SURVEY ON NON-MARKOVIAN QUEUE WITH PHASE SERVICE USING SUPPLEMENTARY VARIABLE TECHNIQUE

Binay Kumar[1], Sudeep Singh Sanga[2], Aditya Kumar[1]

•

[1] Patna University, India.
[2] SV National Institute of Technology, Surat, India.
bkmathasr@gmail.com, ssanga@amhd.svnit.in, ak606295@gmail.com

Abstract

In the present article, we provide a comprehensive overview and literature review on performance modeling of queueing systems with phase service, utilizing the supplementary variable technique. We discuss the scenarios that necessitate phase service and explain how the supplementary variable technique is applied to analyze non-Markovian models. Additionally, we offer a summary of the fundamental concepts and existing literature on queueing systems with phase service. Our review covers research conducted over the past decade (2014-2024) on queues featuring phase service, including aspects such as bulk arrival, service interruptions, and discouragement. The review systematically explores models incorporating bulk arrivals, service interruptions, customer discouragement, and other complex operational features. By consolidating recent advancements and identifying common modeling approaches, this work aims to provide valuable insights for researchers and practitioners engaged in the analysis and design of advanced queueing systems.

Keywords: bulk arrival, phase service, supplementary variable, non-Markovian.

I. Introduction

In the queueing theory literature, extensive research has been conducted on models in which both interarrival and service times follow the exponential distribution. These models are commonly referred to as Markovian queueing systems, as they can be effectively analyzed using the Markov property, which simplifies the modeling process due to its memoryless nature. However, in many real-world situations, the assumption of exponential interarrival or service times (or both) may not hold. For instance, service times in manufacturing systems, healthcare facilities, or computer networks often follow general probability distributions. Queueing systems in which the interarrival time, service time, or both deviate from the exponential form are referred to as non-Markovian queueing models. Due to their increased complexity and the absence of the memoryless property, these models cannot be solved directly using conventional Markovian techniques. To analyze such systems, a common approach is to transform them into Markovian models by introducing additional state variables through the supplementary variable technique. This technique involves augmenting the state space typically by including variables like the elapsed or remaining service time—thereby enabling the use of Markovian methods to study the system's behavior.

In real-life scenarios, we often encounter situations where customers have preferences for multiple types of services from a single service provider. This is particularly evident on various online platforms that offer a wide range of services under one roof. For instance, platforms like Amazon and Flipkart provide an extensive array of products and services, allowing customers to choose from diverse categories such as electronics, clothing, and household goods all in one place. Similarly, travel and booking apps like Make My Trip provide a variety of services, such as flight bookings, hotel reservations, and car rentals, all conveniently available on a single platform to meet diverse customer needs. Phase queueing models adeptly handle such complexity by managing different service stages and preferences, thus providing a robust framework for optimizing service delivery and efficiency in these multifaceted environments.

The exponential and Poisson distributions are widely utilized in real-time queue models due to their broad applicability. The exponential distribution, in particular, is widely preferred due to its practical applicability and the ease of analyzing queueing models based on it. To handle complex scenarios, mixtures of exponential distributions are often employed. The most common among these are those which help model services provided in multiple phases with varying levels of complexity.

Erlangian Distribution: The Erlangian distribution is an effective model for handling phase-type service processes. In this distribution, the service process is divided into several sequential phases, each following an i.i.d. exponential distribution. This structure ensures that the overall service time is the sum of multiple exponential phases, resulting in an Erlangian distribution. The design of the Erlangian distribution is inherently sequential, meaning that service must progress through each phase in a fixed order before completion. This makes it particularly useful for modeling structured service processes in queueing systems, telecommunications, and reliability analysis.

Hyper exponential distribution: The hyper exponential distribution models service times with high variability by employing a parallel arrangement of multiple exponential distributions. Unlike the Erlangian distribution, where service phases occur sequentially, the hyper exponential distribution allows each arriving customer to be randomly assigned to one of several exponential phases, each with a distinct mean service rate and a predefined selection probability.

A more generalized approach to phase-type service models considers cases where service times follow a general distribution rather than strictly exponential phases. These models incorporate a mixture of different service distributions to provide a more flexible and realistic representation of service processes. The selection of specific distributions may be either mandatory or optional, depending on the system's design. Some commonly used phase-type queueing models are:

(i). **Single-Phase Multiple options:** In this system, service consists of a single stage, after which the unit exits. This single-phase service can be either single or multi-optional. If the service is single optional, the arriving unit has no choice and must undergo that service. However, if the service is multi-optional, the unit can choose one service from the available options before exiting the system. For example, at a billing counter, a customer is allowed to pay the bill either in cash or online.

