Dynamical modelling and cost optimization of a 5G base station for energy conservation using feedback retrial queue with sleeping strategy

R. Harini[1] · K. Indhira[1]

Abstract

Dense network deployment is now being evaluated as one of the viable solutions to meet the capacity and connectivity requirements of the fifth-generation (5G) cellular system. The goal of 5G cellular networks is to offer clients with faster download speeds, lower latency, more dependability, broader network capacities, more accessibility, and a seamless client experience. However, one of the many obstacles that will need to be overcome in the 5G era is the issue of energy usage. For energy efficiency in 5G cellular networks, researchers have been studying at the sleeping strategy of base stations. In this regard, this study models a 5G BS as an $M^{[X]}/G/1$ feedback retrial queue with a sleeping strategy to reduce average power consumption and conserve power in 5G mobile networks. The probability-generating functions and steady-state probabilities for various BS states were computed employing the supplementary variable approach. In addition, an extensive palette of performance metrics have been determined. Then, with the aid of graphs and tables, the resulting metrics are conceptualized and verified. Further, this research is accelerated in order to bring about the best possible (optimal) cost for the system by adopting a range of optimization approaches namely particle swarm optimization, artificial bee colony and genetic algorithm.

Keywords Energy conservation · 5G Base station · Sleeping strategy · Retrial queues · Cost optimization

Mathematics Subject Classification 68M20 · 90B22 · 60K25 · 60K30

1 Introduction

A cellular network, also known as a mobile network, is a form of wireless communications that operates over discrete geographic areas, or "cells", each of which is connected to the rest of the network by at least one permanently installed transceiver called the base station (BS). Moreover, the advent of 5G networks heralds a new era of connectivity, unprecedented speed, low latency and vast capacity for data transmissions. Further, the evolution and merits of 5G network have been illustrated in Figs. 1 and 2 respectively.

✉ K. Indhira
  kindhira@vit.ac.in

  R. Harini
  harini.r2020@vitstudent.ac.in

[1] Department of Mathematics, School of Advanced Sciences, Vellore Institute of Technology, Vellore, Tamil Nadu 632 014, India

Energy consumption by 5G base stations

As mobile data traffic has skyrocketed over the past decade, BSs have been rapidly deployed to increase cellular system capacity and expand network coverage. Large amounts of power are required to run many BSs. Data servers, BSs, and backhaul devices utilise considerable energy in cellular mobile radio networks. Holtkamp et al. [14] stated that nearly 80% of the energy utilized by a network is used by BSs. The rising demand for high-data-rate services and the rise in the amount of small cell BSs are likely to cause BSs to use even more energy and at higher frequencies, more antennas and more dense layer of small cells are needed, driving up energy costs even further. As a result, most attempts to improve energy consumption in mobile radio networks are focused on BSs.

Fig. 1 Evolution of 5G cellular network

1979	• Using a frequency range of 800 MHz to 900 MHz and a bandwidth of 30 kilohertz, Japan launched its 1G mobile network • With its limited data capabilities, the technology was primarily utilized for voice communication.
1991	• Global System for Mobile Communication (GSM) was launched as part of the 2G mobile communication system with a maximum data rate of 14.4 to 64 kbps, which is sufficient to comply with SMS and email • The first mobile network to offer internet access, albeit at a slower data rate,
Early 2000	• Universal Mobile Terrestrial / Telecommunication Systems (UMTS) emerged as the first step towards 3G mobile communication with higher data rate of 384 kbps. • The availability of video calling on mobile devices boosted its popularity around the world.
2009	• With a transmission throughput of up to 1 Gbps 4G systems are introduced as an upgraded version of 3G networks. • Advanced wireless technology known as Long Term Evolution (LTE) is employed in 4G networks which reduced latency, improved security, and provided greater mobility
Early 2019	• 5G networks were deployed with a notable improvement in the areas of latency, speed, and connectivity. • With significantly better data rates than 4G, 5G allows mobile users to download and upload data much faster.

5G base station	• 5G networks can handle over a thousand more devices per square meter thanks to the **several directional antennas that may be housed in a single 5G BS** which links **various wireless endpoints to a centralized network.** • One way 5G networks will be able to keep up with the demand for high-speed data transmission, data sharing, and constant coverage is by incorporating numerous small cell BSs into macro cell BSs because small cell BSs **have higher throughput and less delay** than traditional ones.
Radio Access Network (RAN)	• RAN **connects the devices** used by clients and other apps in 5G cellular networks. • Since the end-to-end delay (latency) of client request in the RAN and core networks, as well as the link between the RAN and core networks, is high, 5G cellular networks are being used to **lower the latency**.
Low Latency Tactile Internet (LLTI)	• For applications that demand tremendous dependability and responsiveness, like health care, intense gaming, the intelligent transport system (ITS), and education, 5G assures **incredibly dependable and low-latency connection.**
ITS	• LLTI is crucial for **autonomous driving and traffic efficiency**. • For autonomous cars to perform tasks like platooning and overtaking a packet loss rate of 10^{-3} to 10^{-5} and a **latency of 10 ms to 100 ms** are necessary.
Health care	• The most promising usage of LLTI in healthcare are **tele-diagnosis, tele-surgery, and tele-rehabilitation** which enables remote **palpation examinations and robotic surgery**. • Round-trip **latencies of 1-10 ms** are required for such purposes
Intense gaming	• Applications in fields could benefit from the **problem-solving challenges and goal-oriented motivation** featured in these games. • Perceivable human contact with high-quality visualization requires a **round trip time on the order of 1 ms**.
Education	• LLTI facilitates a haptic **overlay between the teacher and the learner,** making it possible to learn and teach remotely. • A round-trip **latency of 5-10 ms** is allowed for perceivable visual, aural, and tactile interaction.

Fig. 2 Merits of 5G cellular network

A brief note on sleeping strategies

A variety of strategies for awakening and falling asleep were proposed. Obviously, it's crucial to examine the various sleep methods employed by BSs with a view towards maximising efficiency and minimising delay. Single sleep, multiple sleeps, an N limited scheme, light sleep, deep sleep, short sleeps, long sleeps, etc. are all common sleeping patterns. In most cases, people's sleeping habits are dictated by the length of time they have to wait for service or by

how much client request (CR) they have accumulated. In order to minimise power consumption, Hawasli and Colak [13] implemented a sleeping strategy in 5G heterogeneous small cell BSs. According to the CRs in the cellular BSs' service region, Yang et al. [36] addressed two distinct sleeps namely, light and deep. Despite this, implementing sleeping methods in both 4G and 5G small cell BSs isn't sufficient to achieve significantly improved energy efficiency. Thus, the energy efficacy of 5G small cell BSs could be optimized by employing the N limited scheme. Pursuant to the N lim-

ited approach, the BS will become active and initiate setup or wake up if a certain threshold of accumulated CRs has been attained. Raising N reduces power usage yet enhances the delay. Zhang et al. [38] proposed that in 5G cellular networks, BSs regularly shift into sleep mode and remain inactive till N CRs have gathered in the queue. Even so, N limited schemes that employ a sleeping technique might trigger more CR delays. Therefore, one could showcase the trade-off between power consumption, power savings, and delay to locate the optimality with an N limited method.

Motivation

Although 5G BSs offer greater dependability, speed, and reduced latency, they also consume four times as much power as 4G BSs. That's why there's reason to be concerned about 5G BS's high power consumption. Specifically, the authors [8, 9, 19] have tackled this problem by simulating the BS as a $M/G/1$ queue using a unique sleeping strategy. Certain circumstances, such as the bulk arrival of CR and retry cases, are not considered, though. With this motivation, the proposed model considers a queueing system with a sleeping strategy and two unique factors: the retrial factor (i.e., if any arriving CRs discover that the BS is already in use, they will be redirected back to wait in a fictitious location known as orbit, and after a few while, they will attempt the service again) and the feedback factor (i.e., if any CRs are not satisfied with the service they received, they may retry repeatedly until the job is completed to their satisfaction). Furthermore, while we have discussed the bulk arrival of CR, authors in the literature to date have solely examined the single arrival of CR. In real life, we frequently encounter situations where service stations may malfunction and require repair, and thus we have accounted for this possibility as well.

Research objective

Our objective in developing the system is to reduce fixed energy usage through the efficient utilisation of BS downtime. To accomplish this, the BS's transceivers are disabled during sleep mode but fully operational during active mode. If implemented correctly, this feature can bring us closer to our eventual goal of "very less energy consumption at zero traffic". Considering this, in the proposed work we find,

- The repercussions of implementing an N limited sleeping strategy on a BS in 5G mobile systems which has been characterized as a feedback retrial queueing system depicted in Figs. 3 and 4.
- The probability-generating functions and steady-state probabilities for various base station states were computed employing the supplementary variable approach.

- The base station's average energy consumption during a certain time period has been estimated.
- A range of optimization approaches, namely PSO, ABC, and GA, have been employed to obtain the best possible (optimal) cost for the system.

Novelty

5G BSs cost around four times as much power as 4G but offer significantly faster speeds, latency, dependability, and data service availability. As a result, 5G BS's excessive need for power is a major cause for alarm. As a result, we characterize the CR queue in a BS as a, $M^{[X]}/G/1$ feedback retrial queueing system with sleeping and waking approaches. Further, the suggested layout incorporates two distinct sleep modes (two distinct vacations) - sleep mode 1 (SM1) and sleep mode 2 (SM2) - with an active state and a set-up stage, respectively, to maximize energy efficiency. The first sleep mode lasts for a shorter amount of time than the second. SM1 has a limit of M short sleeps, followed by SM2 having a lengthy sleep with N policy. While in either of its slumber states, the BS will not activate any CRs. It ought to be noticed that when a service ends, the BS initiates SM1 if there are no CRs waiting in the queue, or else it continues the next service. The system wakes up once per SM1 sleep cycle to count the number of CR waiting to be processed. The BS activates if the number of CR is at least N; else, it enters the second sleep in SM1 and continues sleeping for up to another M times. After M cycles of SM1, the BS enters an SM2 (long sleep) if there are less than N CRs in the queue. When there are N CRs waiting to be dealt with, the SM2 BS wakes up and enters setup mode. Once the initial setup is complete, the server will begin serving the CRs.

1.1 Literature survey

Recently, a study of the underpinnings of 5G cellular networks was proposed by Zheng [37]. This review covered aspects like its advantages, potential threats, bandwidth and spectrum efficiency. López-Pérez et al. [25] highlighted massive multiple-input multiple-output, the lean carrier design, and 5G sleep modes. According to Chih-Lin et al. [5], this estimate of electricity usage was intolerable owing to the associated financial and ecological impacts. Thereby, green communication is feasible in 5G cellular networks through increased energy efficiency. One simple sleeping conduct was offered by Liu et al. [24] to maximize energy efficiency in small cell BSs. Fatima et al. [30] addressed the potential for adapting the various sleep mode enabling approaches to ultra-dense networks and the resulting financial implications as well as the impact on the global carbon footprint. Deepa et al. [9] modelled the BS as a $M/G/1$ queue and covered how a sleeping strategy and N policy might reduce the energy
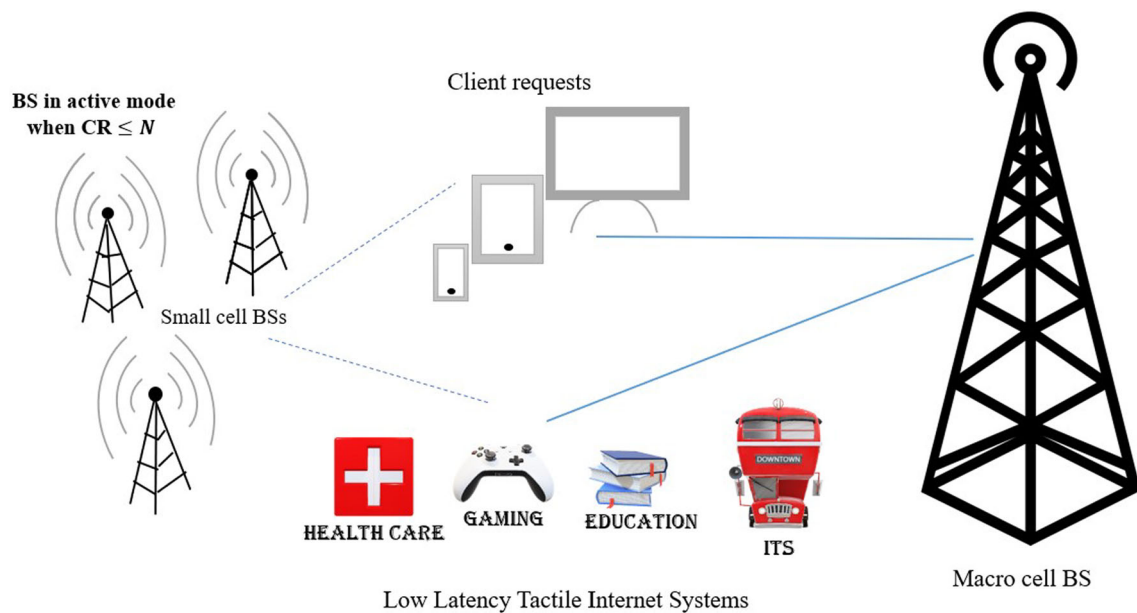
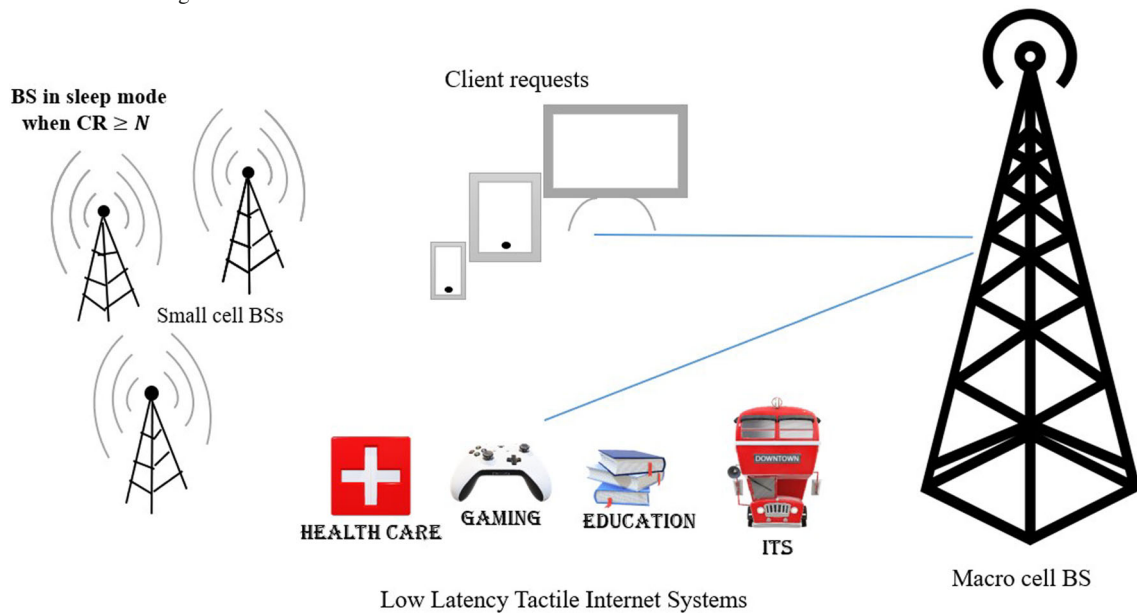**Fig. 3** 5G base station during active mode

Fig. 4 5G base station during sleep mode

consumption of a BS in 4G/5G networks. Further, Kalita and Selvamuthu [19] have analysed a 5G BS by using 2 distinct sleep modes with N policy.

Despite the fact that BS needs some initial setup time by default, this delay was not taken into account in any of the aforementioned studies. Niu et al. [28] modelled the BS as an $M/G/1$ queue with setup and close down times in order to examine the sleep mode operations. Lately, an M/G/1 queue in a 5G BS with N policy in the MS scheme was addressed by authors [6, 7]. Moreover, energy-saving strategies for virtual machines in cloud data centers have been proposed by [27] in order to improve the energy-saving capability of cloud systems.

Research in a queue and retrial queue had been conducted by several authors such as Falin and Templeton [10], considering a variety of combinations of shut down time, setup time, feedback, repair, and N policy. Consumer's impatience with an unreliable bulk queuing system under vacation, setup time close down time with N policy was analysed recently by Ayyappan and Nirmala [2, 3]. A batch arrival sole server retrial queue with modified vacations under N policy was scrutinized by Haridass and Arumuganathan [12]. Moreover, experts like Jain and Upadhyaya [16], Jain and Kaur [17], Jain and Kumar [18], Jain and Bhargava [15] and others have conducted substantial investigations on the feedback and repair situation as well. In this article, the CR retrial queue in a 5G

BS with feedback CR, setup time, repair time, and N policy in the MS scheme has been investigated by implementing the supplementary variable technique.

And here's how the balance of the paper pans out: The proposed 5G BS is described in detail in Sect. 2. The prerequisites for stability, steady state outcomes, and accompanying equations are established in Sect. 3. In Sect. 4, we examine some primary system metrics. In Sects. 5 and 6, we discussed a few cases of exceptional significance and used statistical analysis to understand the interplay between the system's underlying structure and its many components respectively. An evaluation of the suggested model's energy efficiency is provided in Sect. 7. The analysis of cost optimization using three different methods is addressed in Sect. 8. Finally, Sect. 9 concludes with an outline of the findings and future scope.

2 Description of the model

In a cellular network, a signal (CR) from a mobile phone is picked up by a BS antenna and sent via transceivers to the destination. A stochastic model of the 5G BS's sleeping mechanism has been built so that its energy consumption may be analysed. The framework under discussion is illustrated and depicted pictorially in Fig. 5 as follows:

Arrival process: CR streaming in batches from the outside abide a Poisson process, with an arrival rate ν. Presuming that "$\mathcal{U}_k, k = 1, 2, 3, \dots$" follows a standard dist., let $\mathcal{U}_k$ be the no. of CRs affiliated with the kth batch arrival. "$Pr[\mathcal{U}_k = n] = \mathcal{A}_n, n = 1, 2, 3, \dots$" and $\mathcal{U}(\check{\varepsilon})$ indicate the PGF of $\mathcal{U}$. Further, the first and second factorial moments of r.v. $\mathcal{U}$ are denoted by $\mathcal{E}(X_1)$ and $\mathcal{E}(X_2)$ respectively.

Retrial process: Any incoming CRs who encounter a busy, on sleep modes, on setup state, or down BS are assumed to be promptly booted and thrown into a group of blocked users known as an orbit. We presume that the CRs' retrial times distributed equally around the orbit with a random dist. $\mathcal{Q}(\hat{\varpi})$ with a corresponding Laplace Stieltjes Transform (LST) $\mathcal{Q}^*(t)$.

Service process: As soon as the CR reaches the BS and begins to be transmitted, service has been established. If the group of incoming CRs stumbles upon an unoccupied BS, one of them will be given approval to start the service, and the others will enter the orbit. Each incoming CR is processed with first come, first served (FCFS) basis. It is assumed that service time S is assumed to have dist. function $\mathcal{L}(\hat{\varpi})$ with a corresponding LST $\mathcal{L}^*(t)$ with finite k^{th} moment $l^k, k = 1, 2, \dots$.

Vacation process: After a CR's service completion epoch, BSs will enter SM1 with prob. p_0 for a brief length of time (S1) if there is no new CR been assigned. In this mode, the BSs can sleep for up to M times. After first sleep in SM1, if CR is less than N, then the BS will undergo M sleeps with prob. p_m. Prob. dist. function $\mathcal{G}_1(\hat{\varpi})$ with a corresponding LST $\mathcal{G}_1^*(t)$ with finite k^{th} moment $g_1^k, k = 1, 2, \dots$ characterize the distribution of S1's duration, which is i.i.d. r.v. After M sleeps through SM1, if no CRs are available, the BSs will transition into SM2 with prob. s for an extended period of time of random duration S2 or when N or more CRs have been accumulated at the conclusion of any sleeping phase, the BS reactivates with prob. 1 - s and begins its setup procedure, and S2 is an r.v with PDF $\mathcal{G}_2^*(t)$ and a corresponding LST $\mathcal{G}_2^*(t)$ with finite k^{th} moment $g_2^k, k = 1, 2, \dots$
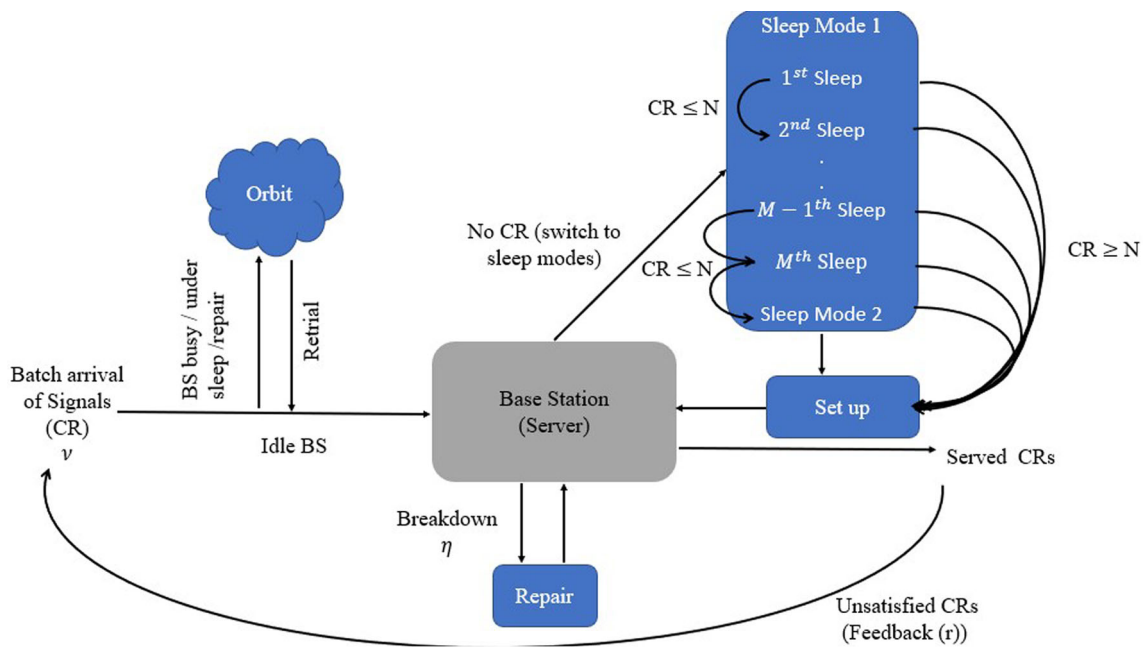

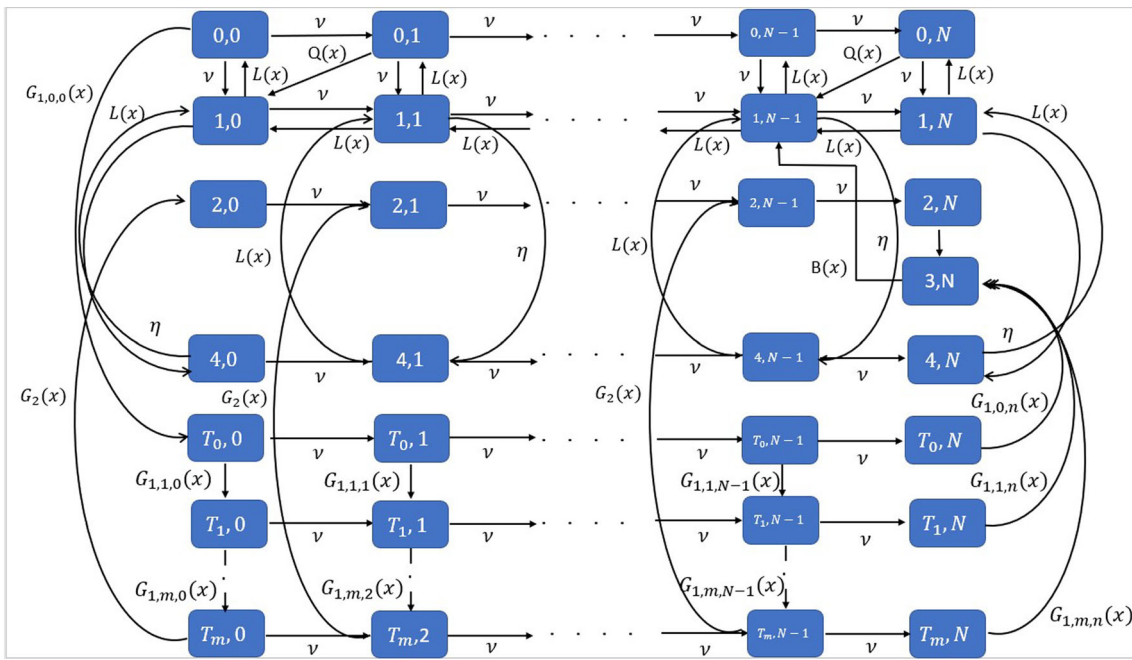**Fig. 5** A pictorial depiction of the model

Fig. 6 State Transition Diagram for the proposed 5G BS system

Setup process: The BS begins its setup process if it detects at least 'N' CRs waiting for service while it is in either of its two sleep states. After initialization is complete, the BS begins to process CRs. There is an implicit assumption that setup time has the dist. function $\mathcal{B}(\hat{\varpi})$ and a corresponding LST $\mathcal{B}^*(t)$.

