. Queueing and computer Science Applications, 2nd edition. John Wiley & Sons, New York, NY (2002).

[31] K.H. Wang and Y.J. Chen, Comparative analysis of availability between three systems with general repair times, reboot delay and switching failures. *Appl. Math. Comput.* **215** (2009) 384–394.

[32] J. Wang and F. Zhang, Strategic joining in *M/M/1* retrial queues. *Eur. J. Oper. Res.* **240** (2013) 76–87.

[33] K.H. Wang, T.C. Yen and Y.C. Fang, Comparison of availability between two systems with warm standby units and different imperfect coverage. *Qual. Technol. Quant. Manage.* **9** (2011) 265–282.

[34] K.H. Wang, T.C. Yen and J.J. Jian, Reliability and sensitivity analysis of a repairable system with imperfect coverage under service pressure condition. *J. Manuf. Syst.* **32** (2013) 357–363.

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

[36] D.Y. Yang, F.M. Chang and J.C. Ke, On an unreliable retrial queue with general repeated attempts and J optional vacations. *Appl. Math. Model.* **40** (2016) 3275–3288.

[37] T.C. Yen and K.H. Wang, Cost benefit analysis of three systems with imperfect coverage and standby switching failures. *Int. J. Math. Oper. Res.* **12** (2018) 253–272.