(ii). **Multi-phase queueing models:** In such a type of system, there is a single server that provides services in more than one phase. The multi-phase may be further classified into the following subcategories:

a) **Two-phase multi-optional service:** In such a system, an arriving customer, after availing essential service, chooses only one optional service among the available optional services. After availing of that optional service unit exits from the system.

b) **Multi-optional phase service:** In such a system of systems, sequences of optional services are provided one after the other as per the demand of the unit after completion of the initial essential service.

c) **Multi-essential Phase Queueing Models:** In such a type of system, a service is provided in a sequence of essential services, one after the other. The service will not be complete until all the phases are completed.

In real-world queueing systems, various factors such as balking, server breakdowns, delayed repairs, vacations, and retrials significantly impact service efficiency and system performance. These complexities, often occurring alongside phase-type service, introduce challenges in queueing analysis, requiring advanced modeling techniques for accurate representation. For instance, customers may opt not to join a queue due to excessive wait times (balking), while system disruptions such as server failures or temporary unavailability (vacations) create additional unpredictability. Addressing these dynamics is crucial for optimizing service operations across various industries, including manufacturing, healthcare, telecommunications, and computer networks.

The scope of this study includes:

- Analyzing queueing models with phase-type service, where service is completed in multiple stages rather than a single step. These models more accurately capture real-world processes such as multi-stage production lines, sequential healthcare procedures, and layered data processing in communication networks.
- Investigating non-Markovian queueing systems, where service time distributions are more general and require advanced analytical methods like supplementary variable technique.
- Reviewing key literature from 2014 to 2024, focusing on advancements in queueing models that incorporate bulk arrivals, customer discouragement, service interruptions, and other complex features.

The layout of the remaining portion of the present study is organized as follows: Section II introduces the supplementary variable technique (SVT) and its application in general service queueing models, particularly those with batch arrivals. Section III applies SVT to specific queueing models featuring different types of phase-type service. Section IV provides an in-depth review of literature on single-server queueing models with phase-type service, covering extensions like unreliable servers, server vacations, retrial queues, balking, and Bernoulli feedback. Section V presents the main conclusions and discusses possible directions for future research.

II. Supplementary Variable Technique (SVT)

When inter-arrival and service times are exponentially distributed, the system is classified as a Markovian system. These systems can be modeled using the memoryless property of exponential distributions, which simplifies the analysis and allows for the use of Markov processes to describe the system's behavior.

However, when the service time or inter-arrival time follows a more general, non-exponential distribution, the system is referred to as non-Markovian. Analyzing non-Markovian systems is more complex because the future state of the system depends not only on the current state but also on the history of past events. To address this complexity, the SVT is employed. This method introduces an additional variable, known as the supplementary variable, which accounts for the elapsed service time or the remaining time of the non-exponentially distributed random variable. By doing so, the non-Markovian system can be transformed into a Markovian system, making it easier to analyze. This technique was pioneered by Cox [1], specifically to study the M/G/1 queue, where "M" stands for Markovian (exponential inter-arrival times), "G" denotes a general distribution for service times, and "1" represents a single server in the system. Cox's method provides a powerful tool for analyzing and understanding a broader range of queueing systems that do not conform to the simple exponential assumptions of Markovian models.

I. SVT for Single Arrival Queueing model

It is helpful to outline the procedure for studying the distribution of queue size in the M/G/1 queue using SVT (Cox [1]). Consider an M/G/1 model where units arrive in Poisson fashion with rate λ. The service times are governed by a general distribution described by the function $B(x)$, its Laplace transform $B^*(s)$, and its first moment $E\,(B)$. Let $N(t)$ represent the total number of units in the system at time t, and $B^0(t)$ be the variable added that corresponds to the elapsed service time. Let us define a random variable as follows

$$\xi(t) = \begin{cases} 0 & : \text{ if the server is idle} \\ 1 & : \text{ if the server is busy} \end{cases}$$