Feedback rule: Once CR's service is cut off, discontented CR have a probability of $r(0 \leq r \leq 1)$ that they will return to the orbit as a feedback CR in order to undergo another regular service, or a prob. of $1 - r$ that they decide to exit the system outright.

Breakdown process: While the BS is functioning, there is always a chance that it will malfunction, leading to a temporary loss of service (a server outage). The failures, or BS lifetimes, are caused by exogenous Poisson processes with rates η, which can be viewed as a catastrophic event.

Repair process: In the event of a BS failure, it is immediately dispatched for repair, at which point it stops serving CR altogether until the corresponding service channel is restored to working order. A CR who was being served when the BS went down must now wait for the rest of their order to be processed. The repair time of the BS is presumed to be arbitrarily distributed with dist. function $\mathcal{S}(\grave{\varsigma})$ and LST $\mathcal{S}^*(t)$ with finite k^{th} moment s^k.

All stochastic processes in the system are considered to be unrelated of one another. In addition, the state transition diagram for the proposed 5G BS system is presented in Fig. 6.

3 Steady state probability analysis

When formulating the SS diff. eqns. for the RQ system in this section, we take the elapsed retrial times, elapsed service times, elapsed vacation time, and elapsed repair times as supplementary variables.

3.1 Steady-state equations

In SS, we take into consideration that $\mathcal{Q}(0) = 0$, $\mathcal{Q}(\infty) = 1$, $\mathcal{L}(0) = 0$, $\mathcal{L}(\infty) = 1$, $\mathcal{G}_{1,m}(0) = 0$, $\mathcal{G}_{1,m}(\infty) = 1$, $\mathcal{G}_2(0) = 0$, $\mathcal{G}_2(\infty) = 1$, $\mathcal{B}(0) = 0$, $\mathcal{B}(\infty) = 1$ are continuous at $\hat{\varpi} = 0$ and $\mathcal{S}(0) = 0$, $\mathcal{S}(\infty) = 1$ is continuous at $\grave{\varsigma} = 0$. Hence, the functions $\vartheta(\hat{\varpi}), \theta(\hat{\varpi}), \upsilon_{1,m}(\hat{\varpi}), \upsilon_2(\hat{\varpi})$, $\varphi(\hat{\varpi})$ and $\zeta(\grave{\varsigma})$ are the "conditional completion rates (hazard rates)" for retrial, service, 1st and 2nd sleep modes, setup and repair, respectively.

$$\vartheta(\hat{\varpi})d\hat{\varpi} = \frac{d\mathcal{Q}(\hat{\varpi})}{1 - \mathcal{Q}(\hat{\varpi})}; \quad \theta(\hat{\varpi})d\hat{\varpi} = \frac{d\mathcal{L}(\hat{\varpi})}{1 - \mathcal{L}(\hat{\varpi})};$$

$$\upsilon_{1,m}(\hat{\varpi})d\hat{\varpi} = \frac{d\mathcal{G}_{1,m}(\hat{\varpi})}{1 - \mathcal{G}_{1,m}(\hat{\varpi})} (m = 1, 2, \ldots M);$$

$$\upsilon_2(\hat{\varpi})d\hat{\varpi} = \frac{d\mathcal{G}_2(\hat{\varpi})}{1 - \mathcal{G}_2(\hat{\varpi})};$$

$$\varphi(\hat{\varpi})d\hat{\varpi} = \frac{d\mathcal{B}(\hat{\varpi})}{1 - \mathcal{B}(\hat{\varpi})};$$

$$\zeta(\grave{\varsigma})d\grave{\varsigma} = \frac{d\mathcal{S}(\grave{\varsigma})}{1 - \mathcal{S}(\grave{\varsigma})};$$

At time $\widetilde{\beta}$, the elapsed retrial times, service times, vacation times, and repair times are denotes as $\mathcal{Q}^0(\widetilde{\beta}), \mathcal{L}^0(\widetilde{\beta}), \mathcal{G}_{1,m}^0(\widetilde{\beta})$,

$\mathcal{G}_2^0(\widetilde{\beta})$, $\mathcal{B}^0(\widetilde{\beta})$ and $\mathcal{S}^0(\widetilde{\beta})$ respectively. Additionally, construct the random variable,

$$
\Xi(\widetilde{\beta}) = \begin{cases}
0, & \text{in the situation when BS is under inactive} \\
 & \quad \text{mode at time } \widetilde{\beta} \\
1, & \text{in the situation when BS is in active mode} \\
 & \quad \text{at time } \widetilde{\beta} \\
2, & \text{in the situation when BS is under sleep} \\
 & \quad \text{mode 2 at time } \widetilde{\beta} \\
3, & \text{in the situation when BS is in set up at} \\
 & \quad \text{time } \widetilde{\beta} \\
4, & \text{in the situation when BS is under repair} \\
 & \quad \text{at time } \widetilde{\beta} \\
T_0, & \text{in the situation when BS is taking 1st} \\
 & \quad \text{sleep under sleep mode 1} \\
 & \quad \text{at time } \widetilde{\beta} \\
T_1, & \text{in the situation when BS is taking 2nd} \\
 & \quad \text{sleep under sleep mode 1} \\
 & \quad \text{at time } \widetilde{\beta} \\
\vdots & \\
T_m, & \text{in the situation when BS is taking mth} \\
 & \quad \text{sleep under sleep mode 1} \\
 & \quad \text{at tim e } \widetilde{\beta}
\end{cases}
$$

We additionally highlight how the bivariate Markov process $\{\Xi(\widetilde{\beta}),\ \Theta(\widetilde{\beta});\ \widetilde{\beta} \geq 0\ \}$ can be applied to characterize the state of the system at a specific instant of time $\widetilde{\beta}$, wherein $\Xi(\widetilde{\beta})$ symbolises the BS state $(0, 1, 2, 3, 4, T_0, T_1, \ldots T_m)$ contingent upon whether the BS is inactive, engaged, under sleeping modes, setup or under repair. At time $\widetilde{\beta}$, the no. of CRs in the orbit is symbolised by $\Theta(\widetilde{\beta})$.

Further, the subsequent notations and probabilities were defined:

$\mathcal{P}_0(\widetilde{\beta})$ Probability that the BS is in its usual busy phase and the BS is inactive at time $\widetilde{\beta}$

$\Psi_n(\hat{\varpi}, \widetilde{\beta})$ Prob. of having precisely n CRs in the orbit at time $\widetilde{\beta}$ with the elapsed retrial time of the trail CR undergoing retrial is $\hat{\varpi}$

$\Gamma_n(\hat{\varpi}, \widetilde{\beta})$ Prob. of having approximately n CRs in the orbit at time $\widetilde{\beta}$ and BS is engaged with batch of CR with the elapsed service time of CR undergoing service is $\hat{\varpi}$

$\Pi_{1,n}(\hat{\varpi}, \widetilde{\beta})$ Prob. of having exactly n CRs in the orbit that at time $\widetilde{\beta}$ in which the BS is under SM1 with the elapsed vacation time is $\hat{\varpi}$

$\Pi_{2,n}(\hat{\varpi}, \widetilde{\beta})$ Prob. of having exactly n CRs in the orbit that at time $\widetilde{\beta}$ where the BS is under SM2 with the elapsed vacation time is $\hat{\varpi}$

$\Omega_n(\hat{\varpi}, \widetilde{\beta})$ Prob. of having precisely n CRs in the orbit at time $\widetilde{\beta}$ and BS is under setup with the elapsed setup time of CR experiencing service is $\hat{\varpi}$

$\mathcal{R}_n(\hat{\varpi}, \varsigma, \widetilde{\beta})$ Prob. of having identically n CR in the orbit at time $\widetilde{\beta}$ with elapsed essential service time of the batch of CRs undergoing service $\hat{\varpi}$ and the elapsed repair time of BS for repair ς

Let the collection of epochs whereby a regular service period, vacation period, or repair period terminates be $\{\widetilde{\beta}_n;\ n = 1, 2, \ldots\}$. Thus, using a sequence of random vectors $F_n = \{\Xi(\widetilde{\beta}_n+), \Theta(\widetilde{\beta}_n+)\}$, an MC is created and incorporated in the retry QS. Our system must thus be stable according to Appendix A, which states that $F_n;\ n \in N$ is ergodic iff $\rho < 1$, The probs. $\mathcal{P}_0(\widetilde{\beta}) = P\{\Xi(\widetilde{\beta}) = 1, \Theta(\widetilde{\beta}) = 0\}$; and prob. densities for the method $\{\Theta(\widetilde{\beta}), \widetilde{\beta} \geq 0\}$, are stated below,

$$
\begin{aligned}
\Psi_n(\hat{\varpi}, \widetilde{\beta})d\hat{\varpi} &= P\{\Theta(\widetilde{\beta}) = 0, \Xi(\widetilde{\beta}) = n, \hat{\varpi} \\
&\leq \mathcal{Q}^0(\widetilde{\beta}) < \hat{\varpi} + d\hat{\varpi}\}, \text{for } \widetilde{\beta}, \hat{\varpi} \geq 0,\ n \geq 1 \\
\Gamma_n(\hat{\varpi}, \widetilde{\beta})d\hat{\varpi} &= P\{\Theta(\widetilde{\beta}) = 1, \Xi(\widetilde{\beta}) = n, \hat{\varpi} \\
&\leq \mathcal{L}^0(\widetilde{\beta}) < \hat{\varpi} + d\hat{\varpi}\}, \text{for } \widetilde{\beta}, \hat{\varpi} \geq 0,\ n \geq 0 \\
\Pi_{1,m,n}(\hat{\varpi}, \widetilde{\beta})d\hat{\varpi} &= P\{\Theta(\widetilde{\beta}) = T_m, \Xi(\widetilde{\beta}) = n, \hat{\varpi} \\
&\leq \mathcal{G}_{1,m}^0(\widetilde{\beta}) < \hat{\varpi} + d\hat{\varpi}\}, \text{for } \widetilde{\beta}, \hat{\varpi} \geq 0,\ n \geq 0 \\
\Pi_{2,n}(\hat{\varpi}, \widetilde{\beta})d\hat{\varpi} &= P\{\Theta(\widetilde{\beta}) = 2, \Xi(\widetilde{\beta}) = n, \hat{\varpi} \\
&\leq \mathcal{G}_2^0(\widetilde{\beta}) < \hat{\varpi} + d\hat{\varpi}\}, \text{for } \widetilde{\beta}, \hat{\varpi} \geq 0,\ n \geq 0 \\
\Omega_n(\hat{\varpi}, \widetilde{\beta})d\hat{\varpi} &= P\{\Theta(\widetilde{\beta}) = 3, \Xi(\widetilde{\beta}) = n, \hat{\varpi} \\
&\leq \mathcal{B}^0(\widetilde{\beta}) < \hat{\varpi} + d\hat{\varpi}\}, \text{for } \widetilde{\beta}, \hat{\varpi} \geq 0,\ n \geq 0 \\
\mathcal{R}_n(\hat{\varpi}, \varsigma, \widetilde{\beta})d\hat{\varpi} &= P\{\Theta(\widetilde{\beta}) = 4, \Xi(\widetilde{\beta}) = n, \varsigma \\
&\leq \mathcal{S}^0(\widetilde{\beta}) < \varsigma + d\varsigma/\mathcal{L}^0(\widetilde{\beta}) = \hat{\varpi}\}, \\
&\quad \text{for } \widetilde{\beta} \geq 0, (\hat{\varpi}, \varsigma) \geq 0, n \geq 0
\end{aligned}
$$

We presume that the sequel complies with the stability requirements, thus we can assign limiting probs. for $\hat{\varpi} > 0$, $n \geq 0$ and $(i = 1, 2)$

$$
\begin{aligned}
\mathcal{P}_0 &= \lim_{\widetilde{\beta}\to\infty}\mathcal{P}_0(\widetilde{\beta});\ \Psi_n(\hat{\varpi}) \\
&= \lim_{\widetilde{\beta}\to\infty}\Psi_n(\hat{\varpi}, \widetilde{\beta});\ \Gamma_n(\hat{\varpi}) \\
&= \lim_{\widetilde{\beta}\to\infty}\Gamma_n(\hat{\varpi}, \widetilde{\beta}); \\
\Pi_{1,m,n}(\hat{\varpi}) &= \lim_{\widetilde{\beta}\to\infty}\Pi_{1,m,n}(\hat{\varpi}, \widetilde{\beta});\ \Pi_{2,n}(\hat{\varpi}) \\
&= \lim_{\widetilde{\beta}\to\infty}\Pi_{2,n}(\hat{\varpi}, \widetilde{\beta}) \\
\Omega_n(\hat{\varpi}) &= \lim_{\widetilde{\beta}\to\infty}\Omega_n(\hat{\varpi}, \widetilde{\beta});\ \text{and} \mathcal{R}_n(\hat{\varpi}, \varsigma) \\
&= \lim_{\widetilde{\beta}\to\infty}\mathcal{R}_n(\hat{\varpi}, \varsigma, \widetilde{\beta})
\end{aligned}
$$

The eqns. below, which govern the dynamics of the system's behaviour, are generated by using a supplementary variable

approach.

$$\nu \mathcal{P}_0 = p_0 \int_0^\infty \Pi_{1,0,0}(\hat{\varpi}) \upsilon_{1,0}(\hat{\varpi}) d\hat{\varpi}$$

$$+ \int_0^\infty \Pi_{1,1,0}(\hat{\varpi}) \upsilon_{1,1}(\hat{\varpi}) d\hat{\varpi} + \ldots$$

$$+ \int_0^\infty \Pi_{1,M,0}(\hat{\varpi}) \upsilon_{1,M}(\hat{\varpi}) d\hat{\varpi} \tag{1}$$

$$\frac{d\Psi_n(\hat{\varpi})}{d\hat{\varpi}} + (\nu + \vartheta(\hat{\varpi}))\Psi_n(\hat{\varpi}) = 0; \, n \geq 1 \tag{2}$$

$$\frac{d\Gamma_n(\hat{\varpi})}{d\hat{\varpi}} + (\nu + \eta + \theta(\hat{\varpi}))\Gamma_n(\hat{\varpi}) = 0; \, n < N \tag{3}$$

$$\frac{d\Gamma_n(\hat{\varpi})}{d\hat{\varpi}} + (\nu + \eta + \theta(\hat{\varpi}))\Gamma_n(\hat{\varpi})$$

$$= \nu \sum_{k=1}^n a_k \Gamma_{n-k}(\hat{\varpi}); \, n \geq N \tag{4}$$

$$\frac{d\Pi_{1,m,0}(\hat{\varpi})}{d\hat{\varpi}} + (\nu + \upsilon_{1,m}(\hat{\varpi}))\Pi_{1,m,0}(\hat{\varpi}) = 0; \tag{5}$$

$$\frac{d\Pi_{1,m,n}(\hat{\varpi})}{d\hat{\varpi}} + (\nu + \upsilon_{1,m}(\hat{\varpi}))\Pi_{1,m,n}(\hat{\varpi})$$

$$= \nu \sum_{k=1}^n a_k \Pi_{1,m,n-k}(\hat{\varpi});$$

$$(m = 1, 2, \ldots M); \, n \leq N - 1 \tag{6}$$

$$\frac{d\Pi_{2,n}(\hat{\varpi})}{d\hat{\varpi}} + (\nu + \upsilon_2(\hat{\varpi}))\Pi_{2,n}(\hat{\varpi})$$

$$= \nu \sum_{k=1}^n a_k \Pi_{2,n-k}(\hat{\varpi}); \, n \leq N - 1 \tag{7}$$

$$\frac{d\Omega_n(\hat{\varpi})}{d\hat{\varpi}} + (\nu + \varphi(\hat{\varpi}))\Gamma_n(\hat{\varpi}) = 0; n < N; \tag{8}$$

$$\frac{d\Omega_n(\hat{\varpi})}{d\hat{\varpi}} + (\nu + \varphi(\hat{\varpi}))\Gamma_n(\hat{\varpi})$$

$$= \sum_{k=1}^n a_k \Omega_{n-k}(\hat{\varpi}); n \geq N; \tag{9}$$

$$\frac{d\mathcal{R}_n(\hat{\varpi}, \zeta)}{d\zeta} + (\kappa + \zeta(\zeta))\mathcal{R}_n(\hat{\varpi}, \zeta)$$

$$= \kappa \sum_{k=1}^n a_k \mathcal{R}_{n-k}(\hat{\varpi}, \zeta); n \geq 0 \tag{10}$$

At $\hat{\varpi} = 0$, $\zeta = 0$ the SS boundary conditions are as follows:

$$\Psi_n(0) = \bar{r} \int_0^\infty \Gamma_n(\hat{\varpi}) \theta(\hat{\varpi}) d\hat{\varpi}$$

$$+ r \int_0^\infty \Gamma_{n-1}(\hat{\varpi}) \theta(\hat{\varpi}) d\hat{\varpi}$$

$$+ p_0 \int_0^\infty \Pi_{1,0,n}(\hat{\varpi}) \upsilon_{1,0}(\hat{\varpi}) d\hat{\varpi}$$

$$+ \int_0^\infty \Pi_{1,1,n}(\hat{\varpi}) \upsilon_{1,1}(\hat{\varpi}) d\hat{\varpi}$$

$$+ \cdots + \Pi_{1,M,n}(\hat{\varpi}) \upsilon_{1,M}(\hat{\varpi}) d\hat{\varpi}; n \geq 1 \tag{11}$$

$$\Gamma_1(0) = \int_0^\infty \Psi_1(\hat{\varpi}) \vartheta(\hat{\varpi}) d\hat{\varpi} + \nu a_1 \mathcal{P}_0; \, n = 0 \tag{12}$$

$$\Gamma_n(0) = \int_0^\infty \Psi_{n+1}(\hat{\varpi}) \vartheta(\hat{\varpi}) d\hat{\varpi} + \nu a_{n+1} \mathcal{P}_0$$

$$+ \nu \sum_{k=1}^n a_k \int_0^\infty \Psi_{n-k+1}(\hat{\varpi}) d\hat{\varpi}; \, n < N \tag{13}$$

$$\Gamma_n(0) = \int_0^\infty \Psi_{n+1}(\hat{\varpi}) \vartheta(\hat{\varpi}) d\hat{\varpi}$$

$$+ \nu a_{n+1} \mathcal{P}_0 + \nu \sum_{k=1}^n a_k \int_0^\infty \Psi_{n-k+1}(\hat{\varpi}) d\hat{\varpi}$$

$$+ \int_0^\infty \Omega_n(\hat{\varpi}) \varphi(\hat{\varpi}) d\hat{\varpi}; \, n \geq N \tag{14}$$

$$\Pi_{1,0,0}(0) = \bar{r} \int_0^\infty \Gamma_n(\hat{\varpi}) \theta(\hat{\varpi}) d\hat{\varpi}; \, n = 0 \tag{15}$$

$$\Pi_{1,m,0}(0) = \begin{cases} p_m \int_0^\infty \Pi_{1,0,n}(\hat{\varpi}) \upsilon_{1,0}(\hat{\varpi}) d\hat{\varpi}; \\ \quad (m = 1, 2, \ldots M); \, n \leq N - 1 \\ 0; \, n \geq N \end{cases} \tag{16}$$

$$\Pi_{2,n}(0) = s \left[p_0 \int_0^\infty \Pi_{1,0,n}(\hat{\varpi}) \upsilon_{1,0}(\hat{\varpi}) d\hat{\varpi} \right.$$

$$+ \int_0^\infty \Pi_{1,1,n}(\hat{\varpi}) \upsilon_{1,1}(\hat{\varpi}) d\hat{\varpi} + \ldots$$

$$\left. + \Pi_{1,M,n}(\hat{\varpi}) \upsilon_{1,M}(\hat{\varpi}) d\hat{\varpi} \right]; n \leq N - 1 \tag{17}$$

$$\Omega_n(0) = p_0 \int_0^\infty \Pi_{1,0,n}(\hat{\varpi}) \upsilon_{1,0}(\hat{\varpi}) d\hat{\varpi}$$

$$+ \int_0^\infty \Pi_{1,1,n}(\hat{\varpi}) \upsilon_{1,1}(\hat{\varpi}) d\hat{\varpi} + \ldots$$

$$+ \Pi_{1,M,n}(\hat{\varpi}) \upsilon_{1,M}(\hat{\varpi}) d\hat{\varpi}$$

$$+ \int_0^\infty \Pi_{2,n}(\hat{\varpi}) \upsilon_2(\hat{\varpi}) d\hat{\varpi}; \, n \geq N; \tag{18}$$

$$\mathcal{R}_n(\hat{\varpi}, 0) = \eta \Gamma_n(\hat{\varpi}); n \geq 0 \tag{19}$$

The normalizing condition is

$$\mathcal{P}_0 + \sum_{n=1}^\infty \int_0^\infty \Psi_n(\hat{\varpi}) d\hat{\varpi} + \sum_{n=0}^\infty \left[\sum_{m=1}^M \int_0^\infty \Pi_{1,m,n}(\hat{\varpi}) d\hat{\varpi} \right.$$

$$+ \int_0^\infty \Pi_{2,n}(\hat{\varpi}) d\hat{\varpi}$$

$$+ \int_0^\infty \Gamma_n(\hat{\varpi}) d\hat{\varpi} + \int_0^\infty \Omega_n(\hat{\varpi}) d\hat{\varpi}$$

$$\left. + \int_0^\infty \int_0^\infty \mathcal{R}_n(\hat{\varpi}, \zeta) d\zeta d\hat{\varpi} \right] = 1$$

3.2 The steady state solution

The retrial queueing framework's SS solution is generated with the aid of generating function strategy. In addition, to solve the eqns. above the GFs for $|\breve{\varepsilon}| < 1$ are formulated as below:

$$\Psi(\hat{\varpi}, \breve{\varepsilon}) = \sum_{n=1}^{\infty} \Psi_n(\hat{\varpi})\breve{\varepsilon}^n; \ \Psi(0, \breve{\varepsilon}) = \sum_{n=1}^{\infty} \Psi_n(0)\breve{\varepsilon}^n;$$

$$\Gamma(\hat{\varpi}, \breve{\varepsilon}) = \sum_{n=0}^{\infty} \Gamma_n(\hat{\varpi})\breve{\varepsilon}^n; \ \Gamma(0, \breve{\varepsilon}) = \sum_{n=0}^{\infty} \Gamma_n(0)\breve{\varepsilon}^n$$

$$\Pi_{1,m}(\hat{\varpi}, \breve{\varepsilon}) = \sum_{n=0}^{\infty} \Pi_{1,m,n}(\hat{\varpi})\breve{\varepsilon}^n; \ \Pi_{1,m}(0, \breve{\varepsilon}) = \sum_{n=0}^{\infty} \Pi_{1,m,n}(0)\breve{\varepsilon}^n;$$

$$\Pi_2(\hat{\varpi}, \breve{\varepsilon}) = \sum_{n=0}^{\infty} \Pi_{2,n}(\hat{\varpi})\breve{\varepsilon}^n; \ \Pi_2(0, \breve{\varepsilon}) = \sum_{n=0}^{\infty} \Pi_{2,n}(0)\breve{\varepsilon}^n;$$

$$\Omega(\hat{\varpi}, \breve{\varepsilon}) = \sum_{n=0}^{\infty} \Omega_n(\hat{\varpi})\breve{\varepsilon}^n; \ \Omega(0, \breve{\varepsilon}) = \sum_{n=0}^{\infty} \Omega_n(0)\breve{\varepsilon}^n$$

$$\mathcal{R}(\hat{\varpi}, \varsigma, \breve{\varepsilon}) = \sum_{n=0}^{\infty} \mathcal{R}_n(\hat{\varpi}, \varsigma)\breve{\varepsilon}^n; \ \mathcal{R}(\hat{\varpi}, 0, \breve{\varepsilon}) = \sum_{n=0}^{\infty} \mathcal{R}_n(\hat{\varpi}, 0)\breve{\varepsilon}^n;$$

$$a(\breve{\varepsilon}) = \sum_{n=1}^{\infty} \mathcal{A}_n\breve{\varepsilon}^n$$

From (2) to (19), multiply the SS eqns. and boundary conditions by $\breve{\varepsilon}^n$ and summing on n, and $(n = 0, 1, 2, \dots)$.

$$\frac{\partial}{\partial \hat{\varpi}}\Psi(\hat{\varpi}, \breve{\varepsilon}) + (\nu + \vartheta(\hat{\varpi}))\Psi(\hat{\varpi}, \breve{\varepsilon}) = 0 \tag{20}$$

$$\frac{\partial}{\partial \hat{\varpi}}\Gamma(\hat{\varpi}, \breve{\varepsilon}) + [\nu(1 - a(\breve{\varepsilon})) + \eta + \theta(\hat{\varpi})]\Gamma(\hat{\varpi}, \breve{\varepsilon})$$
$$= \int_0^{\infty} \mathcal{R}(\hat{\varpi}, \varsigma, \breve{\varepsilon})\zeta(\varsigma)d\varsigma \tag{21}$$

$$\frac{\partial}{\partial \hat{\varpi}}\Pi_{1,m}(\hat{\varpi}, \breve{\varepsilon}) + [\nu(1 - a(\breve{\varepsilon}))$$
$$+ \upsilon_{1,m}(\hat{\varpi})]\Pi_{1,m}(\hat{\varpi}, \breve{\varepsilon}) = 0; \ (m = 1, 2, \dots M) \tag{22}$$

$$\frac{\partial}{\partial \hat{\varpi}}\Pi_2(\hat{\varpi}, \breve{\varepsilon}) + [\nu(1 - a(\breve{\varepsilon})) + \upsilon_2(\hat{\varpi})]\Pi_2(\hat{\varpi}, \breve{\varepsilon}) = 0 \tag{23}$$

$$\frac{\partial}{\partial \hat{\varpi}}\Omega(\hat{\varpi}, \breve{\varepsilon}) + [\nu(1 - a(\breve{\varepsilon})) + \varphi(\hat{\varpi})]\Omega(\hat{\varpi}, \breve{\varepsilon}) = 0 \tag{24}$$

$$\frac{\partial}{\partial \hat{\varpi}}\mathcal{R}(\hat{\varpi}, \breve{\varepsilon}) + [\nu(1 - a(\breve{\varepsilon})) + \varphi(\hat{\varpi})]\mathcal{R}(\hat{\varpi}, \breve{\varepsilon}) = 0 \tag{25}$$

$$\Psi(0, \breve{\varepsilon}) = (1 - r + r\breve{\varepsilon})\int_0^{\infty} \Gamma(\hat{\varpi}, \breve{\varepsilon})\theta(\hat{\varpi})d\hat{\varpi}$$
$$+ p_0 \int_0^{\infty} \Pi_{1,0}(\hat{\varpi}, \breve{\varepsilon})\upsilon_{1,0}((\hat{\varpi})d\hat{\varpi}$$
$$+ \sum_{m=1}^{M} \int_0^{\infty} \Pi_{1,m}(\hat{\varpi}, \breve{\varepsilon})\upsilon_{1,m}(\hat{\varpi})d\hat{\varpi} - \nu\mathcal{P}_0 - \Pi_{1,0}(0, \breve{\varepsilon}) \tag{26}$$

$$\Gamma(0, \breve{\varepsilon}) = \frac{1}{\breve{\varepsilon}}\int_0^{\infty} \Psi(\hat{\varpi}, \breve{\varepsilon})\vartheta(\hat{\varpi})d\hat{\varpi}$$
$$+ \frac{\nu}{\breve{\varepsilon}}\mathcal{A}(\breve{\varepsilon})\int_0^{\infty} \Psi(\hat{\varpi}, \breve{\varepsilon})d\hat{\varpi}$$
$$+ \int_0^{\infty} \Omega(\hat{\varpi}, \breve{\varepsilon})\varphi(\hat{\varpi})d\hat{\varpi}$$
$$+ \mathcal{A}(\breve{\varepsilon})\frac{\nu\mathcal{P}_0}{\breve{\varepsilon}} - \int_0^{\infty} \Psi_1(\hat{\varpi})\vartheta(\hat{\varpi})d\hat{\varpi} - \nu a_1\mathcal{P}_0 \tag{27}$$

$$\Pi_m(0, \breve{\varepsilon}) = p_m \int_0^{\infty} \Pi_{1,0}(\hat{\varpi}, \breve{\psi})\upsilon_{1,0}(\hat{\varpi})d\hat{\varpi};$$
$$(m = 1, 2, \dots M) \tag{28}$$

$$\Pi_2(0, \breve{\varepsilon}) = s\left[p_0 \int_0^{\infty} \Pi_{1,0}(\hat{\varpi}, \breve{\varepsilon})\upsilon_{1,0}((\hat{\varpi})d\hat{\varpi} \right.$$
$$\left. + \sum_{m=1}^{M} \int_0^{\infty} \Pi_{1,m}(\hat{\varpi}, \breve{\varepsilon})\upsilon_{1,m}(\hat{\varpi})d\hat{\varpi} \right] \tag{29}$$

$$\Omega(0, \breve{\varepsilon}) = p_0 \int_0^{\infty} \Pi_{1,0}(\hat{\varpi}, \breve{\varepsilon})\upsilon_{1,0}((\hat{\varpi})d\hat{\varpi}$$
$$+ \sum_{m=1}^{M} \int_0^{\infty} \Pi_{1,m}(\hat{\varpi}, \breve{\varepsilon})\upsilon_{1,m}(\hat{\varpi})d\hat{\varpi}$$
$$+ \int_0^{\infty} \Pi_2(\hat{\varpi}, \breve{\varepsilon})\upsilon_2(\hat{\varpi})d\hat{\varpi} \tag{30}$$

$$\mathcal{R}(0, \breve{\varepsilon}) = \eta\Gamma(\hat{\varpi}, \breve{\varepsilon}) \tag{31}$$

From (5), by setting $m = 0$, we have,

$$\Pi_{1,0,0}(\hat{\varpi}) = \Pi_{1,0,0}(0)e^{-\nu\hat{\varpi}}[1 - \mathcal{G}_{1,0}(\hat{\varpi})] \tag{32}$$

When we multiply the aforementioned eqns. by $\upsilon_{1,0}(\hat{\varpi})$ on both sides and integrating with regard to $\hat{\varpi}$ from 0 to ∞ and further, by utilizing (1) we get,

$$\Pi_{1,0,0}(0) = \Pi_{1,0}(0, \breve{\psi}) = \frac{\nu\mathcal{P}_0}{\mathcal{G}_{1,0}^*(\nu)} \tag{33}$$

Solving the partial differential eqns. (20) to (25), we get

$$\Psi(\hat{\varpi}, \breve{\varepsilon}) = \Psi(0, \breve{\varepsilon})[1 - \mathcal{Q}(\hat{\varpi})]e^{-\nu\hat{\varpi}} \tag{34}$$

$$\Gamma(\hat{\varpi}, \breve{\varepsilon}) = \Gamma(0, \breve{\varepsilon})[1 - \mathcal{L}(\hat{\varpi})]e^{-(\mathcal{T}(\breve{\varepsilon})\hat{\varpi})} \tag{35}$$

$$\Pi_{1,m}(\hat{\varpi}, \breve{\varepsilon}) = \Pi_{1,m}(0, \breve{\varepsilon})[1 - \mathcal{G}_{1,m}(\hat{\varpi})]e^{-(\mathcal{C}(\breve{\varepsilon})\hat{\varpi})};$$
$$(m = 1, 2, \dots M) \tag{36}$$

$$\Pi_2(\hat{\varpi}, \breve{\varepsilon}) = \Pi_2(0, \breve{\varepsilon})[1 - \mathcal{G}_2(\hat{\varpi})]e^{-(\mathcal{C}(\breve{\varepsilon})\hat{\varpi})} \tag{37}$$

$$\Omega(\hat{\varpi}, \breve{\varepsilon}) = \Omega(0, \breve{\varepsilon})[1 - \mathcal{B}(\hat{\varpi})]e^{-(\mathcal{C}(\breve{\varepsilon})\hat{\varpi})} \tag{38}$$

$$\mathcal{R}(\hat{\varpi}, \breve{\varepsilon}, \breve{\varepsilon}) = \mathcal{R}(\hat{\varpi}, 0, \breve{\varepsilon})[1 - \mathcal{S}(\breve{\varepsilon})]e^{-(\mathcal{C}(\breve{\varepsilon})\hat{\varpi})} \tag{39}$$

where, $\mathcal{C}(\breve{\varepsilon}) = \nu(1 - \mathcal{A}(\breve{\varepsilon}))$; $\mathcal{T}(\breve{\varepsilon}) = \mathcal{C}(\breve{\varepsilon}) + \eta[1 - \mathcal{S}^*(\mathcal{C}(\breve{\varepsilon}))]$

Theorem 1 *Utilising $\rho < 1$ as the stability criterion, the stationary dist. of the system's CR count when the BS is inactive, engaged, under two sleep modes, setup and repair are estimated by,*

Similarly, by substituting (62) mentioned in Appendix B in (39) and further by integrating it with regard to $\hat{\omega}$ and $\check{\zeta}$ and by determining the PGF as

$$\Psi(\check{\varepsilon}) = \frac{\begin{array}{c}\mathcal{P}_0\{\check{\varepsilon}(1-M) - (1-r+r\check{\varepsilon})\mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))[\mathcal{A}(\check{\varepsilon})) - \check{\varepsilon}(1-\mathcal{A}(\check{\varepsilon}))) \\ + \check{\varepsilon}U(1+s\mathcal{G}_2^*(\mathcal{C}(\check{\varepsilon})))\mathcal{B}^*(\mathcal{C}(\check{\varepsilon}))]\}[1-\mathcal{Q}^*(\nu)]\end{array}}{[(1-r+r\check{\varepsilon})\mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))[\mathcal{Q}^*(\nu) + \mathcal{A}(\check{\varepsilon})(1-\mathcal{Q}^*(\nu))] - \check{\varepsilon}]} \tag{40}$$

$$\Gamma(\check{\varepsilon}) = \frac{\begin{array}{c}\nu\mathcal{P}_0\{(1-M)[\mathcal{Q}^*(\nu) + \mathcal{A}(\check{\varepsilon}))(1-\mathcal{Q}^*(\nu)) - \check{\varepsilon}] - [\mathcal{A}(\check{\varepsilon})) - \check{\varepsilon}(1-\mathcal{A}(\check{\varepsilon}))) \\ + \check{\varepsilon}U(1+s\mathcal{G}_2^*(\mathcal{C}(\check{\varepsilon})))\mathcal{B}^*(\mathcal{C}(\check{\varepsilon}))]\}[1-\mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))]\end{array}}{(\mathcal{T}(\check{\varepsilon}))[(1-r+r\check{\varepsilon})\mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))[\mathcal{Q}^*(\nu) + \mathcal{A}(\check{\varepsilon})(1-\mathcal{Q}^*(\nu))] - \check{\varepsilon}]} \tag{41}$$

$$\Pi_{1,0}(\check{\varepsilon}) = \frac{\nu\mathcal{P}_0}{\mathcal{G}_{1,0}^*(\nu)} \frac{[1-\mathcal{G}_{1,0}^*(\mathcal{C}(\check{\varepsilon})))]}{C(\check{\varepsilon})} \tag{42}$$

$$\Pi_{1,m}(\check{\varepsilon}) = p_m \frac{\nu\mathcal{P}_0}{\mathcal{G}_{1,0}^*(\nu)} \mathcal{G}_{1,0}^*(\mathcal{C}(\check{\varepsilon}))) \frac{[1-\mathcal{G}_{1,m}^*(C(\check{\varepsilon}))]}{C(\check{\varepsilon})}; \quad (m=1,2,\ldots M) \tag{43}$$

$$\Pi_2(\check{\varepsilon}) = s[\frac{\nu\mathcal{P}_0}{\mathcal{G}_{1,0}^*(\nu)} \mathcal{G}_{1,0}^*(\mathcal{C}(\check{\varepsilon})))[p_0 + \sum_{m=1}^{M} p_m \mathcal{G}_{1,m}^*(\mathcal{C}(\check{\varepsilon})))]] \frac{[1-\mathcal{G}_2^*(C(\check{\varepsilon}))]}{C(\check{\varepsilon})} \tag{44}$$

$$\Omega(\check{\varepsilon}) = [\frac{\nu\mathcal{P}_0}{\mathcal{G}_{1,0}^*(\nu)} \mathcal{G}_{1,0}^*(\mathcal{C}(\check{\varepsilon})))[p_0 \sum_{m=1}^{M} p_m \mathcal{G}_{1,m}^*(\mathcal{C}(\check{\varepsilon})))]][1+s\mathcal{G}_2^*(\mathcal{C}(\check{\varepsilon})))] \frac{[1-\mathcal{B}^*(C(\check{\varepsilon}))]}{C(\check{\varepsilon})} \tag{45}$$

$$\mathcal{R}(\check{\varepsilon}) = \eta \times \frac{\begin{array}{c}\nu\mathcal{P}_0\{(1-M)[\mathcal{Q}^*(\nu) + \mathcal{A}(\check{\varepsilon}))(1-\mathcal{Q}^*(\nu)) - \check{\varepsilon}] - [\mathcal{A}(\check{\varepsilon})) - \check{\varepsilon}(1-\mathcal{A}(\check{\varepsilon}))) \\ + \check{\varepsilon}[1-\mathcal{G}_{1,m}^*(\mathcal{C}(\check{\varepsilon}))]\}[1-\mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))][1-\mathcal{S}^*(\mathcal{C}(\check{\varepsilon}))]\end{array}}{(\mathcal{C}(\check{\varepsilon}))(\mathcal{T}(\check{\varepsilon}))[(1-r+r\check{\varepsilon})\mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))[\mathcal{Q}^*(\nu) + \mathcal{A}(\check{\varepsilon})(1-\mathcal{Q}^*(\nu))] - \check{\varepsilon}]} \tag{46}$$

where,

$$\mathcal{P}_0 = \frac{1 - \mathcal{E}(X_1)[1-\mathcal{Q}^*(\nu)] + r - \nu\mathcal{E}(X_1)[1+\eta\mathcal{E}(S^1)]\mathcal{E}(L^1)}{r - 1 + [1-\mathcal{Q}^*(\nu)][1-\frac{r}{\mathcal{G}_{1,0}^*(\nu)}] + [\nu\mathcal{E}(L^1)(1+\mathcal{E}(S^1))\eta - 1]}$$

$$[\mathcal{E}(X_1)[1-\mathcal{Q}^*(\nu)] - M - U'(1+s)] - \nu\mathcal{E}(X_1)\mathcal{E}(L^1)(1+\eta\mathcal{E}(S^1))$$

$$[1 - \frac{[1-\mathcal{Q}^*(\nu)]}{\mathcal{G}_{1,0}^*(\nu)}] - 2\nu\mathcal{E}(X_1)\mathcal{E}(L^1)(1+\mathcal{E}(S^1))\eta + [(1-\mathcal{Q}^*(\nu))$$

$$-_\nu (1+\mathcal{E}(L^1)\mathcal{E}(S^1))\eta][\frac{1}{\mathcal{G}_{1,0}^*(\nu)}[(1+s)[1-\nu\mathcal{E}(X_1)\mathcal{E}(B^1)] + \nu s\mathcal{E}(X_1)\mathcal{E}(G_2^1)]]$$

$$U = \frac{\nu[p_0 + \sum_{m=1}^{M} p_m \mathcal{G}_{1,m}^*(\mathcal{C}(\check{\varepsilon}))]}{\mathcal{G}_{1,0}^*(\nu)}\mathcal{G}_{1,0}^*(\mathcal{C}(\check{\varepsilon}))$$

$$M = \frac{[p_0 + \sum_{m=1}^{M} p_m \mathcal{G}_{1,m}^*(\mathcal{C}(\check{\varepsilon})) - 1]}{\mathcal{G}_{1,0}^*(\nu)}\mathcal{G}_{1,0}^*(\mathcal{C}(\check{\varepsilon}))$$

$$and \quad U' = \nu\left[\frac{\nu\mathcal{E}(G_{1,0}^1) + \nu\sum_{m=1}^{M} p_m\mathcal{E}(G_{1,m}^1)}{\mathcal{G}_{1,0}^*(\nu)}\right] \tag{47}$$

Proof By substituting (57) to (61) mentioned in Appendix B in (34) to (38) and further by integrating them with respect to $\hat{\omega}$ and further by formulating the PGF as

$\Psi(\check{\varepsilon}) = \int_0^\infty \Psi(\hat{\omega},\check{\varepsilon})d\hat{\omega}$; $\Gamma(\check{\varepsilon}) = \int_0^\infty \Gamma(\hat{\omega},\check{\varepsilon})d\hat{\omega}$; $\Pi_{1,m}(\check{\varepsilon}) = \int_0^\infty \Pi_{1,m}(\hat{\omega},\check{\varepsilon})d\hat{\omega}$; $(m=1,2,\ldots M)$; $\Pi_2(\check{\varepsilon}) = \int_0^\infty \Pi_2(\hat{\omega},\check{\varepsilon})d\hat{\omega}$; $\Omega(\check{\varepsilon}) = \int_0^\infty \Omega(\hat{\omega},\check{\varepsilon})d\hat{\omega}$

$\mathcal{R}(\hat{\omega},\check{\varepsilon}) = \int_0^\infty \mathcal{R}(\hat{\omega},\check{\zeta},\check{\varepsilon})d\check{\zeta}$, $\mathcal{R}(\check{\varepsilon}) = \int_0^\infty \mathcal{R}(\hat{\omega},\check{\varepsilon})d\hat{\omega}$, we get the above mentioned eqns.