The bivariate random variable $\{N(t), \sigma(t)\}$ forms a Markov process, where $\sigma(t) = 0 \; if \; \xi(t) = 0$ and $\sigma(t) = B^0(t)$. To study the system's dynamics, the governing equations for the present model can be formulated based on probabilistic principles (Cox [1]). We introduce the following definitions:

$$P_0(t) = Prob.\{\, N(t) = 0, \sigma(t) = 0\} \tag{1}$$

$$P_n(t,x) = Prob.\{N(t) = n, \sigma(t) = B^0(t), x < B^0(t) \le x + dx\} \;\; ; x > 0 \tag{2}$$

And the steady state probability is given by

$$P_0 = \lim_{t\to\infty} P_0(t) \quad and \;\; P_n(x) = \lim_{t\to\infty} \; P_n(t,x) \tag{3}$$

The hazard function for a state in which the server is actively serving is given by

$$\mu(x)dx = \frac{dB(x)}{1-B(x)} \tag{4}$$

The governing equation for the transient process can be derived by examining the transitions occurring over a time interval Δt (Chaudhry and Templeton [2]).

$$P_0(t+\Delta t) = \; P_0(t)(1-\lambda\Delta t) + (1-\lambda\Delta t)\int_0^\infty \mu(x)P_1(x,t)dx + o(\Delta t)$$

which will give

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t) + \int_0^\infty \mu(x)P_1(x,t)dx \tag{5}$$

$$P_n(x+\Delta t, t+\Delta t) = \; P_n(x,t)(1-\lambda\Delta t)(1-\mu(x)\Delta t) + \; \lambda\Delta t P_n(x,t)(1-\mu(x)\Delta t) + o(\Delta t), \quad \text{for } n \ge 1$$

which gives

$$\frac{\partial P_n(x,t)}{\partial t} + \frac{\partial P_n(x,t)}{\partial x} = -\big(\lambda + \mu(x)\big)P_n(x,t) + \lambda P_{n-1}(x,t) \quad \text{for n} \ge 1, \text{x} > 0. \tag{6}$$

The steady state equation of (5) and (6) is given by taking $\to \infty$, we have

$$\lambda P_0 = \int_0^\infty \mu(x)P_1(x)dx \tag{7}$$

$$\frac{dP_n(x)}{dx} = -\big(\lambda + \mu(x)\big)P_n(x) + \lambda P_{n-1}(x) \quad \text{for n} \ge 1, \text{x} > 0. \tag{8}$$

The equations (7) and (8) are solved by using the boundary conditions

$$P_n(0) = \lambda\delta_{n,1}P_0 + \int_0^\infty \mu(x)P_{n+1}(x)dx \tag{9}$$

These equations are to be solved subject to the normalization condition

$$P_0 + \sum_{n=1}^\infty \int_0^\infty P_n(x)dx = 1 \tag{10}$$

Additionally, we introduce the following probability-generating functions.

$$P(x,z) = \sum_{n=1}^\infty P_n(x)z^n \qquad \text{and} \;\; P(0,z) = \sum_{n=1}^\infty P_n(0)\, z^n \tag{11}$$

Solving equations (7)-(9) and using (10) and (11), we get the PGF of the queue size distribution

$$\varphi(z) = (1-\lambda E(B))(z-1)B^*(\lambda(1-z))/[z - B^*\big(\lambda(1-z)\big)] \tag{12}$$

Using equation (12), we can find the system performance indices, such as average queue length and mean waiting time.

II. SVT for Bulk Queueing Model

In some queueing systems, units arrive in groups of random size. For a detailed study on batch arrival, we refer to the work done by Chaudhury and Templeton [2], Bhorthakur and Medhi [3]. Let X be the random variable representing the batch size, with $c_k \; = \; P\,(X = k)$ and an expected batch size of E(X). Then, the governing equation discussed earlier is modified accordingly.

$$\lambda P_0 = \int_0^\infty \mu(x) P_1(x) dx \tag{13}$$

$$\frac{dP_n(x)}{dx} = -(\lambda + \mu(x)) P_n(x) + \lambda(1 - \delta_{n,1}) \sum_{j=1}^{n} c_j P_{n-j}(x) \text{ , n} > 0, \text{x} > 0 \tag{14}$$

To solve equations (13) and (14), we will use the boundary condition

$$P_n(0) = \lambda c_n P_0 + \int_0^\infty \mu(x) P_{n+1}(x) dx \tag{15}$$

The probability distribution for the queue size is given by

$$\varphi(z) = (1 - \lambda E(X)(E(B))(z-1) B^*(\lambda(1 - X(z)))/[z - B^*(\lambda(1 - X(z)))] \tag{16}$$

III. Different Phase type models using SVT

I. Two phase service

This section presents a concise approach to analyzing a single-arrival queueing system with two service phases. The first phase is compulsory for all customers, whereas the second phase is optional. Customers can opt for the additional service phase based on their specific needs.