Since the normalising condition can be used to compute the single unknown, $\mathcal{P}_0$, the prob. that the server stays inactive, if there aren't any CR in orbit. As a result, by using the rule of L-Hospitals whenever necessary and by utilising

$\check{\varepsilon} = 1$ in the eqns. (40)–(46), we get,

$$\mathcal{P}_0 + \Psi(1) + \Gamma(1) + \Pi_{1,0}(1)$$
$$+ \sum_{m=1}^{M} \Pi_{1,m}(1) + \Pi_2(1) + \Omega(1) + \mathcal{R}(1) = 1.$$

□

Theorem 2 *Assuming $\rho < 1$. as a stability constraint, we can estimate the PGF of no. of CRs in the BS and orbit size dist. at a fixed point in time as follows:*

$$K(\check{\varepsilon}) = \frac{Nr(\check{\varepsilon})}{Dr(\check{\varepsilon})}$$

$$\begin{aligned}
Nr(\check{\varepsilon}) = \mathcal{P}_0\big\{&[(1 - r + r\check{\varepsilon})\mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))[\mathcal{Q}^*(v) \\
&+ \mathcal{A}(\check{\varepsilon})(1 - \mathcal{Q}^*(v))] - \check{\varepsilon}]\mathcal{T}(\check{\varepsilon})[M + \mathcal{C}(\check{\varepsilon}) \\
&+ vU(1 + s - \mathcal{B}^*(\mathcal{C}(\check{\varepsilon}))(1 + s\mathcal{G}_2^*(\mathcal{C}(\check{\varepsilon}))))] \\
&+ \check{\varepsilon}(1 - M)[\mathcal{T}(\check{\varepsilon})\mathcal{C}(\check{\varepsilon})(1 - \mathcal{Q}^*(v)) \\
&+ v(\mathcal{Q}^*(v) + \mathcal{A}(\check{\varepsilon})(1 - \mathcal{Q}^*(v)))[1 - \mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))][\mathcal{C}(\check{\varepsilon}) \\
&+ [1 - \mathcal{S}^*(\mathcal{C}(\check{\varepsilon}))\eta]]] - [\mathcal{A}(\check{\varepsilon}) \\
&- \check{\varepsilon}(1 - \mathcal{A}(\check{\varepsilon})) \\
&+ U(1 + s\mathcal{G}_2^*(\mathcal{C}(\check{\varepsilon}))\mathcal{B}^*(\mathcal{C}(\check{\varepsilon}))\check{\varepsilon}][(1 - r \\
&+ r\check{\varepsilon})\mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))(1 - \mathcal{Q}^*(v))\mathcal{C}(\check{\varepsilon}) \\
&\mathcal{T}(\check{\varepsilon}) + \check{\varepsilon}v[1 - \mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))][\mathcal{C}(\check{\varepsilon}) \\
&+ [1 - \mathcal{S}^*(\mathcal{C}(\check{\varepsilon}))\eta]]]\big\}
\end{aligned}$$

$$\begin{aligned}
Dr(\check{\varepsilon}) = &[(1 - r + r\check{\varepsilon})\mathcal{L}^*(\mathcal{C}(\check{\varepsilon}))[\mathcal{Q}^*(v) \\
&+ \mathcal{A}(\check{\varepsilon})(1 - \mathcal{Q}^*(v))] - \check{\varepsilon}]\mathcal{T}(\check{\varepsilon})\mathcal{C}(\check{\varepsilon})
\end{aligned} \qquad (48)$$

$$H(\check{\varepsilon}) = \frac{Nr_e(\check{\varepsilon})}{Dr(\check{\varepsilon})}$$

$$\begin{aligned}
Nr_e(\check{\varepsilon}) = \mathcal{P}_0\big\{&[(1 - r + r\check{\varepsilon})\mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))[\mathcal{Q}^*(v) \\
&+ \mathcal{A}(\check{\varepsilon})(1 - \mathcal{Q}^*(v))] - \check{\varepsilon}]\mathcal{T}(\check{\varepsilon})[M + \mathcal{C}(\check{\varepsilon}) \\
&+ vU(1 + s - \mathcal{B}^*(\mathcal{C}(\check{\varepsilon}))(1 + s\mathcal{G}_2^*(\mathcal{C}(\check{\varepsilon}))))] \\
&+ (1 - M)[\mathcal{T}(\check{\varepsilon})\mathcal{C}(\check{\varepsilon})(1 - \mathcal{Q}^*(v)) \\
&+ v(\mathcal{Q}^*(v) + \mathcal{A}(\check{\varepsilon})(1 - \mathcal{Q}^*(v)))[1 - \mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))][\mathcal{C}(\check{\varepsilon}) \\
&+ [1 - \mathcal{S}^*(\mathcal{C}(\check{\varepsilon}))\eta]]] - [\mathcal{A}(\check{\varepsilon}) \\
&- \check{\varepsilon}(1 - \mathcal{A}(\check{\varepsilon})) \\
&+ U(1 + s\mathcal{G}_2^*(\mathcal{C}(\check{\varepsilon})))\mathcal{B}^*(\mathcal{C}(\check{\varepsilon}))\check{\varepsilon}][(1 - r \\
&+ r\check{\varepsilon})\mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))(1 - \mathcal{Q}^*(v))\mathcal{C}(\check{\varepsilon}) \\
&\mathcal{T}(\check{\varepsilon}) + v[1 - \mathcal{L}^*(\mathcal{T}(\check{\varepsilon}))][\mathcal{C}(\check{\varepsilon}) \\
&+ [1 - \mathcal{S}^*(\mathcal{C}(\check{\varepsilon}))\eta]]]\big\}
\end{aligned} \qquad (49)$$

where $\mathcal{P}_0$ is represented by eqn. (47)

Proof The subsequent eqns. are utilized to calculate the PGF for the no. of CRs in the 5G BS ($K(\check{\varepsilon})$) and in the orbit ($H(\check{\varepsilon})$).

$K(\check{\varepsilon}) = \mathcal{P}_0 + \Psi(\check{\varepsilon}) + \Pi_{1,0}(\check{\varepsilon}) + \sum_{m=1}^{M} \Pi_{1,m}(\check{\varepsilon}) + \Pi_2(\check{\varepsilon}) + \Omega(\check{\varepsilon}) + \check{\varepsilon}[\Gamma(\check{\varepsilon}) + \mathcal{R}(\check{\varepsilon})]$
and
$H(\check{\varepsilon}) = \mathcal{P}_0 + \Psi(\check{\varepsilon}) + \Pi_{1,0}(\check{\varepsilon}) + \sum_{m=1}^{M} \Pi_{1,m}(\check{\varepsilon}) + \Pi_2(\check{\varepsilon}) + \Omega(\check{\varepsilon}) + \Gamma(\check{\varepsilon}) + \mathcal{R}(\check{\varepsilon})$

The eqns. (48) and (49) may be computed directly when the eqns. (40) to (46) are substituted in the preceding findings.

□

4 System performance measures

Note that the eqn. (47) gives the SS prob. when the BS is empty but available in the system. Thus, the probabilities of different BS state are derived from (40) to (46) whenever the system meets the stability criterion $\rho < 1$, which is stated below,

1. Let Ψ represent the SS prob. when the BS is idle during the retrial time

$$\begin{aligned}
\Psi &= lim_{\check{\varepsilon} \to 1}\Psi(\check{\varepsilon}) \\
&= \frac{\begin{aligned}\mathcal{P}_0[1 - M' - \frac{1}{\mathcal{G}_{1,0}^*(v)}[r - v\mathcal{E}(X_1)\mathcal{E}(L^1)(1 + \eta\mathcal{E}(S^1)) - vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)] \\ - (1+s)[U' - \frac{v\mathcal{E}(X_1)\mathcal{E}(B^1)+1}{\mathcal{G}_{1,0}^*(v)}] - 2\mathcal{E}(X_1)[1 - \mathcal{Q}^*(v)]\end{aligned}}{r + \mathcal{E}(X_1)[1 - \mathcal{Q}^*(v)] - v\mathcal{E}(X_1)[1 + \eta\mathcal{E}(S^1)]\mathcal{E}(L^1) - 1}
\end{aligned} \qquad (50)$$

2. Let Γ represent the SS prob. when the BS is active

$$\begin{aligned}
\Gamma &= lim_{\check{\varepsilon} \to 1}\Gamma(\check{\varepsilon}) \\
&= \frac{\begin{aligned}v\mathcal{P}_0[\mathcal{E}(X_1)[1 - \mathcal{Q}^*(v)] - M' - 2\mathcal{E}(X_1) + U' + \frac{1+s}{\mathcal{G}_{1,0}^*(v)} \\ [-v\mathcal{E}(X_1)\mathcal{E}(B^1) + 1] - \frac{vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)}{\mathcal{G}_{1,0}^*(v)}]\mathcal{E}(L^1)\end{aligned}}{r + \mathcal{E}(X_1)[1 - \mathcal{Q}^*(v)] - v\mathcal{E}(X_1)[1 + \eta\mathcal{E}(S^1)]\mathcal{E}(L^1) - 1}
\end{aligned} \qquad (51)$$

3. Let $\Pi_{1,m}$ represent the SS prob. when the BS is under SM1

$$\begin{aligned}
\Pi_{1,m} &= lim_{\check{\varepsilon} \to 1}\Pi_{1,m}(\check{\varepsilon}) \\
&= \frac{p_m v\mathcal{P}_0}{\mathcal{G}_{1,0}^*(v)}\mathcal{E}(G_{1,0}^1)[v\mathcal{E}(X_1)\mathcal{E}(G_{1,m}^1)]; \quad (m = 1, 2, \dots M)
\end{aligned} \qquad (52)$$

4. Let Π_2 represent the SS prob. when the BS is under SM2

$$\begin{aligned}
\Pi_2 &= lim_{\check{\varepsilon} \to 1}\Pi_2(\check{\varepsilon}) \\
&= s\mathcal{P}_0\frac{v\mathcal{E}(G_{1,0}^1) + v\sum_{m=1}^{M} p_m\mathcal{E}(G_{1,m}^1)}{\mathcal{G}_{1,0}^*(v)}\mathcal{E}(G_2^1)
\end{aligned} \qquad (53)$$

5. Let Ω represent the SS prob. that the BS is under set up

$$\begin{aligned}
\Omega &= lim_{\check{\varepsilon} \to 1}\Omega(\check{\varepsilon}) \\
&= \mathcal{P}_0\frac{v\mathcal{E}(G_{1,0}^1) + v\sum_{m=1}^{M} p_m\mathcal{E}(G_{1,m}^1)}{\mathcal{G}_{1,0}^*(v)}sv\mathcal{E}(X_1)\mathcal{E}(G_2^1)\mathcal{E}(B^1)
\end{aligned} \qquad (54)$$

6. Let $\mathcal{R}$ represent the SS prob. that the BS is under repair

$$\mathcal{R} = lim_{\breve{\varepsilon} \to 1} \mathcal{R}(\breve{\varepsilon})$$

$$= \frac{\begin{array}{c} \nu\eta\mathcal{P}_0[\mathcal{E}(X_1)[1 - \mathcal{Q}^*(\nu)] - M' - 2\mathcal{E}(X_1) + U' + \frac{1+s}{\mathcal{G}^*_{1,0}(\nu)} \\ [-\nu\mathcal{E}(X_1)\mathcal{E}(B^1) + 1] - \frac{\nu s \mathcal{E}(X_1)\mathcal{E}(G_2^1)}{\mathcal{G}^*_{1,0}(\nu)}]\mathcal{E}(L^1)\mathcal{E}(S^1) \end{array}}{r + \mathcal{E}(X_1)[1 - \mathcal{Q}^*(\nu)] - \nu\mathcal{E}(X_1)[1 + \eta\mathcal{E}(S^1)]\mathcal{E}(L^1) - 1}$$

(55)

where,

$$M' = \frac{\nu\mathcal{E}(G^1_{1,0}) + \nu\sum_{m=1}^{M} p_m\mathcal{E}(G^1_{1,m})}{\mathcal{G}^*_{1,0}(\nu)}$$

4.1 Mean size of a system and its orbit

When the system is in a SS,

(i) By differentiating (48) with regard to $\breve{\varepsilon}$ and providing $\breve{\varepsilon} = 1$ under stability conditions, the mean no. of clients in the orbit (L_q) is computed.

$$L_q = H'(1) = \lim_{\breve{\varepsilon} \to 1} \frac{d}{d\breve{\varepsilon}} H(\breve{\varepsilon})$$
$$= \mathcal{P}_0 \left[\frac{Nr_q''''(1)Dr_q'''(1) - Dr_q''''(1)Nr_q'''(1)}{4(Dr_q'''(1))^2} \right]$$

(ii) By differentiating (49) with regard to $\breve{\varepsilon}$ and providing $\breve{\varepsilon} = 1$ under stability conditions, the mean no. of clients in the system (L_s) is computed.

$$L_s = K'(1) = \lim_{\breve{\varepsilon} \to 1} \frac{d}{d\breve{\varepsilon}} K(\breve{\varepsilon})$$
$$= \mathcal{P}_0 \left[\frac{Nr_s''''(1)Dr_q'''(1) - Dr_q''''(1)Nr_q'''(1)}{4(Dr_q'''(1))^2} \right]$$

All the above mentioned values are given in Appendix C.

(iii) With the aid of Little's approach, it is possible to predict the amount of time a user can expect to spend in the system (W_s) and the amount of time a user can expect to spend in the queue (W_q), (i.e) $W_s = \frac{L_s}{\nu\mathcal{E}(X_1)}$ and $W_q = \frac{L_q}{\nu\mathcal{E}(X_1)}$.

5 Special cases

Case (i): No batch arrival, no retrial, no repair, no feedback and no setup. $Pr[\mathcal{U}_k = n] = 1$, $\mathcal{Q}^*(\nu) \to 1$, $\eta = 0$, $r = 0$, $\varphi(\hat{\varpi}) = 0$. Then our model will be reduced to a $M/G/1$ queue with M short sleep and single long sleep under N policy which coincides with the results of Deepa et al. [9].

Case (ii): No batch arrival, no retrial, no repair, no feedback and N=1. $Pr[\mathcal{U}_k = n] = 1$, $\mathcal{Q}^*(\nu) \to 1$, $\eta = 0$, $r = 0$. Then our model will be reduced to a $M/G/1$ queue with M short sleep and single long sleep which coincides with the outcome of a Kalita and Selvamuthu [19].

6 Numerical examples

To demonstrate the many alternatives for the system's dynamic responsiveness, we'll employ MATLAB in this part. In addition, the exponentially distributed retrial times, service, dual sleeps, setup, and repair times have been significantly analysed. Random selection is made in order to determine whether numerical measures meet the stability requirements. The calculated values of certain aspects of the model, including the prob. that the server is inactive Π_0, the mean queue size (L_q), and the mean queue waiting time (W_q), are presented in Tables 1, 2, 3, and 4.

Table 1 depicts that as retrial rate (ϑ) rises, L_q also rises, whereas $\mathcal{P}_0$, Γ, $\mathcal{R}$ subsides for the value of $\nu = 0.4$, $\eta = 5$, $r = 0.5$, $s = 0.9$, $\theta = 5$, $p_0 = 0.5$, $\Pi_{1,m} = 0.1$.

Table 2 shows that for the as the arrival rate (ν) of CR goes up, so does $\mathcal{P}_0$, L_s, W_s and Γ for the value of $\nu = 0.5$, $\eta = 9$, $r = 0.2$, $s = 0.7$, $\theta = 6$, $p_0 = 0.5$, $\Pi_2 = 0.5$, $\upsilon_2 = 4.5$.

Table 3 demonstrates that for the mounting value of the service rate (ϑ), L_q and W_q also mounts whereas $\mathcal{P}_0$, and Π_2 decline for the value of $\nu = 0.2$, $\eta = 4$, $r = 0.9$, $s = 0.1$, $\vartheta = 5$, $p_m = 0.9$, $\Pi_2 = 0.5$, $\varphi = 4.5$.

As illustrated in Table 4, if boost the SM2 rate (υ_2), L_s, and W_s also rises but, Γ and Ω reduces for the value of $\nu = 0.1$, $\eta = 7$, $r = 0.2$, $s = 0.7$, $\theta = 8$, $p_0 = 0.5$, $\zeta = 0.5$, $\vartheta = 4.5$.

Here, Figs. 7 and 8 depict the 2D and 3D graphs, respectively representing how the parameters ν, ϑ, υ_2, θ influence the key metrics. In Fig. 7a as the retrial rate ϑ rises, L_q also rises, whereas $\mathcal{P}_0$ and $\mathcal{R}$ decline. The arrival rate ν mounts for the increasing values of L_s, W_s and Γ is presented in Fig. 7b. In Fig. 7c, for the mounting value of service rate θ, L_q. W_q and Π_2 declines. Figure 7d depicts that as SM2 rate υ_2 elevates, L_s and W_s also mount, whereas Γ diminish.

Table 1 The effect of retrial rate (ϑ) on $\mathcal{P}_0$, L_q, Ω, $\mathcal{R}$

Retrial rate(ϑ)	$\mathcal{P}_0$	L_q	Ω	$\mathcal{R}$
2	0.0196	0.4945	0.0879	0.0698
2.5	0.0189	0.5167	0.0846	0.0692
3	0.0182	0.5375	0.0816	0.0687
3.5	0.0176	0.5572	0.0788	0.0683
4	0.0170	0.5757	0.0763	0.0679

Table 2 The impact of arriving CR (ν) on $\mathcal{P}_0$, L_s, W_s, Γ

CR rate (ν)	$\mathcal{P}_0$	L_s	W_s	Γ
0.4	0.0060	0.2038	3.0264	0.0060
0.5	0.0083	0.2452	3.2841	0.0077
0.6	0.0106	0.2775	3.6040	0.0084
0.7	0.0128	0.3045	4.0790	0.0089
0.8	0.0170	0.3304	6.6726	0.0096

Table 3 The effect of service rate (θ) on $\mathcal{P}_0$, L_q, W_q, Π_2

BS (θ)	$\mathcal{P}_0$	L_q	W_q	Π_2
1	0.0128	0.2458	5.2427	0.0114
2	0.0114	0.2114	5.1412	0.0102
3	0.0110	0.2003	4.9921	0.0099
4	0.0108	0.1945	4.7298	0.0097
5	0.0107	0.1907	4.0680	0.0096

Table 4 The effect of SM2 (υ_2) on L_s, W_s, Γ, Ω

SM2 (υ_2)	L_s	W_s	Γ	Ω
1	0.5643	1.3490	0.0113	0.3877
1.5	0.6411	1.3861	0.0104	0.3021
2	0.6900	1.4493	0.0098	0.2203
2.5	0.7214	1.5599	0.0093	0.1409
3	0.7413	1.7721	0.0090	0.0763

Table 5 Energy consumption in various states

Server states	Values J/s
P_{OS}	11,577
P_{SM1}	5788.5
P_{SM2}	1736.55
P_{SS}	10,419.3
P_{RP}	9840.45

Figure 8a depicts that as ϑ mounts, L_q mount, whereas $\mathcal{P}_0$ decreases. Figure 8b depicts that as ν rises L_s and W_s also elevate. In Fig. 8c as θ mounts, L_q and W_q diminish. As υ_2 elevates in Fig. 8d, L_s and W_s also rise.

The influence of diverse characteristics on the system's performance standards can perhaps be ascertained using the numerical findings shown above.

7 Evaluation of energy conservation

7.1 Power savings factor

The percentage of total time that BSs remained in SM1 and SM2 is the power savings factor. By using the (52) and (53), we obtain the following,

$$\mathcal{PS} = (\Pi_{1,m} + \Pi_2) \times 100\%$$

7.2 Expected energy conservation

The BS is portrayed as an $M^{[X]}/G/1$ feedback retrial queue in the proposed framework, which includes dual sleep modes, setup, and repair phases. New CRs show up at a constant rate, ν, according to a Poisson distribution. The newly arrived CRs have an immediate active service time S. In the operation state, it is believed that power consumption is P_{OS}. The BS will keep running until there are no more CRs on the BS. Once the BS becomes empty, it will switch to sleep mode. P_{SM1} and P_{SM2} are the power consumed when the device is in sleep modes 1 and 2 respectively. Until N CRs are discovered at the end of any sleep, the BS sleep will keep looping. If the amount of gathered CRs is greater than N at the end of any sleep, the BS will begin setting up. Setup time power conservation is provided by P_{SS}. Also, if there is any breakdown in the BS, then it will be sent for repair immediately. Thus, the energy consumed during repair time is denoted as P_{RP}. For simplicity, we'll call BS's power consumption as PC per second. As a result, we may employ the following formula, to estimate the BS's average energy consumption during a certain time period and moreover, based on analysis by [24, 28, 35] the energy consumption values for various states has been addressed in Table 5.

$$\begin{aligned} E(PC) = {} & P_{OS}\Gamma + P_{SM1}\Pi_{1,m} + P_{SM2}\Pi_2 \\ & + P_{SS}\Omega + P_{RP}\mathcal{R} \end{aligned}$$

When BS is fully functioning, all of its components will be using energy at once. Hence, it will consume full energy. Further, if a breakdown occurs and the damaged BS components are sent for repair, the impact on energy consumption will be reduced a bit during the repair stage since the repaired components will not be functioning. Then, if there is no CR available, BS will go to SM1. Energy usage is drastically reduced in SM1 stage due to the fact that many non-essential components will be disabled. At the end of the Mth brief sleep period under SM1, if the number of requests remains below the limit of N, the BS transitions from SM1 to SM2 (long sleep). In this state, nearly all of the system's features are deactivated to save energy. This results in a much more reduction in power usage compared to SM1 stage. Hence maximum energy will be saved in this stage. This is because the BS will undergo sleep for a long period. However, even in either of the two sleep modes, the energy spent will not be zero because certain aspects, like the one that helps tally the number of CRs in the queue, are still active. When the
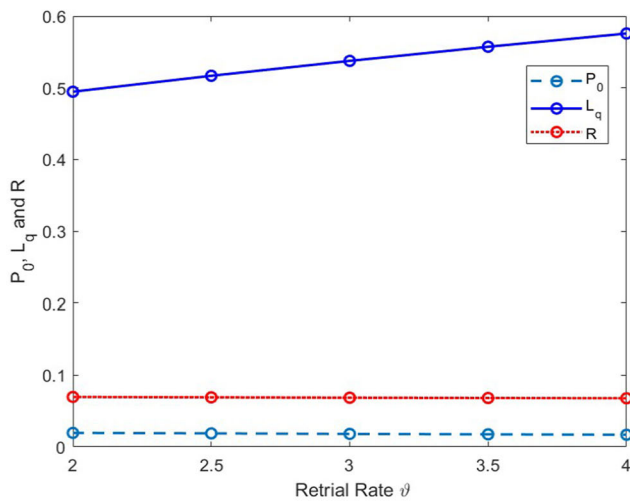
(a) $\mathcal{P}_0$, L_q and $\mathcal{R}$ vs retrial rate ϑ
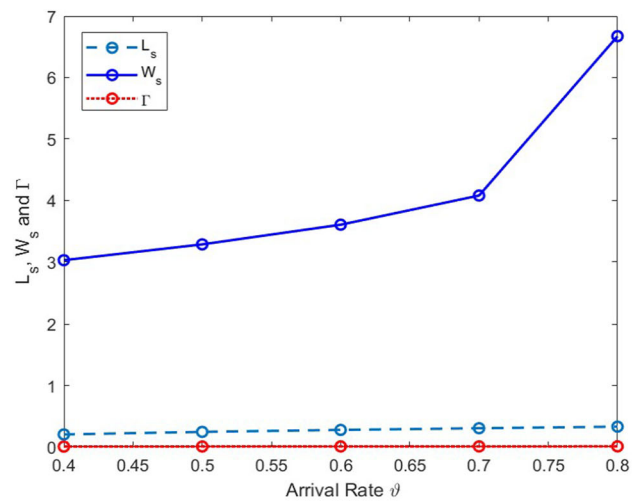
(b) L_s, W_s and Γ vs CRs arrival rate ν
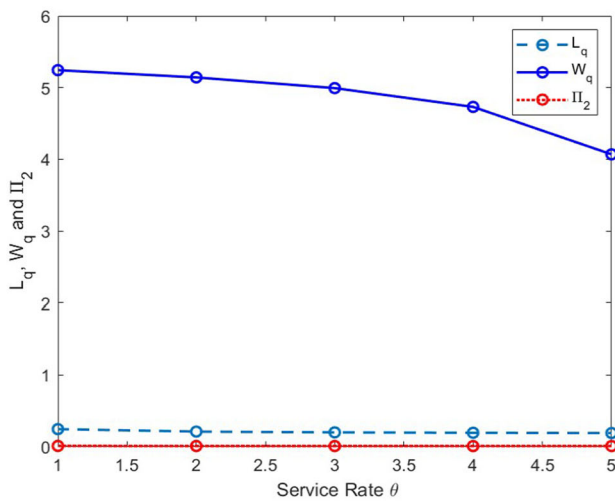
(c) L_q, W_q and Π_2 vs service rate θ
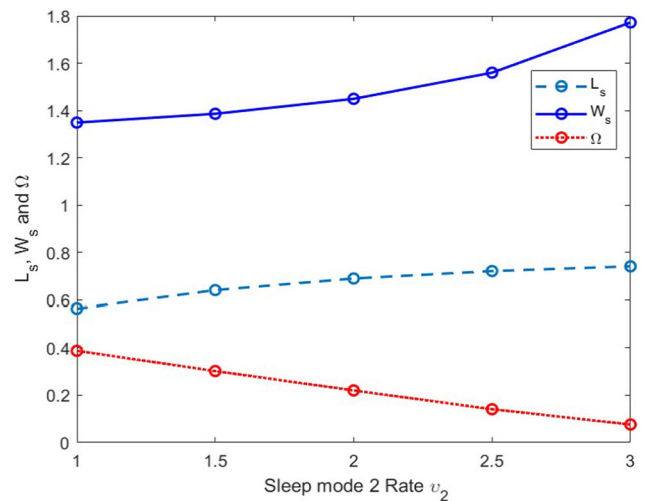
(d) L_s, W_s and Ω vs SM2 rate υ_2

Fig. 7 2D representation of graphs

N CRs have gathered at the conclusion of any sleep, the BS will use some energy to begin setting up and preparing for service.

7.3 Expected delay

The amount of time a UR waits in queue before the service starts is known as the delay. Queue delay occurs when a newly arrived CR must wait in queue to receive service if the BS is already serving another CR. When the BS is in setup, sleep or repair mode, the server in the proposed model won't offer service. We refer to this delay as an additional delay. For an $M^{[X]}/G(a,b)/1$ queuing system with N-policy, multiple vacations, and setup, Krishna Reddy et al. [22] computed

additional delay as the total of idle duration owing to multiple vacation processes and mean length of setup time.

As a result it is clear that queue delays are the sole thing affecting CRs when BS is in active state. In the meantime, sleep, queue and setup delays impact the arriving CRs during both the sleep mode BS. Also, there will be setup and queue delays for the CRs that show up during setup. Further, the CRs will be impacted only by the queue delay when they are in the repair mode. Therefore, in light of this, the expected delay or the mean latency E(L) for a CR is determined as follows:

$$E(L) = E(D) + E(AD) + E(SD) + E(RD)$$
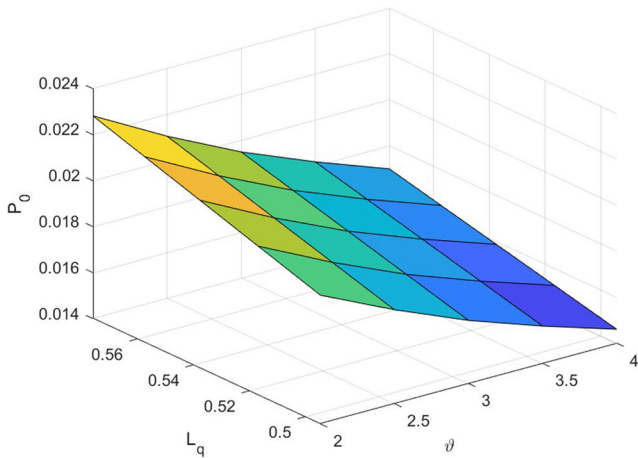
(a) L_q, $\mathcal{P}_0$ vs retrial rate ϑ
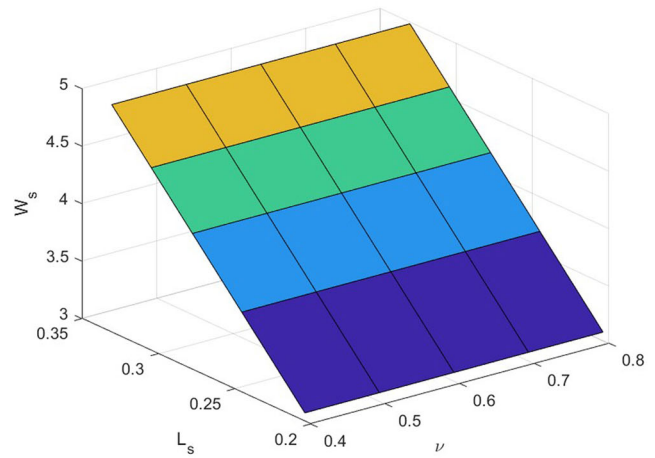
(b) L_s, W_s vs arrival rate ν
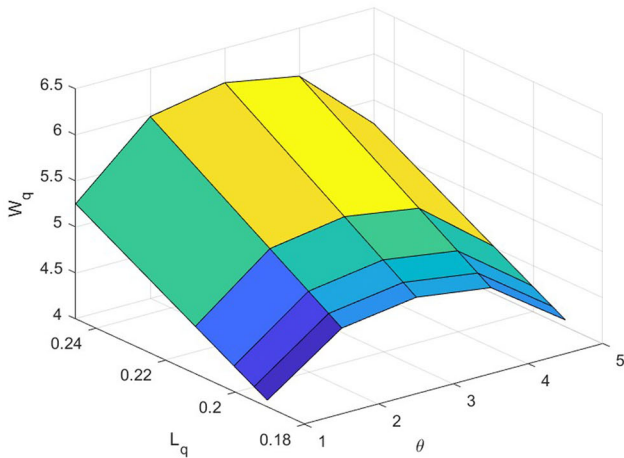
(c) L_q, W_q verse service rate θ
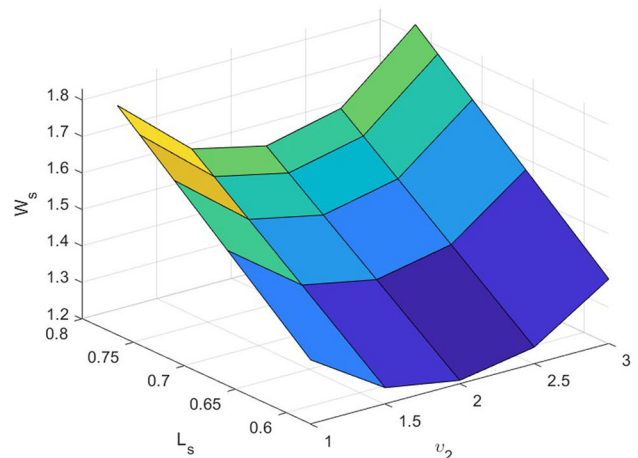
(d) L_s, W_s vs SM1 rate υ_1

Fig. 8 3D representation of graphs

where, E(D) = $\frac{v\mathcal{E}(X_2)\mathcal{E}(L^2)}{2(1-v\mathcal{E}(X_1)\mathcal{E}(L^1))}$ is the expected queue delay stated by Shortle et al. [32]; E(AD) denotes the expected additional delay (i.e.) delay due to both the sleep mode, E(SD) represents the expected setup delay and E(RD) denotes the expected repair delay.

Here, by differentiating eqns. (59), (60) and (61) w.r.t ($\check{\varepsilon}$) and taking $lim_{(\check{\varepsilon})\to 1}$ and utilizing L'hospital rule, we have

$$E(AD) = \frac{p_m v}{\mathcal{G}_{1,0}^*(v)}\mathcal{E}(G_{1,0}^1)[v\mathcal{E}(X_1)\mathcal{E}(G_{1,m}^1)]$$
$$+\frac{v\mathcal{E}(G_{1,0}^1) + sv\sum_{m=1}^M p_m\mathcal{E}(G_{1,m}^1)}{\mathcal{G}_{1,0}^*(v)}\mathcal{E}(\mathcal{G}_2^1)$$

and

$$E(SD) = \frac{v\mathcal{E}(G_{1,0}^1) + v\sum_{m=1}^M p_m\mathcal{E}(G_{1,m}^1)}{\mathcal{G}_{1,0}^*(v)}$$

$$E(RD) = \frac{sv\mathcal{E}(X_1)\mathcal{E}(G_2^1)\mathcal{E}(B^1)}{\mathcal{E}(X_1)[1-\mathcal{Q}^*(v)] - v\mathcal{E}(X_1)[1+\eta\mathcal{E}(S^1)]\mathcal{E}(L^1)}$$

Therefore, from the aforementioned three equations, we can get the mean latency E(L) of the system from which the maximum latency that the system experience can be obtained.

Springer

7.3.1 Numerical analysis on expected delay

In this section, the mathematical investigation and computational study of different components of the delay in the systems have been presented, along with the optimization of E(PC) and E(L) which primarily aim to model, analyze, and optimize their effects through numerical methods.

Tables 6, 7, 8, and 9 explore the influence of various system parameters on different aspects of delay. It is evident from Table 6 that as the arrival rate of the CR increases, E(D), E(AD), E(SD), E(RD), and E(L) also increase. This relationship is further illustrated in the 2D graph presented in Fig. 9a. Notably, when the BS serves any CR, there is an automatic reduction in delay. This is discussed in Table 7, which indicates that as the service rate increases, both E(D) and E(L) decrease. This trend is also depicted in the 2D graph shown in Fig. 9b.

Furthermore, by examining Tables 8 and 9, it becomes evident that an increase in SM1 results in a corresponding rise in E(AD), E(SD), and E(L). This is attributed to the fact that when the BS enters sleep mode, no services are provided, consequently causing delays to escalate. Additionally, it is noticeable that as the SM1 rate increases, the expected delay E(L) also increases, albeit at a less pronounced rate compared to the steep increase observed with the SM2 rate. This stark rise in delay during SM2 is attributed to its longer duration of sleep compared to SM1. However, a slight variation can be observed in E(AD) and E(SD) between SM1 and SM2 as their rates increase. Furthermore, the influence of SM1 and SM2 rates is illustrated through 2D graphs in Fig. 10a, b. Moreover, Fig. 10c depicts the impact of E(L) by raising the number of sleeps in SM1. It is clearly observed that by elevating the number of sleeps in SM1, automatically increases the E(L) due to the raise in the duration of sleep.

7.4 Numerical analysis

The implications of SM1 on expected energy conservation [E(PC)] and power saving factor [$\mathcal{PS}$] at various maximum sleeps in SM1 of 5G BS are presented in Fig. 11a, b, by

Table 6 The effect of CR rate (ν) on $E(D)$, $E(AD)$, $E(SD)$, $E(RD)$, $E(L)$

CR rate (ν)	$E(D)$	$E(AD)$	$E(SD)$	$E(RD)$	$E(L)$
0.5	0.0632	1.8845	4.8632	0.9128	54.8795
0.6	0.0778	2.1201	5.4712	1.0269	62.7182
0.7	0.0941	2.4229	6.2528	1.1736	73.1700
0.8	0.1125	2.8267	7.2949	1.3692	87.8028
0.9	0.1333	3.3921	8.7538	1.6431	109.7523

Table 7 The effect of service rate (θ) on $E(D)$, $E(L)$

Service (ν)	$E(D)$	$E(L)$
0.2	0.3522	65.0820
0.3	0.2045	64.6701
0.4	0.1125	63.6647
0.5	0.0555	62.6503
0.6	0.0220	61.6400

Table 8 The effect of SM1 rate (υ_1) on $E(AD)$, $E(SD)$, $E(L)$

SM1 rate (υ_1)	$E(AD)$	$E(SD)$	$E(L)$
2	6.9367	12.7283	36.1828
3	8.9220	13.5650	39.3328
4	10.9033	15.0617	42.4828
5	12.8867	16.2283	45.6382
6	14.8700	17.9350	48.7828

Table 9 The effect of SM2 rate (υ_2) on $E(AD)$, $E(SD)$, $E(L)$

SM2 rate (υ_2)	$E(AD)$	$E(SD)$	$E(L)$
2	8.9200	13.8950	55.6828
3	10.9050	17.8425	68.3470
4	12.8900	22.7900	81.6111
5	14.8750	29.7375	94.8753
6	16.8600	31.6850	108.1395

assuming the parametric values $\nu = 0.4$, $\eta = 5$, $r = 0.5$, $s = 0.9$, $\theta = 5$, $p_0 = 0.5$.

In Fig. 11a E(PC) falls as SM1 rate rises. This occurs because the majority of the components are not in a working state and just a few of the components operate in SM1. Similarly, Fig. 11b depicts that as $\mathcal{PS}$ of the 5G BS falls as SM1 rate increases. This occurs because a shorter SM1 results in a reduced $\mathcal{PS}$.

Further, the implication of SM2 on expected energy conservation [E(PC)] and power saving factor [$\mathcal{PS}$] is presented in Fig. 11c, d. It is clear from Fig. 11c that there is a drastic fall in E(PC) compared to Fig. 11a as the SM2 rate elevates. This is because most of the components will be in sleep mode during SM2, and since it is longer sleep, the E(PC) declines drastically compared to SM1. However, $\mathcal{PS}$ initially declines and gradually rises as the SM2 elevates, which is presented in Fig. 11d. This is due to a longer sleep mode (SM2), which leads to a higher power saving of the 5G BS, while a shorter sleep mode (SM1) leads to a lower power saving. The effects of SM2 on three distinct system state probabilities are depicted in Fig. 11e. It turns out that while the probability of the BS under active state (Γ) decreases as the SM2 rate increases, the probability of the BS under retrial state (Ψ)
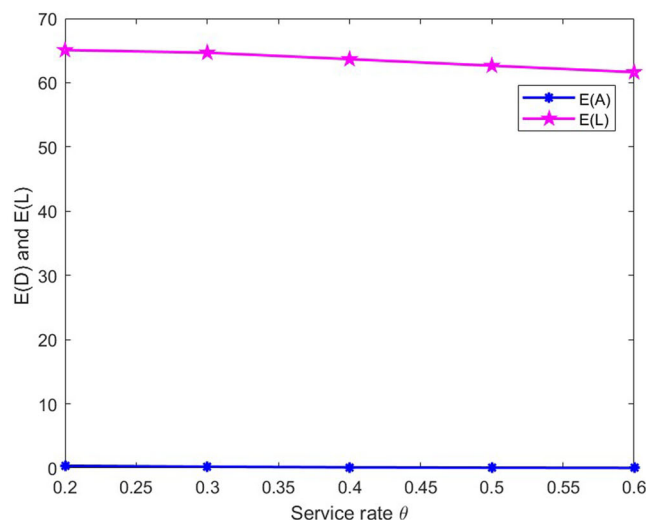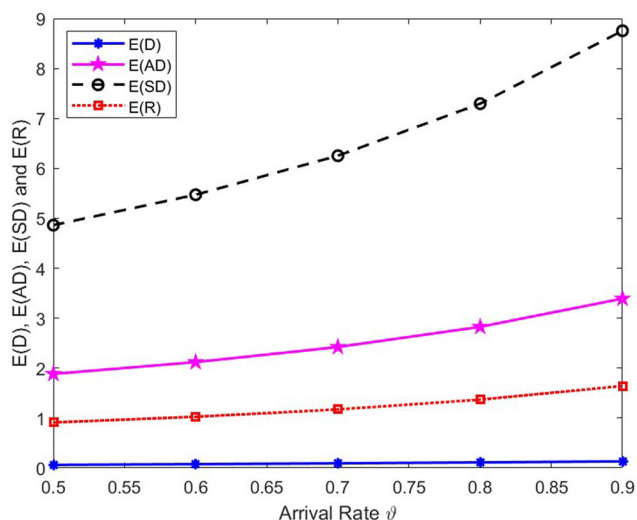
(a) Arrival rate vs E(D), E(AD), E(SD), E(RD), and E(L)

(b) Service rate vs E(D) and E(L)

Fig. 9 Implications of arrival and service rate on distinct components of delay

and setup state (Ω) rises. This is because the longer the BS is in sleep mode, the less service it provides, which causes Γ to decrease while Ψ and Ω to increase.

Figure 12 depicts the impact of SM1 and SM2 on $E(PC)$ and $\mathcal{PS}$ at different setup rates in a BS. Here, Fig. 12a, b clearly address the fact that as SM1 rate increases, $E(PC)$ declines, however $\mathcal{PS}$ rises. Similarly, as SM2 rate increases, $E(PC)$ declines, but $\mathcal{PS}$ rises which is depicted in Fig. 12c, d respectively. This occurs because, as the 5G setup rate evolves, it takes longer for the 5G network to begin serving the request, which lowers the 5G BS's average power consumption and raises its power saving factor. Also, from the above Fig. 12b, d, we can state that between SM1 and SM2 more power has been saved in SM2 due to the long sleep duration in SM2 compared to SM1.

Fig. 13a illustrates how the arrival rate affects the probability of the BS under the setup state (Ω) at various set-up rates. It is clear from Fig. 13a that the probability of the BS in the set-up state reduces as the arrival rate increases for different set-up rate. This is triggered by the fact that as the arrival rate increases, there is a greater chance that the BS will stay in an active state, which lowers the probability that it will be in a set-up state. Moreover, we know that the probability of the BS under set-up state increases with set-up duration. As a result raises the system's additional latency will also increase.

Fig. 13b addresses the effect of the arrival rate on various system state probabilities namely Γ, Ψ, $\Pi_{1,m}$, Π_2 and Ω. The probability that the BS under both the active (Γ) and retrial (Ψ) states rises with an elevation in arrival rate, while the probability of other system states such as $\Pi_{1,m}$, Π_2 and

Ω falls. This is due to the fact that as the arrival rate rises, the BS undergoes a set-up time transition, which leads them to go into a frequent off-on (sleep-active) state. Yet, we find that the amount of time spent in this state is quite small when compared to the amount of time spent in the other states. Therefore, it can be inferred from Fig. 13a, b that there is a negligible additional latency in relation to the on-off (sleep-active) state.

7.5 Comparative analysis

This subsection juxtaposes the suggested model's performance outcomes with the existing literature. Sleeping methods, which enable part of the components to be turned off when there is minimal traffic, are adopted to reduce power usage. Kalita and Selvamuthu [19] have analysed 5G BG by using a $M/G/1$ queue with a distinct sleeping strategies. According to their model, BS will reactivate if any CR arrives while it is in the sleep phase. Instead of waking for a single CR, the BS can be made to wake up after accruing a predetermined amount of CRs in order to save power. In such cases, N policy can be very beneficial to conserve energy which was addressed by Deepa et al. [9] and Deena Merit et al. [7]. However, bulk arrival was not taken into account in any of the aforementioned works. Though, authors [2] and [3] have addressed the bulk arrival, they haven't analysed with a real life example. In the light of this, the proposed study aimed to analyse a 5G BS for energy conservation by implementing N policy in two different sleep modes while accounting for bulk arrival. Additionally, consideration was given to the feedback case and the retrial case.
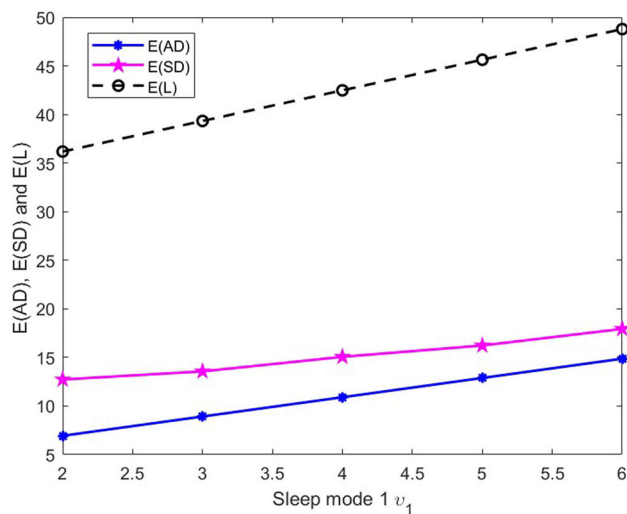
(a) SM1 rate vs E(AD), E(SD), and E(L)
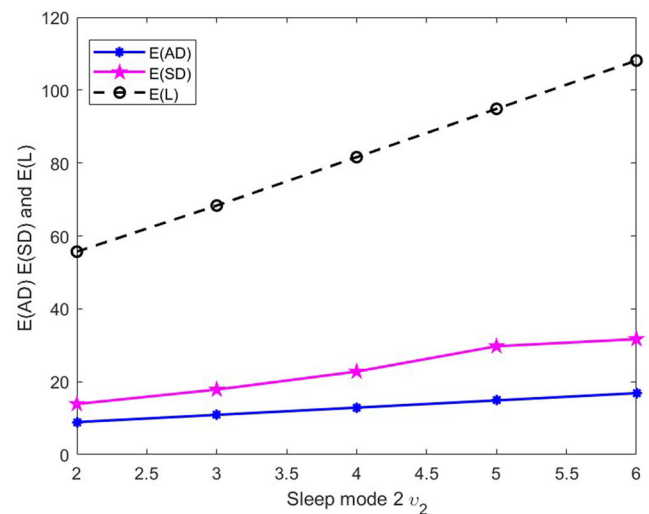
(b) SM2 rate vs E(AD), E(SD), and E(L)
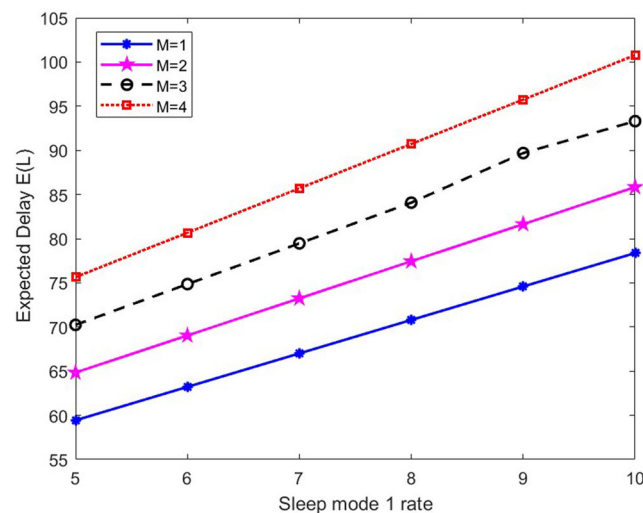
(c) SM1 rate vs expected delay E(L)

**Fig. 10** Implications of SM1 and SM2 rate on distinct components of delay

A numerical comparison with the existing work has been offered to illustrate the efficacy of the suggested model by defining the following two models,

Existing Model: An $M/G/1$ queue with M short sleep, a long sleep, and setup time without N policy was employed to depict a 5G BS by Kalita and Selvamuthu [19].

Proposed Model: An $M^{[x]}/G/1$ retrial queue with M short sleep, a long sleep, and setup time under N policy subjected to feedback and repair was proposed to simulate a 5G BS.

We are able to observe the effect of SM1 and SM2 on $E(PC)$ and $\mathcal{PS}$ at different maximum sleeps in a BS based on the Fig. 11 and the existing figure. We can infer from the data that, in contrast to the previous study, there is a significant decrease in $E(PC)$ and $\mathcal{PS}$ with an increase in SM1 rate

which is presented in Fig. 11a, b respectively. Moreover, $E(PC)$ likewise falls when the rate of SM2 rises as shown in Fig. 11c. Nevertheless, $\mathcal{PS}$ gradually rises after declining first presented in Fig. 11d. This occurs because a SM2 (long sleep) duration results in a higher power saving than SM1 (shorter sleep). Therefore, it turns out that the suggested approach will be more beneficial for energy conservation in the actual 5G networks based on the aforementioned observation and calculation.
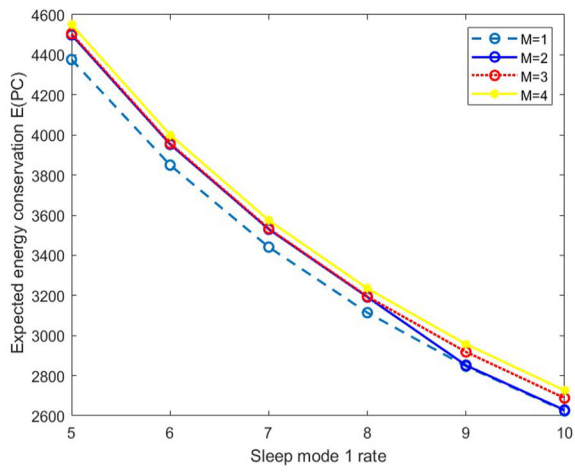
(a) SM1 rate vs expected energy conservation $E(PC)$
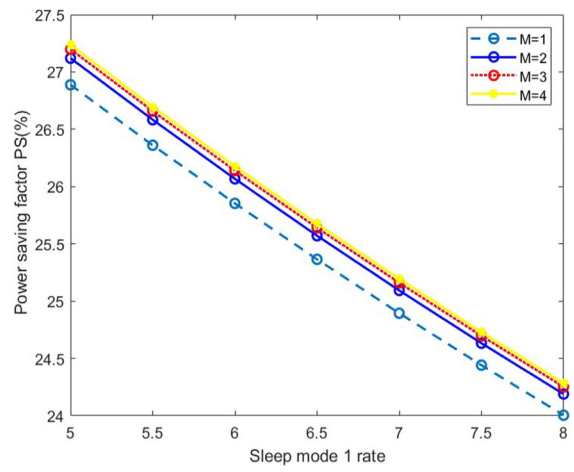
(b) SM1 rate vs power saving factor $\mathcal{PS}$
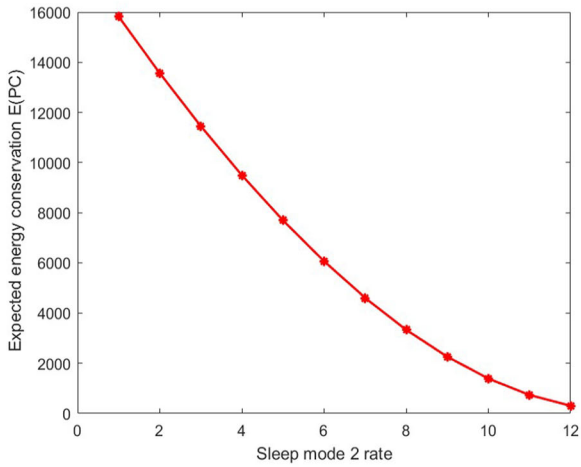
(c) SM2 rate vs expected energy conservation $E(PC)$
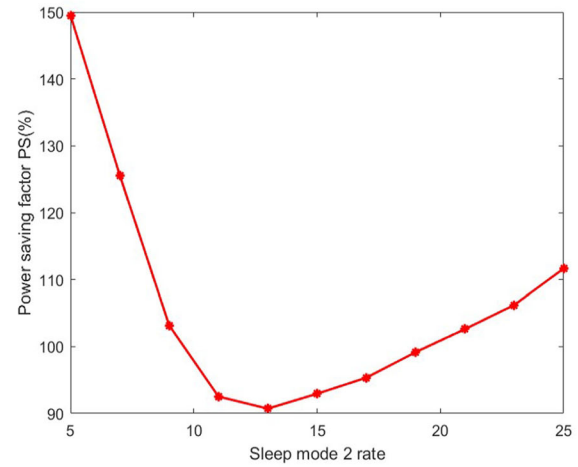
(d) SM2 rate vs power saving factor $\mathcal{PS}$
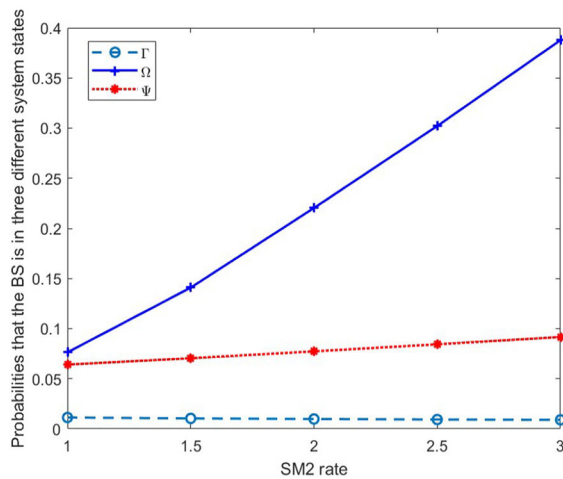
(e) SM2 vs three different system state probabilities

Fig. 11 Implications of SM1 at various maximum sleeps and SM2
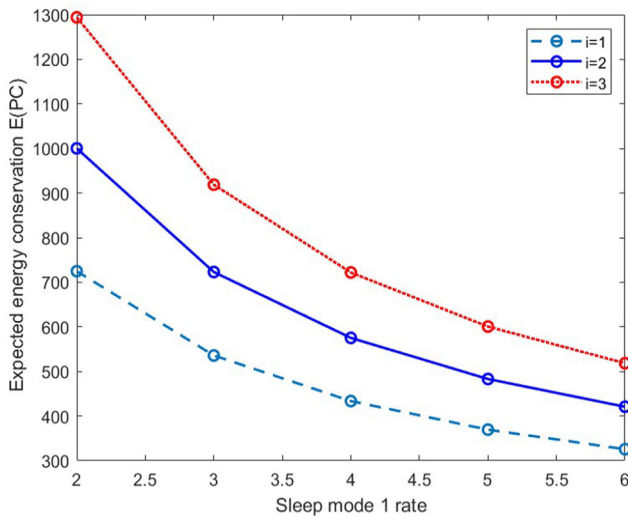
(a) SM1 rate vs expected energy conservation E(PC)
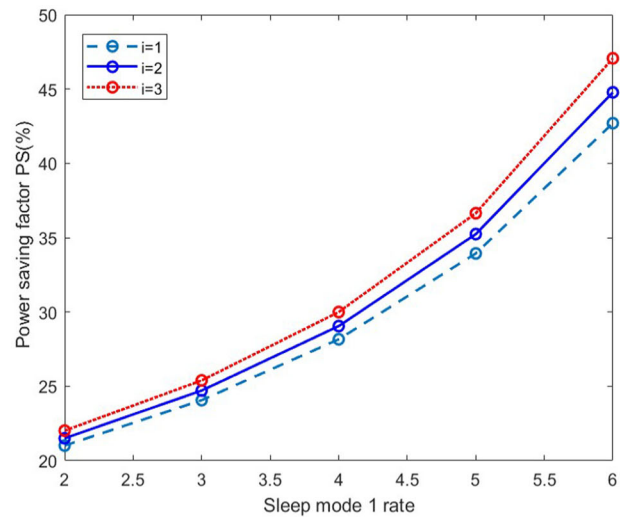
(b) SM1 rate vs power saving factor $\mathcal{PS}$
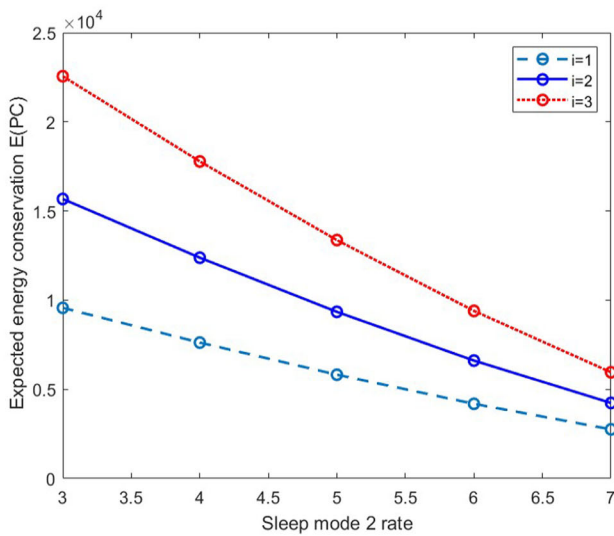
(c) SM2 rate vs expected energy conservation E(PC)
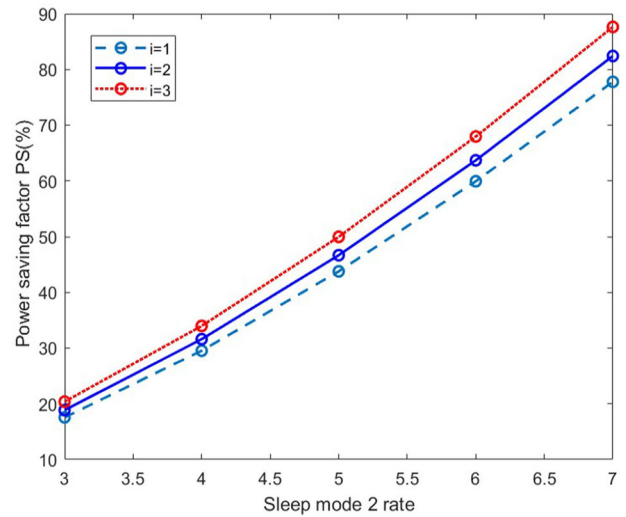
(d) SM2 rate vs power saving factor $\mathcal{PS}$

Fig. 12 Implications of sleep modes at various setup rate

7.6 Enhancing the efficiency of E(PC) and E(L)

The percentage of fluctuation (PF), as shown in Table 10, is computed using the formula below,

$$PF = \frac{\text{Measure without N-policy} - \text{Measure with N-policy}}{\text{Measure without N-policy}}$$

We know that, increasing the threshold value, raises the E(L), however decreases the E(PC). Thus, to strike a balance between these factors - E(PC) and E(L) - we assess the

percentage of fluctuation. This metric reveals the percentage difference between the energy consumption/delay with and without the N-policy. Higher N values tend to elevate the mean delay while minimizing the anticipated power usage. Therefore, selecting an appropriate value for N effectively reduces power usage while maintaining a reasonable latency. By implementing the proposed sleep and wake schedule results in at least an 16.5% reduction in energy usage.
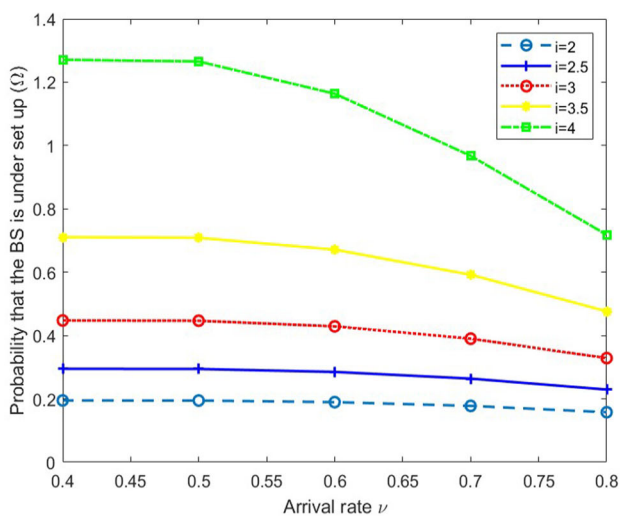
(a) Arrival rate vs probability that the BS under setup state
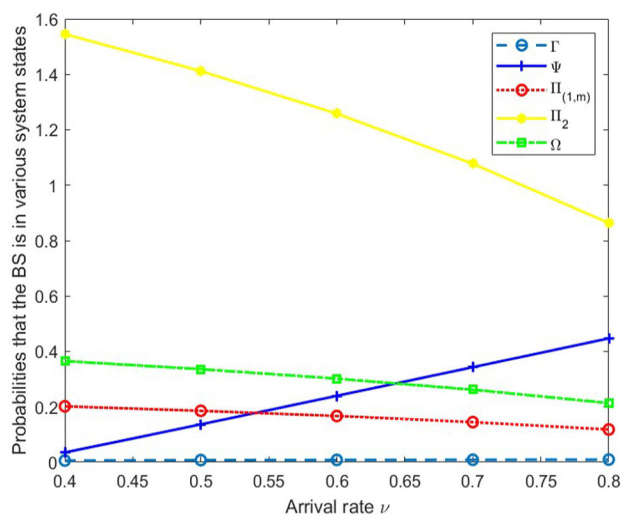
(b) Arrival rate vs different system state probabilities

Fig. 13 Implication of arrival rate ν

Table 10 Impact of N on E(L) and E(PC)

Parameters	N	E(PC)		E(L)	
		$\overline{E(L)}$	% of fluctuation	$\overline{E(PC)}$	% of fluctuation
Sleep without N policy	1	28.6390	22.2	637.4583	16.5
	2	31.3168	28.7	876.5941	21.7
	3	34.0007	35.4	1129.4655	27.2
Sleep with N policy	4	36.6847	41.1	1396.2974	33.1
	5	39.3686	49.2	1677.2982	42.3
	6	42.0525	58.5	1972.6581	51.2

8 Cost optimization: a perspective

Assuming a model has been built and is poised to go into functioning, the initial query a manufacturer or analyst will be curious about is whether it is cost-effective. Therefore, "cost analysis"is a key consideration in determining a model's adequacy. System designers are aided in their decision-making, and risk is mitigated as a result. Considering the importance and pressing requirement for cost minimization in this study, we implemented three distinct optimization methods namely particle swarm optimization (PSO), artificial bee colony (ABC) and genetic algorithm (GA). Since cost is the most vital variable in establishing economic interpretation in a system, we have provided a cost analysis that will help us avoid this issue. First, a workable cost function is constructed, and then the optimal parameters are determined using the above acceptable optimization method. Cost-cutting strategies such as those listed above have been widely used by researchers [23, 26, 34] in recent years.

8.1 Cost analysis

The cost optimization approach is employed to estimate the optimal parameters for the service rates (τ_b, τ_1). It presumes that the estimated cost function has a linear form in terms of cost components tied to diverse system activities.
For per unit of time (PUT), the succeeding cost components are included in the expected total cost function TC (τ_b, τ_1):

$\mathcal{D}_h$ Each CR's holding cost per unit time spent within the BS
$\mathcal{D}_b$ Cost per unit time provided the BS is normally active
$\mathcal{D}_v$ Cost per unit time provided the BS is under vacation (both sleep modes)
$\mathcal{D}_f$ Cost per unit time provided the BS is under repair
$\mathcal{D}_1$ Cost per CR served during busy hours
$\mathcal{D}_2$ Cost per CR served during BS's sleep modes

Therefore, the expected cost function is defined as,

$$TC = \mathcal{D}_h L_q + \mathcal{D}_b \Gamma + \mathcal{D}_v (\Pi_{1,m} + \Pi_2) + \mathcal{D}_f \mathcal{R}$$
$$+ \mathcal{D}_1 \tau_b + \mathcal{D}_2 \tau_1 \qquad (56)$$

The difficulty in analytically optimizing the cost function described by (56) stems from its considerable non-linearity. Thus, we infer that total cost is a function of service rates (τ_b, τ_1) and employ a heuristic method to optimize it.

8.2 Cost optimization

For any given objective function, "optimization" refers to the process of discovering its "optimal" or "minimal" collection of inputs. Cost optimization is a firm-centric, iterative process that seeks to lessen outlays and boost profitability. Our primary goal is to minimise the total cost function to find out what service rates τ_b^* and τ_1^* should be set at for the server's busy mode and its vacation mode, respectively.

The cost-minimization problem can be expressed mathematically as:

$$TC(\tau_b^*, \tau_1^*) = \underset{\tau_b^*, \tau_1^*}{Minimize} TC(\tau_b, \tau_1)$$

Table 11 contains the values of the cost elements used to conduct a graphical analysis of the cost function's sensitivity.

Plenty of optimization strategies have been constructed since the 1960s. Each of these algorithms has proven its mettle in implementing various optimization problems. We ought to employ a local optimization strategy when we have an inkling that we are close to the global optimum, or when our objective function has a single optimum, as in the case of a unimodal distribution; we should use a global optimization method when we have a poor idea of the structure of the objective function, or when the function has multiple local optimums. When a situation calls for a nationwide search algorithm, a local search might be tricked by local optima into producing subpar outcomes. We conducted this research using certain global search optimization approaches like PSO, ABC, GA given that we acknowledge the necessity and virtue of cost.

Each one of these three algorithms has advantages and disadvantages of its own. Compared to certain other optimization techniques, PSO, ABC, and GA are comparatively easy to implement. They are usable by users of different skill levels due to their simple ideas and intuitive parameters. They don't require much changes to be applied to a variety of optimization situations. They are adaptable and have been effectively applied in an array of fields, notably biology, machine learning, technology, and business. PSO draws inspiration from the social interactions of fish and birds, ABC from honeybees' foraging habits, and GA from genetics and the process of natural selection. These algorithms' biological roots frequently result in sensible exploration-exploitation

Table 11 Different cost sets for cost analysis

Cost sets	$\mathcal{D}_h$	$\mathcal{D}_b$	$\mathcal{D}_v$	$\mathcal{D}_f$	$\mathcal{D}_1$	$\mathcal{D}_2$	$\mathcal{D}_3$
Set 1	15	95	35	20	15	12	10
Set 2	20	75	45	10	25	15	12
Set 3	25	65	55	5	30	20	17

trade-offs when looking for optimal solutions. It's crucial to remember that no optimization technique is inherently better than any other, even though PSO, ABC, and GA all offer benefits. The particulars of the problem, the computational power at hand, and the method's user-friendliness all play a role in the selection of an algorithm.

8.3 Particle swarm optimization (PSO)

In 1995, Kennedy and Eberhart [21] introduced the world to PSO, a technique for optimizing continuous non-linear functions that was driven by initial research on the behaviour of bird flocks as simulated in computer programmes. Personal bests (pbest) and global bests (gbest) are obtained by assigning a fitness value to each particle. After then, the 'pbest' value of a particle is utilised to replace the 'gbest' value if it is more precise. Until the optimum amount of iterations has been achieved, this procedure is repeated. A substantial quantity of power is used by 5G BS. Radio transmitters and processors are a couple of base station components whose power consumption can be optimized with the use of PSO. PSO can assist in lowering the consumption of energy while preserving network performance by modifying parameters like transmission power and duty cycles.

Upadhyaya [33] elaborated on this method to talk about cost optimization in a discrete-time retrial queue with Bernoulli feedback and starting failure. For further information on how PSO works, the study by Malik et al. [26] has been referred to. Algorithm 1 lays out the PSO algorithm's pseudo coded sequence of operations. In addition, Table 12 details the effects of v, r, φ on total minimal cost TC using PSO optimization approaches, on the optimal service rate pairs (τ_b*, τ_1*).

8.4 Artificial bee colony optimization (ABC)

ABC technique was proposed by Karaboga and Basturk [20], and it is a swarm-based solution for optimization issues that takes its cues from the smart conduct of honey bees during foraging. There are two main parts to the algorithm: " foraging "and " food source."Foraging bees might be classified as workers, observers, or scouts, depending on their roles in the current situation. Both working and jobless bees use foraging to find good food sources. In this algorithm (a good answer), a

Table 12 Effect of ν, r, φ on (TC, τ_b^*, τ_1^*) using PSO

Parameters	(TC, τ_b^*, τ_1^*)			
		Cost set 1	Cost set 2	Cost set 3
ν	0.52	(2.3543,1.8555,210.7060)	(1.9452,2.3343,276.2257)	(1.5449,3.1892,343.1923)
	0.53	(2.3792,1.7524,210.5543)	(1.9585,2.2353,276.2904)	(1.5467,3.0631,343.4782)
	0.54	(2.4096,1.6281,210.3284)	(1.9790,2.1699,276.2596)	(1.5497,2.8620,343.6415)
r	0.3	(2.3715,1.6953,209.7231)	(1.9833,2.1181,272.8109)	(1.6246,2.6779,344.9708)
	0.4	(2.3560,1.7432,209.4125)	(1.9590,2.1661,276.2585)	(1.5526,2.9239,344.3083)
	0.5	(2.3306,1.7595,209.0943)	(1.9531,2.1583,275.6982)	(1.5598,2.8132,343.6436)
φ	8.5	(2.3258,1.7700,209.0945)	(1.9809,2.1363,276.8108)	(1.6242,2.8420,344.9317)
	8.6	(2.3276,1.8896,210.6988)	(1.9866,2.2412,278.8914)	(1.6038,2.8928,347.0337)
	8.7	(2.3381,1.9831,212.2810)	(2.9895,2.3275,279.9473)	(1.6384,3.0862,349.1272)

Algorithm 1 Pseudo Code for PSO Algorithm

INPUT: Objective function $=TC(\tau_b, \tau_1)$,
OUTPUT: The cost function's value $TC(\tau_b^*, \tau_1^*)$
Initialization
for each particle j=1,2,…N, **do**
    Start with a uniform distribution of particle positions as $M_j(0) \sim (LB, UB)$, in which LB and UB are the lower and upper limits of the search area.
    Setup abest to its starting position abest$(j, 0)= M_j(0.)$
    Initialize cbest to the minimal value of the swarm:$cbest(0) = argmin\ f[M_j(0).]$
end for
Initiate the velocity: $V_j \sim U(-|UB - LB|, |UB - LB|)$.
Iterate until the endpoint is reached
for each particle j=1,2,…N **do**
    Choose arbitrary numbers:$s_1, s_2 \sim U(0, 1)$.
    Enhance the particle's velocity.
    Enhance the particle's position.
    Suppose $[M_j(0).]\leq [abest(j, t)(0)]$
    Update the particle's current predicted position. $j : abest(j, t) = M_j(t)$.
    Suppose $[M_j(0)]\leq [cbest(t)(0)]$, enhance the swarm's best known position: $j : abest(t) = M_j(t)$. t ←(t+1;)
end for
$cbest(t)$ holds the optimal found solution.

Moreover, the impacts of ν, r, φ on optimal service rate pairs (τ_b^*, τ_1^*), respectively, as well as total minimal cost TC via ABC optimization methods are laid out in Table 13.

Algorithm 2 Pseudo Code for ABC Algorithm

INPUT: Objective function $=TC(\tau_b, \tau_1)$,
OUTPUT: The cost function's value $TC(\tau_b^*, \tau_1^*)$
Generate an assembly of outcome $M_i, i = 1$
Check out the assembly, period 1, $h = 0$
Choose the optimal outcome, M best and fix M best1 $= M$ best
Redo
Obtain a new way out Mnew $= M_i$ for the worker bees and to get them.
Use the greedy selection method for the worker bees.
Give each result a grade, and then choose the best one.
Find the probability P_i of the solution M_i.
With P_1 as a starting point, generate a fresh outcome M_i for the onlookers.
Use the greedy selection method to the onlookers.
If the scout's result has been cancelled, then proceed with a newly created result M_i.
Keep in mind the M new conclusion you've reached thus far.
Put $h = h + 1$ period = period + 1.
Until (the end condition has been met, i.e., period = MCN)

colony of robotic forager bees searches for plentiful synthetic food sources. In order to optimize the objective function with this method, the correct parameter vector must be used. Then, the robotic bees find a random distribution of starting solution vectors. The closest neighbour search method is employed during the iterative process of reducing the error. ABC can be utilized to maximize the choice of locations for the installation of 5G BS. ABC can assist in determining the best sites for BS to save deployment costs by taking into account variables such as population density, signal coverage needs, and the expense of infrastructure implementation. The findings indicate that ABC is either superior to, equivalent to, or offers the advantage of fewer control factors than some of the other population approaches. The phases of the ABC method are described in pseudocode in Algorithm 2.

8.5 Genetic algorithm (GA)

During the 1960 s and 1970 s, Bremermann, Holland [4, 11] and their colleagues developed the GA, an approach for solving optimization issues inspired by natural selection, the process through which organisms evolve. They come in handy for effective remedies for unreliable search queries. The full algorithm depicts the parameters incorporated to pick the healthiest and most successful people to have children. Natural selection is simulated by genetic algorithms, which favour organisms with the ability to survive, reproduce, and pass its traits on to offspring. In order to find a remedy to a problem, researchers replicate " the survival of the fittest "through multiple human generations. In addition,

Table 13 Effect of ν, r, φ on (TC, τ_b^*, τ_1^*) using ABC

Parameters	(TC, τ_b^*, τ_1^*)			
		Cost set 1	Cost set 2	Cost set 3
ν	0.52	(2.7571,2.8115,273.6573)	(2.6282,3.1965,376.1871)	(2.5425,3.594,467.7676)
	0.53	(2.7476,2.7313,272.9675)	(2.5826,3.1089,375.9223)	(2.5394,2.9736,467.4577)
	0.54	(2.7208,2.6261,272.1762)	(2.5348,3.0374,375.4594)	(2.5297,2.9190,466.9847)
r	0.3	(2.8241,2.6030,275.8688)	(2.8104,3.0323,378.2474)	(2.4807,3.0349,469.9546)
	0.4	(2.8031,2.6469,275.5177)	(2.6337,3.2465,377.1020)	(2.5352,3.0774,468.8927)
	0.5	(2.7895,2.6791,275.1595)	(2.5887,3.2744,376.3267)	(2.5536,3.2535,468.0455)
φ	8.5	(2.7281,2.8117,273.6499)	(2.7577,2.9492,376.6902)	(2.5093,2.9237,462.1726)
	8.6	(2.7488,2.8852,275.9271)	(2.6884,3.3789,379.0621)	(2.5308,2.9902,466.0346)
	8.7	(2.7701,2.9570,278.1969)	(2.6518,3.4546,382.1096)	(2.5510,3.0522,469.8845)

Arqub [1] has detailed how a continuous genetic algorithm was used in his study. Research analysts are drawn to the aforementioned method due to its potential value in a broad variety of areas, including data centres, cryptanalysis, electrical circuit design, etc. Moreover, the deployment of 5G base stations is limited by things like financial constraints, technological limits, and regulatory regulations. GA can be modified to efficiently manage these limitations, guaranteeing that the solutions produced are workable and meet the required standards. Algorithm 3 provides the pseudocode for the GA algorithm's order of procedures. Moreover, the influence of ν, r, φ on optimal service rate pairs (τ_b^*, τ_1^*), respectively, as well as total minimal cost TC via GA optimization methods, are laid out in Table 14.

Algorithm 3 Pseudo Code for GA Algorithm

INPUT: Objective function $= TC(\tau_b, \tau_1)$,
OUTPUT: The cost function's value $TC(\tau_b^*, \tau_1^*)$
Calculate the objective function (OF)
Set the number of generations to 0 $(t = 0)$
Generate an initial sample of users at random $P(t)$
Evaluate an assembly of users by employing OF
while termination criterion is not satisfied **do**
　　t=t+1
　　Choose the users to assembly $P(t)$ from $P(t-1)$
　　Alter users of $P(t)$ by employing crossover and mutation
　　Evaluate an assembly of users by employing OF
end while
Return the best user found during the evaluation

At the outset of any optimization method- PSO, ABC or GA-the particles are not stable, and it is vital to know whether or not they return to normal and whether or not they roam around in search of a better solution. That's why convergence is so important to cost analysis. Here Fig. 15 depicts the convergence of the cost function via PSO, ABC and GA techniques. In addition, by employing the aforementioned techniques, we have been able to arrive at combined optimum

values, which offer the lowest predicted costs. Moreover, we have made an effort to investigate the cost model through the use of 3D graphs, which is presented in Fig. 14.

8.6 Comparative analysis within PSO, ABC, and GA

Here we examine three methods, namely PSO, ABC and GA to estimate the lowest possible cost using their respective MATLAB programmes. In this case, we analyse three distinct sets of costs, depicted in Table 11. The MATLAB programmes for each of these algorithms are then executed in turn. As a result of carrying this approach forward, Tables 12, 13, and 14 have been created.

There was considerable similarity between the outcomes of the three programmes, but there was also a bit of deviation. Thus, the optimal solutions and associated lowest costs of those three strategies are adjacent to one another. This proves that the aforementioned heuristics provide trustworthy, optimal solutions. Based on the data shown in Tables 12, 13, and 14, PSO has the lowest ideal cost value. Thus, we can use any strategy to determine the ideal cost; nevertheless, as we juxtapose for the proposed framework, PSO is a highly profitable approach for determining the best feasible cost. It is simple to configure, performs admirably in global queries, and is insensitive to scaling changes in design variables. PSO tends to lead to a slow and quick convergence in mid-optimal locations, as well as a ponderous convergence in a widened search domain.

8.7 Convergence in PSO, ABC and GA

Knowing whether or not a particle lays off and how it will wander over in search of a more effective outcome is actually crucial when applying optimization techniques like PSO, ABC, or GA, because the constituent parts (people, groups, or signals) are inherently out of equilibrium. Hence, cost analysis convergence is crucial here since it eliminates all


**Table 14** Effect of ν, r, φ on (TC, τ_b^*, τ_1^*) using GA

Parameters	(TC, τ_b^*, τ_1^*)			
		Cost set 1	Cost set 2	Cost set 3
ν	0.52	(2.8614,3.0806,322.5623)	(2.4458,3.2909,429.3605)	(2.5115,3.0575,535.6779)
	0.53	(2.8760,3.0591,322.7222)	(2.4596,3.2713,429.6836)	(2.5234,3.8366,535.8236)
	0.54	(2.8897,3.0370,322.7863)	(2.4726,3.2512,429.8757)	(2.5347,3.0152,535.8118)
r	0.3	(2.9027,3.0145,322.7551)	(2.5302,3.2005,432.4459)	(2.4988,3.0779,535.3740)
	0.4	(2.8802,3.0285,321.9419)	(2.5079,3.2154,431.1988)	(2.4737,3.0911,533.8794)
	0.5	(2.8572,3.0427,321.1204)	(2.4850,3.2306,429.9371)	(2.4480,3.1044,532.3252)
φ	8.5	(2.8167,3.0112,318.0792)	(2.4602,3.1671,423.6458)	(2.4480,3.1044,532.3252)
	8.6	(2.8378,3.0609,321.1670)	(2.4792,3.2160,427.5879)	(2.4672,3.1527,537.2758)
	8.7	(2.8589,3.1106,324.2572)	(2.4980,3.2650,431.5319)	(2.4863,3.2010,542.2288)
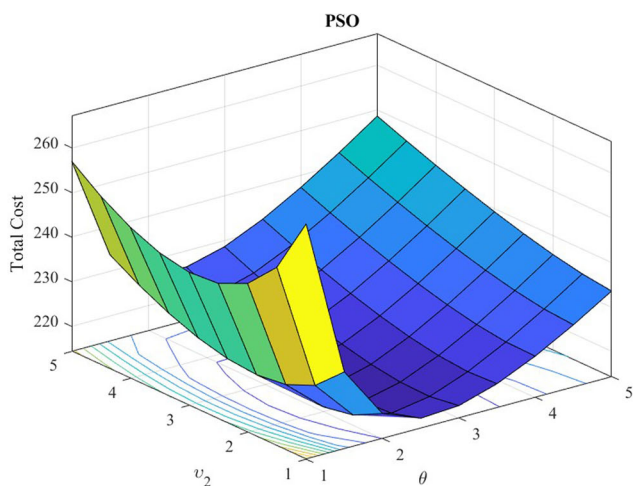
(a) θ, υ_2 vs TC in PSO
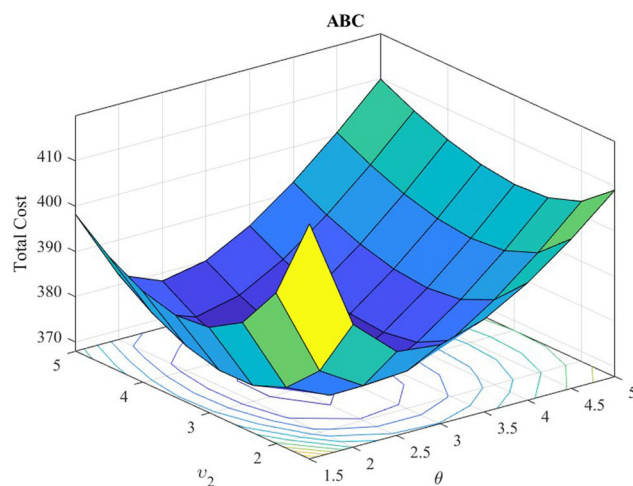
(b) θ, υ_2 vs TC in ABC
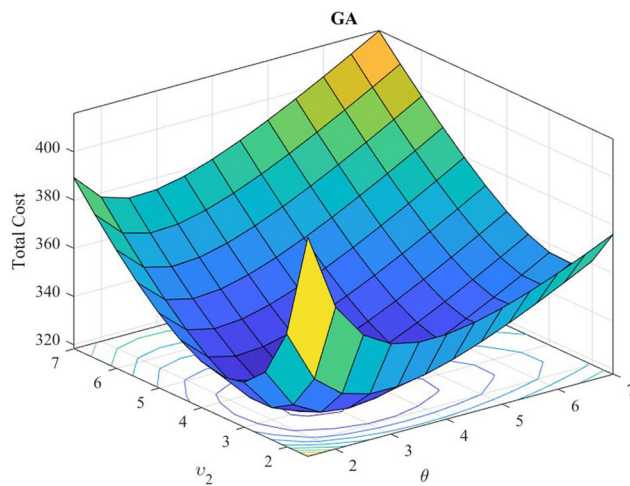
(c) θ, υ_2 vs TC in GA

Fig. 14 Optimality of the cost function

of these challenges. Tables 12, 13, and 14 display the typical outcomes of applying the methods PSO, ABC, and GA to every one of the cost sets listed in Table 11. Further, as shown in Fig. 15, particles converge to the optimal solution following certain trials (generations) in PSO, ABC, and GA. After a brief amount of instances (generations), the particles in all three methods converge on the minimum possible cost. Overall, the cost study above suggests that PSO converges faster. We acknowledge that the overall estimated cost is sensitive to the parameters chosen for priority, voluntary service, and vacation, and that admins should exercise great caution when making these choices. According to the results of the analysis, the model we have outlined works very well with the 5G BS model. The analysts can afford this system in its entirety, which will help alleviate some of their financial strain. Besides demonstrating the reliability of our model, the cost analysis's close resemblance to real-world examples helps system designers and analysts minimise problems.

9 Conclusion

By simulating the 5G BS sleeping mechanism, we were able to lower the 5G BS's average power usage and achieve the greatest possible power savings. In order to develop a better-performing model that can be employed in real time systems, we blended the 5G BS with the concepts of

- Retrial queue
- Feedback
- Repair

simultaneously which is worth emphasizing. Combining a sleeping strategy with an N policy and a retrial queuing framework is a comprehensive and very successful way to address issues with energy usage in 5G BSs. When the sleeping strategy and N policy work together, BSs are able to dynamically switch between sleep and active modes according to the arrival of the CR, which maximizes energy efficiency without sacrificing acceptable service levels. The useful analytical results for 5G BS are validated by the numerical examples. Moreover, cost minimization techniques such as PSO, ABC, and GA are also employed. The outcomes are evaluated across methods, and the lowest possible price is determined.

Future research could expand on this study by charactering various sleeping strategies. Moreover, the proposed QS can also be analysed with 6G network BS. The research could potentially be extended to discrete systems as well. Further, the matrix geometric method was not applied to the study of these systems, which is also a promising avenue for future investigation.

Appendix A

In the case where $\rho < 1$, the embedded Markov chain (MC) $\{F_n; n \in N\}$ can be said to be ergodic, where $\rho = \mathcal{E}(X_1)[1 - \mathcal{Q}^*(\nu)] + r - \nu\mathcal{E}(X_1)[1 + \eta\mathcal{E}(S^1)]\mathcal{E}(L^1)$

Proof It is straightforward to verify that ergodicity is a sufficient condition by applying Foster's criteria [29], which points out that the chain might be irreducible and aperiodic. If a non-negative fn. $c(d)$, $d \in N$ and $\epsilon > 0$, exists that guarantees that the mean drift $\varsigma_d = H[c(f_{n+1}) - c(f_n)/f_n = d]$ is limited for all $d \in N$ and $d \in N$ and $\varsigma_d \leq -\epsilon$, excluding possibly for a finite no. of d's, then the MC is ergodic. When considering the function $c(d) = d$ in our scenario, we get

$$\varsigma_d = \begin{cases} r - \nu\mathcal{E}(X_1)[1 + \eta\mathcal{E}(S^1)]\mathcal{E}(L^1), & \text{if } c = 0 \\ \mathcal{E}(X_1)[1 - \mathcal{Q}^*(\nu)] + r - \nu\mathcal{E}(X_1)[1 + \eta\mathcal{E}(S^1)] \\ \quad \mathcal{E}(L^1) - 1, & \text{if } c = 1, 2, \ldots \end{cases}$$

It is clear that ergodicity must exist in order for the inequality below to exist.

$$\mathcal{E}(X_1)[1 - \mathcal{Q}^*(\nu)] + r - \nu\mathcal{E}(X_1)[1 + \eta\mathcal{E}(S^1)]\mathcal{E}(L^1) < 1$$

If the MC $\{F_n; n\epsilon N\}$ meets Kaplan's criterion, we may easily ensure non-ergodicity in accordance with Sennott et al. [31] particularly $\varsigma_d < \infty$ for all $d \geq 0$ and $\exists d_0 \in N$ s.t $\varsigma_d \geq 0$ for $d \geq d_0$. The fact that "Kaplan's condition" is met in our scenario is underscored by the existence of a w such that $g_{ld} = 0$ for $d < l - m$ and $l > 0$, in which g_{ld} is the one-step transition matrix of $\{F_n; n \in N\}$. Consequently, it is indicated that the MC is non-ergodic by

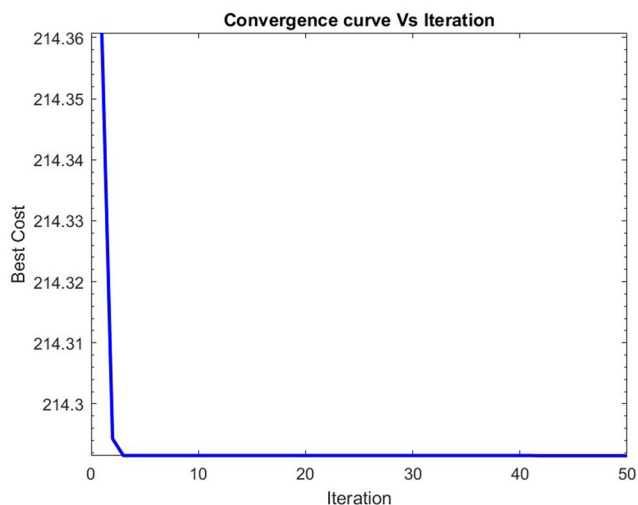$$\mathcal{E}(X_1)[1 - \mathcal{Q}^*(\nu)] + r - \nu\mathcal{E}(X_1)[1 + \eta\mathcal{E}(S^1)]\mathcal{E}(L^1) \geq 1$$
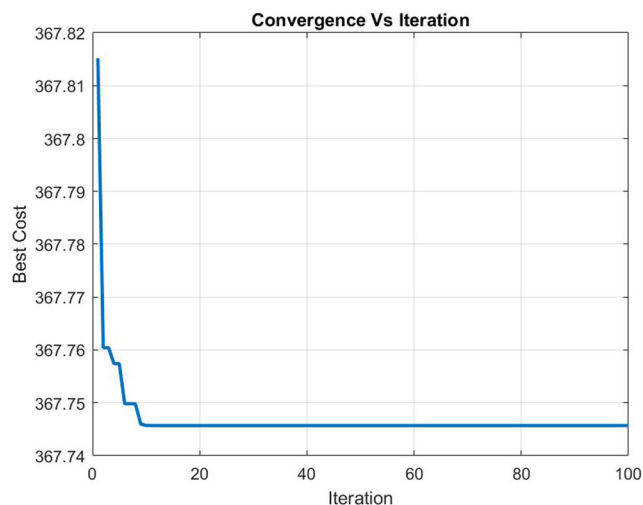
□

Appendix B

By substituting the eqns. (35), (36) and (33) in (26) and performing some calculations, we eventually arrive at,

$$\Psi(0, \breve{\varepsilon}) = \frac{\nu\mathcal{P}_0 \{\breve{\varepsilon}(1 - M) - (1 - r + r\breve{\varepsilon})\mathcal{L}^*(\mathcal{T}(\breve{\varepsilon}))[\mathcal{A}(\breve{\varepsilon})) - \breve{\varepsilon}(1 - \mathcal{A}(\breve{\varepsilon})) + \breve{\varepsilon}U(1 + s\mathcal{G}_2^*(\mathcal{C}(\breve{\varepsilon})))\mathcal{B}^*(\mathcal{C}(\breve{\varepsilon}))]\}}{(1 - r + r\breve{\varepsilon})\mathcal{L}^*(\mathcal{T}(\breve{\varepsilon}))[\mathcal{Q}^*(\nu) + \mathcal{A}(\breve{\varepsilon})(1 - \mathcal{Q}^*(\nu))] - \breve{\varepsilon}}$$
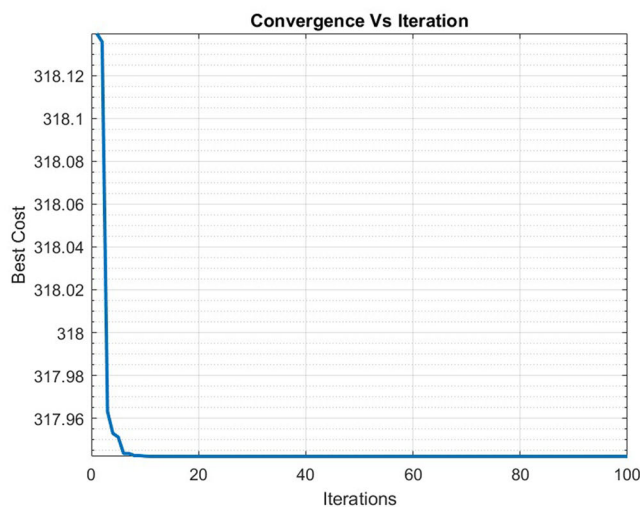
(57)

Here, by substituting the eqns. (35) to (39) into the eqns. (27) to (31), we get,

(a) Convergence in PSO

(b) Convergence in ABC

(c) Convergence in GA

Fig. 15 Convergence of the cost function

$$\Gamma(0,\breve{\varepsilon}) = \frac{\nu\mathcal{P}_0\left\{(1-M)[\mathcal{Q}^*(\nu)+\mathcal{A}(\breve{\varepsilon}))(1-\mathcal{Q}^*(\nu))-\breve{\varepsilon}]-[\mathcal{A}(\breve{\varepsilon}))-\breve{\varepsilon}(1-\mathcal{A}(\breve{\varepsilon}))) \\ +\breve{\varepsilon}U(1+s\mathcal{G}_2^*(\mathcal{C}(\breve{\varepsilon})))\mathcal{B}^*(\mathcal{C}(\breve{\varepsilon}))]\right\}}{(1-r+r\breve{\varepsilon})\mathcal{L}^*(\mathcal{T}(\breve{\varepsilon}))[\mathcal{Q}^*(\nu)+\mathcal{A}(\breve{\varepsilon})(1-\mathcal{Q}^*(\nu))]-\breve{\varepsilon}} \tag{58}$$

$$\Pi_{1,m}(0,\breve{\varepsilon}) = p_m \frac{\nu\mathcal{P}_0}{\mathcal{G}_{1,0}^*(\nu)}\mathcal{G}_{1,0}^*(\mathcal{C}(\breve{\varepsilon}))); \; (m=1,2,\ldots M) \tag{59}$$

$$\Pi_2(0,\breve{\varepsilon}) = s\left[\frac{\nu\mathcal{P}_0}{\mathcal{G}_{1,0}^*(\nu)}\mathcal{G}_{1,0}^*(\mathcal{C}(\breve{\varepsilon})))\left[p_0+\sum_{m=1}^M p_m \mathcal{G}_{1,m}^*(\mathcal{C}(\breve{\varepsilon})))\right]\right] \tag{60}$$

$$\Omega(0,\breve{\varepsilon}) = \left[\frac{\nu\mathcal{P}_0}{\mathcal{G}_{1,0}^*(\nu)}\mathcal{G}_{1,0}^*(\mathcal{C}(\breve{\varepsilon})))\left[p_0 + \sum_{m=1}^M p_m \mathcal{G}_{1,m}^*(\mathcal{C}(\breve{\varepsilon})))\right]\right][1+s\mathcal{G}_2^*(\mathcal{C}(\breve{\varepsilon})))] \tag{61}$$

$$\mathcal{R}(0,\breve{\varepsilon}) = \eta\Gamma(0,\breve{\varepsilon})[1-\mathcal{L}(\hat{\varpi})]e^{-(\mathcal{T}(\breve{\varepsilon})\hat{\varpi})} \tag{62}$$

Appendix C

$$Nr_q'''(1) = -6v\mathcal{E}(X_1)(1 + s\mathcal{E}(S^1))[r + \mathcal{E}(X_1)(1 - \mathcal{Q}^*(v))$$
$$- v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))\mathcal{E}(L^1)$$
$$- 1][M' - v\mathcal{E}(X_1) + \frac{v}{\mathcal{G}_{1,0}^*(v)}[v\mathcal{E}(X_1)(1 + s)\mathcal{E}(B^1)$$
$$+ vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)]]$$
$$- 9M'v\mathcal{E}(X_1)(1 + s\mathcal{E}(S^1))v\mathcal{E}(X_1)(1 - \mathcal{Q}^*(v))$$
$$+ 2v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))\mathcal{E}(L^1)[-v\mathcal{E}(X_1)$$
$$+ \eta\mathcal{E}(X_1)\mathcal{E}(S^1)]] - 9[2\mathcal{E}(X_1) + U'(1 + s)$$
$$+ \frac{(1 + s)}{\mathcal{G}_{1,0}^*(v)} - \frac{v}{\mathcal{G}_{1,0}^*(v)}[v\mathcal{E}(X_1)(1 + s)\mathcal{E}(B^1)$$
$$+ vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)]][2(1 - \mathcal{Q}^*(v))v\mathcal{E}(X_1)(1$$
$$+ s\mathcal{E}(S^1))v\mathcal{E}(X_1) + v[2\mathcal{E}(X_1)(1$$
$$+ s\mathcal{E}(S^1))v\mathcal{E}(L^1)[-v\mathcal{E}(X_1) + \eta\mathcal{E}(X_1)\mathcal{E}(S^1)]]]$$

$$Nr_q''''(1) = -6[(v\mathcal{E}(X_1))^2(1 + s\mathcal{E}(S^1))][(1 - \mathcal{Q}^*(v))[M''$$
$$- 4M' + 4r[2\mathcal{E}(X_1) + U'(1 + s) + \frac{v}{\mathcal{G}_{1,0}^*(v)}(1 + s)$$
$$- \frac{v}{\mathcal{G}_{1,0}^*(v)}[v(1 + s)\mathcal{E}(X_1)\mathcal{E}(B^1)$$
$$+ vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)]]] - 12[2\mathcal{E}(X_2)$$
$$+ 2\mathcal{E}(X_1) + U''(1 + s) - 2[v(1 + s)\mathcal{E}(X_1)\mathcal{E}(B^1)$$
$$+ vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)][U' + \frac{v}{\mathcal{G}_{1,0}^*(v)}] + 2U'(1 + s)$$
$$- \frac{v}{\mathcal{G}_{1,0}^*(v)}[(1 + s)\mathcal{B}^{*''}(1) + s\mathcal{G}_2^{*''}(1)]$$
$$+ 2\frac{v}{\mathcal{G}_{1,0}^*(v)}v(1 + s)\mathcal{E}(X_1)\mathcal{E}(B^1)vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)]]$$
$$- 12v[M' + [2\mathcal{E}(X_1) + U'(1 + s) + \frac{v}{\mathcal{G}_{1,0}^*(v)}(1 + s)$$
$$- \frac{v}{\mathcal{G}_{1,0}^*(v)}[v(1 + s)\mathcal{E}(X_1)\mathcal{E}(B^1) + vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)]]]$$
$$[(1 - \mathcal{Q}^*(v))\mathcal{E}(X_1)\mathcal{T}''(1)$$
$$+ \mathcal{L}^{*''}(1)[-v\mathcal{E}(X_1) + \eta v\mathcal{E}(X_1)\mathcal{E}(S^1)]]$$
$$- 12[M' + v[2\mathcal{E}(X_1) + U'(1 + s)(U' + \frac{v}{\mathcal{G}_{1,0}^*(v)})$$
$$- \frac{v}{\mathcal{G}_{1,0}^*(v)}[v(1 + s)\mathcal{E}(X_1)\mathcal{E}(B^1)$$
$$+ vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)]]]v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))(\mathcal{E}(L^1)$$
$$+ (1 - \mathcal{Q}^*(v))v\mathcal{E}(X_2))$$
$$- 12v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))\mathcal{E}(L^1)[[-v\mathcal{E}(X_1)$$
$$+ \eta v\mathcal{E}(X_1)\mathcal{E}(S^1)][v[2\mathcal{E}(X_2)$$
$$+ 2\mathcal{E}(X_1) + U''(1 + s) - 2[v(1 + s)\mathcal{E}(X_1)\mathcal{E}(B^1)$$
$$+ vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)][U' + \frac{v}{\mathcal{G}_{1,0}^*(v)}]$$
$$+ 2U'(1 + s) - \frac{v}{\mathcal{G}_{1,0}^*(v)}[(1 + s)\mathcal{B}^{*''}(1) + s\mathcal{G}_2^{*''}(1)]$$
$$+ 2\frac{v}{\mathcal{G}_{1,0}^*(v)}v(1 + s)\mathcal{E}(X_1)\mathcal{E}(B^1)vs$$
$$\mathcal{E}(X_1)\mathcal{E}(G_2^1)]$$

$$+ (1 - \mathcal{Q}^*(v))[M''v - 12M'v\mathcal{E}(X_1)]]$$
$$+ 12(1 - \mathcal{Q}^*(v))(1 + \eta\mathcal{E}(S^1))(\mathcal{E}(X_1))^2[2\mathcal{E}(X_1)$$
$$+ [U' + \frac{v}{\mathcal{G}_{1,0}^*(v)}](1 + s)$$
$$- \frac{v}{\mathcal{G}_{1,0}^*(v)}[v(1 + s)\mathcal{E}(X_1)\mathcal{E}(B^1) + vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)]]]$$
$$- 12[r + \mathcal{E}(X_1)(1 - \mathcal{Q}^*(v))$$
$$- v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))\mathcal{E}(L^1) - 1][-v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))][M''$$
$$- v\mathcal{E}(X_2) + v[2U'[v(1 + s)\mathcal{E}(X_1)\mathcal{E}(B^1) + vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)]$$
$$+ \frac{v}{\mathcal{G}_{1,0}^*(v)}[(1 + s)\mathcal{B}^{*''}(1)$$
$$+ s\mathcal{S}^{*''}(1)]]] + 12[M' - v\mathcal{E}(X_1)$$
$$+ \frac{v}{\mathcal{G}_{1,0}^*(v)}[v(1 + s)\mathcal{E}(X_1)\mathcal{E}(B^1)$$
$$+ vs\mathcal{E}(X_1)\mathcal{E}(G_2^1)]][\mathcal{T}''(1)(r - 1)$$
$$+ v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))[(1 - \mathcal{Q}^*(v))\mathcal{E}(X_2) + \mathcal{L}^{*''}(1)]$$
$$- v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))\mathcal{E}(L^1)[\mathcal{T}''(1)$$
$$- 2v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))[r + (1 - \mathcal{Q}^*(v))\mathcal{E}(X_1)]]$$
$$+ (1 - \mathcal{Q}^*(v))\mathcal{E}(X_1)[2rv\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))$$
$$+ \mathcal{T}''(1)]]$$

$$Dr_q'''(1) = 6v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))v\mathcal{E}(X_1)[r + \mathcal{E}(X_1)(1$$
$$- \mathcal{Q}^*(v)) - v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))$$
$$\mathcal{E}(L^1) - 1]$$

$$Dr_q''''(1) = 12[[r + \mathcal{E}(X_1)(1 - \mathcal{Q}^*(v))$$
$$- v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))\mathcal{E}(L^1) - 1][\mathcal{T}''(1)\mathcal{E}(X_1)$$
$$+ \mathcal{E}(X_2)v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))]$$
$$+ v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))\mathcal{E}(X_1)[2[rv\mathcal{E}(X_1)(1 - \mathcal{Q}^*(v))$$
$$- v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))\mathcal{E}(L^1)[r$$
$$+ \mathcal{E}(X_1)(1 - \mathcal{Q}^*(v))]]$$
$$+ \mathcal{E}(X_2)(1 - \mathcal{Q}^*(v)) - \mathcal{T}''(1)]]$$

$$Nr_s''''(1) = Nr_q''''(1) - 6v\mathcal{E}(X_1)(1 + \eta\mathcal{E}(S^1))\mathcal{E}(L^1)[v\mathcal{E}(X_1)$$
$$+ v\mathcal{E}(X_1)\eta\mathcal{E}(S^1)][1 - 4M'] + 24M'(v\mathcal{E}(X_1))^2(1$$
$$+ \eta\mathcal{E}(S^1))(1 - \mathcal{Q}^*(v))$$

$$\mathcal{T}''(1) = v\mathcal{E}(X_2)(1 + \eta\mathcal{E}(S^1)) + (v\mathcal{E}(X_1))^2\eta\mathcal{E}(S^2)$$

$$\mathcal{L}^{*''}(1) = v(\mathcal{E}(X_2)\mathcal{E}(L^1) + \mathcal{E}(X_1)\mathcal{E}(L^2)); \ \mathcal{B}^{*''}(1)$$
$$= v(\mathcal{E}(X_2)\mathcal{E}(B^1) + \mathcal{E}(X_1)\mathcal{E}(B^2))$$

$$\mathcal{S}^{*''}(1) = v(\mathcal{E}(X_2)\mathcal{E}(S^1) + \mathcal{E}(X_1)\mathcal{E}(S^2)); \ \mathcal{G}_2^{*''}(1)$$
$$= v(\mathcal{E}(X_2)\mathcal{E}(G_2^1) + \mathcal{E}(X_1)\mathcal{E}(G_2^2))$$

$$U'' = \frac{(v)^2}{\mathcal{G}_{1,0}^*(v)}[\mathcal{E}(G_{1,0}^2)$$
$$+ 2\sum_{m=1}^{M} p_m\mathcal{E}(G_{1,m}^1)\mathcal{E}(G_{1,0}^1) + \sum_{m=1}^{M} p_m\mathcal{E}(G_{1,m}^2)]$$

$$M'' = \frac{v}{\mathcal{G}_{1,0}^*(v)}[\mathcal{E}(G_{1,0}^2) + 2\sum_{m=1}^{M} p_m\mathcal{E}(G_{1,m}^1)\mathcal{E}(G_{1,0}^1)$$

$$+ \sum_{m=1}^{M} p_m \mathcal{E}(G_{1,m}^2)]$$

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

1. Arqub, O. A., & Al-Smadi, M. (2020). Fuzzy conformable fractional differential equations: novel extended approach and new numerical solutions. *Soft Computing, 24*(16), 12501–12522. https://doi.org/10.1007/s00500-020-04687-0

2. Ayyappan, G., & Nirmala, M. (2020). An $M^{[X]}/G(a,b)/1$ queue with unreliable server, second optional service, closedown, setup with N-policy and multiple vacation. *International Journal of Operational Research, 16*(1), 53–81. https://doi.org/10.1504/IJMOR.2020.104686

3. Ayyappan, G., & Nirmala, M. (2021). Analysis of customer's impatience on bulk service queueing system with unreliable server, setup time and two types of multiple vacations. *International Journal of Industrial and Systems Engineering, 38*(2), 198–222. https://doi.org/10.1504/IJISE.2021.115321

4. Bremermann, H. J. (1985). *The evolution of intelligence: the nervous system as a model of its environment*. Seattle: Department of Mathematics: University of Washington.

5. Chih-Lin, I., Han, S., & Bian, S. (2020). Energy-efficient 5G for a greener future. *Nature Electronics, 3*(4), 182–184. https://doi.org/10.1038/s41928-020-0404-1

6. CK, D. M., & Haridass, M. (2023). Analysis of multiple sleeps and N-policy on a M/G/1/K user request queue in 5g networks base station. *The Scientific Temper, 14*(02), 375–382. https://doi.org/10.58414/SCIENTIFICTEMPER.2023.14.2.21

7. Ck, D. M., Selvamuthu, D., & Kalita, P. (2023). Energy Efficiency in a Base Station of 5G Cellular Networks using $M/G/1$ Queue with Multiple Sleeps and N-Policy. *Methodology and Computing in Applied Probability, 25*(2), 48. https://doi.org/10.1007/s11009-023-10026-1

8. Deepa, V., & Haridass, M. (2023). Simulation Analysis of a Base Station Using Finite Buffer M/G/1 Queueing System with Variant Sleeps. *Methodology and Computing in Applied Probability, 25*(4), 80. https://doi.org/10.1007/s11009-023-10052-z

9. Deepa, V., Haridass, M., Selvamuthu, D., & Kalita, P. (2023). Analysis of energy efficiency of small cell base station in 4G/5G networks. *Telecommunication Systems, 82*(3), 381–401. https://doi.org/10.1007/s11235-022-00987-y

10. Falin, G. I., & Templeton, J. G. C. (1997). *Retrial queues*. London: Chapman & Hall.

11. Holland, J. H. (1975). *Adaptation in natural and artificial systems* (pp. 390–401). Ann Arbor: University of Michigan Press.

12. Haridass, M., & Arumuganathan, R. (2015). Analysis of a single server batch arrival retrial queueing system with modified vacations and N-policy. *RAIRO-Operations Research, 49*(2), 279–296. https://doi.org/10.1051/ro/2014037

13. Hawasli, M., & Çolak, S. A. (2017). Toward green 5G heterogeneous small-cell networks: power optimization using load balancing technique. *AEU-International Journal of Electronics and Communications, 82*, 474–485. https://doi.org/10.1016/j.aeue.2017.09.012

14. Holtkamp, H., Auer, G., Bazzi, S., & Haas, H. (2014). Minimizing base station power consumption. *IEEE Journal on Selected Areas in Communications, 32*(2), 297–306. https://doi.org/10.1109/JSAC.2014.141210

15. Jain, M., & Bhargava, C. (2008). Bulk arrival retrial queue with unreliable server and priority subscribers. *International Journal of Operations Research, 5*(4), 242–259.

16. Jain, M., & Upadhyaya, S. (2012). Optimal repairable $M^{[X]}/G/1$ queue with Bernoulli feedback and setup. *International Journal of Mathematics in Operational Research, 4*(6), 679–702. https://doi.org/10.1051/ro/2020074

17. Jain, M., & Kaur, S. (2021). Bernoulli vacation model for $M^{[X]}/G/1$ unreliable server retrial queue with bernoulli feedback, balking and optional service. *RAIRO-Operations Research, 55*, S2027–S2053. https://doi.org/10.1051/ro/2020074

18. Jain, M., & Kumar, A. (2022). Unreliable Server $M^{[X]}/G/1$ Retrial Feedback Queue with Balking, Working Vacation and Vacation Interruption. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences, 93*(1), 57–73. https://doi.org/10.1007/s40010-022-00777-w

19. Kalita, P., & Selvamuthu, D. (2023). Stochastic modelling of sleeping strategy in 5G base station for energy efficiency. *Telecommunication Systems, 83*(2), 115–133. https://doi.org/10.1007/s11235-023-01001-9

20. Karaboga, D., & Basturk, B. (2007). A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of Global Optimization, 39*(3), 459–471. https://doi.org/10.1007/s10898-007-9149-x

21. Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization, Proceedings of ICNN'95. *International Conference on Neural Networks, Perth, WA, Australia*, 4, 1942–1948. https://doi.org/10.1109/ICNN.1995.488968.

22. Krishna Reddy, G. V., Nadarajan, R., & Arumuganathan, R. (1998). Analysis of a bulk queue with N-policy multiple vacations and setup times. *Computers & Operations Research, 25*(11), 957–967. https://doi.org/10.1016/S0305-0548(97)00098-1

23. Kumar, A., & Jain, M. (2023). Cost Optimization of an Unreliable server queue with two stage service process under hybrid vacation policy. *Mathematics and Computers in Simulation, 204*, 259–281. https://doi.org/10.1016/j.matcom.2022.08.007

24. Liu, C., Natarajan, B., & Xia, H. (2015). Small cell base station sleep strategies for energy efficiency. *IEEE Transactions on Vehicular Technology, 65*(3), 1652–1661. https://doi.org/10.1109/TVT.2015.2413382

25. López-Pérez, D., Domenico, A.D., Piovesan, N., Baohongqiang, H., Xinli, G., Bao, H., Qitao, S., & Debbah, M. (2022). A survey on 5G energy efficiency: massive MIMO, lean carrier design, sleep modes, and machine learning. *arXiv preprint*, arXiv:2101.11246.

26. Malik, G., Upadhyaya, S., & Sharma, R. (2012). Cost inspection of a $Geo/G/1$ retrial model using particle swarm optimization and Genetic algorithm. *Ain Shams Engineering Journal, 12*(2), 2241–2254. https://doi.org/10.1016/j.asej.2020.11.012

27. Ma, Z., Guo, S., & Wang, R. (2023). The virtual machines scheduling strategy based on M/M/c queueing model with vacation. *Future*

Generation Computer Systems, 138, 43–51. https://doi.org/10.1016/j.future.2022.08.001

28. Niu, Z., Guo, X., Zhou, S., & Kumar, P. R. (2015). Characterizing energy-delay tradeoff in hyper-cellular networks with base station sleeping control. *IEEE Journal on Selected Areas in Communications, 33*(4), 641–650. https://doi.org/10.1109/JSAC.2015.2393494

29. Pakes, A. G. (1969). Some conditions for Ergodicity and recurrence of Markov chains. *Operations Research, 17*, 1058–1061. https://doi.org/10.1287/opre.17.6.1058

30. Salahdine, F., Opadere, J., Liu, Q., Han, T., Zhang, N., & Wu, S. (2021). A survey on sleep mode techniques for ultra-dense networks in 5G and beyond. *Computer Networks, 201*, 108567. https://doi.org/10.1016/j.comnet.2021.108567

31. Sennott, L. I., Humblet, P. A., & Tweedi, R. L. (1983). Average drifts and the non-Ergodicity of Markov chains. *Operations Research, 31*, 783–789. https://doi.org/10.1287/opre.31.4.783

32. Shortle, J. F., Thompson, J. M., Gross, D., & Harris, C. M. (2018). *Fundamentals of queueing theory* (p. 399). John Wiley & Sons.

33. Upadhyaya, S. (2020). Cost optimization of a discrete-time retrial queue with Bernoulli feedback and starting failure. *International Journal of Industrial and Systems Engineering, 36*, 165–196. https://doi.org/10.1504/IJISE.2020.110245

34. Vaishnawi, M., Upadhyaya, S., & Kulshrestha, R. (2022). Optimal Cost Analysis for Discrete-Time Recurrent Queue with Bernoulli Feedback and Emergency Vacation. *International Journal of Applied and Computational Mathematics, 8*(5), 254. https://doi.org/10.1007/s40819-022-01445-8

35. Woon, L. J., Ramasamy, G., & Thiagarajah, S. P. (2021). Peak power shaving in hybrid power supplied 5G base station. *Bulletin of Electrical Engineering and Informatics, 10*(1), 62–69. https://doi.org/10.11591/eei.v10i1.2705

36. Yang, J., Zhang, X., & Wang, W. (2016). Two-stage base station sleeping scheme for green cellular networks. *Journal of Communications and Networks, 18*(4), 600–609. https://doi.org/10.1109/JCN.2016.000083

37. Zebari, G. M., Zebari, D. A., & Al-zebari, A. (2021). Fundamentals of 5G cellular networks: A review. *Journal of Information Technology and Informatics, 1*(1), 1–5.

38. Zhang, H., Liu, H., Cheng, J., & Leung, V. C. (2017). Downlink energy efficiency of power allocation and wireless backhaul bandwidth allocation in heterogeneous small cell networks. *IEEE Transactions on Communications, 66*(4), 1705–1716. https://doi.org/10.1109/TCOMM.2017.2763623

R. Harini has completed her M.Sc. from Loyola college. She is currently pursuing her Ph.D. from Vellore Institute of Technology, Vellore. She is working in the field of queueing theory, operation research and optimization. She has published one literature review paper and a book chapter related to retrial queueing systems.

K. Indhira has pursued her Ph.D. degree from Vellore Institute of Technology, Vellore. In various Journals and Conferences, she has published over 54 articles. The publications widely include 44 Journal publications and 10 Conference publications. She has published eight books like Feedback queueing models with optional vacations and unreliable server, Multivariable calculus and Differential equations, Calculus and Analysis and many such. She has delivered more than 7 Guest Lectures. She has also been awarded the grant of 20000 USD in the year 2015 for her project titled 'Performance Analysis of Single Server Retrial Queueing System with Vacations and Unreliable Server' by the National Board for Higher Mathematics which was established under the Department of Atomic Energy, Government of India. Indhira has successfully guided over 6 Doctor of Philosophy scholars. She has also mentored various Undergraduate and Post Graduate students.