Consider a queueing situation in which units join the system in Poisson fashion with arrival rate λ, and service is provided in two phases. Let both phases' service follow general distribution with distribution functions respectively $B_1(x), B_2(x)$, with Laplace transforms $B_1^*(s), B_2^*(s)$, and first moments $E(B_1), E(B_2)$. Let $N(t)$ denote the number of units in the system at time t, and $B^1(t)$, $B^2(t)$ be supplementary variables that denote the elapsed service time for the first phase and second phase, respectively. Let us define the system state equations as follows

$$\xi(t) = \begin{cases} 0 & : \textit{if the server is idle} \\ 1 & : \textit{if the server is busy in the first phase of service} \\ 2 & : \textit{if the server is busy in the second phase of service} \end{cases}$$

The bivariate random variable $\{N(t), \sigma(t)\}$ form a Markov process, where $\sigma(t) = 0, if\ \xi(t) = 0, \sigma(t) = B^1(t)\ if\ \xi(t) = 1, v(t) = B^2(t)\ if\ \xi(t) = 2$. To study the system's dynamics, the governing equations for the queueing model with two-phase service can be formulated based on probabilistic principles (Cox [1]).

We define

$$P_0(t) = Prob.\{\, N(t) = 0, \sigma(t) = 0\} \tag{17}$$

$$P_n^{(1)}(x,t) = Prob.\{N(t) = n, \sigma(t) = B^0(t), x < B^1(t) \leq x + dx\,\} \text{ for } x > 0 \tag{18}$$

$$P_n^{(2)}(t,x) = Prob.\{N(t) = n, \sigma(t) = B^0(t), x < B^2(t) \leq x + dx\} \text{ for } x > 0 \tag{19}$$

And the steady state probability is given by

$$P_0 = \lim_{t\to\infty} P_0(t) \text{ and } P_n^{(j)}(x) = \lim_{t\to\infty} P_n^{(j)}(t,x),\ j = 1,2. \tag{20}$$

The hazard rate functions for the state in which the server is busy are given by

$$\mu_1(x)dx = \frac{dB_1(x)}{1 - B_1(x)},\ \mu_2(x)dx = \frac{dB_2(x)}{1 - B_2(x)} \tag{21}$$

The equations that govern the system are given by

$$\lambda P_0 = \int_0^\infty \mu_2(x) P_1^{(2)}(x) dx + (1-p) \int_0^\infty \mu_1(x) P_1^{(1)}(x) dx \tag{22}$$

$$\frac{dP_n^{(1)}(x)}{dx} = -(\lambda + \mu_1(x)) P_n^{(1)}(x) + \lambda P_{n-1}^{(1)}(x) \text{ for n} > 0, \text{x} > 0 \tag{23}$$

$$\frac{dP_n^{(2)}(x)}{dx} = -(\lambda + \mu_2(x)) P_n^{(2)}(x) + \lambda P_{n-1}^{(2)}(x) \text{ for n} > 0, \text{x} > 0 \tag{24}$$

To solve equations (23) and (24), we use the following boundary condition

$$P_n^{(1)}(0) = \lambda \delta_{n,1} P_0 + q \int_0^\infty \mu_1(x) P_{n+1}^{(1)}(x) dx + \int_0^\infty \mu_2(x) P_{n+1}^{(2)}(x) dx\ , n > 0 \tag{25}$$

$$P_n^{(2)}(0) = p \int_0^\infty \mu_1(x) P_n^{(1)}(x) dx\ , n > 0 \tag{26}$$

Equation (22)-(26) are solved under the normalization condition

$$P_0 + \sum_{n=1}^{\infty} [P_n^{(1)}(x) + P_n^{(2)}(x)] = 1 \tag{27}$$

For simplification purposes, we define the following probability-generating functions.

$$P_1(x,z) = \sum_{n=1}^{\infty} P_n^{(1)}(x)\, z^n\ ;\ P_1(0,z) = \sum_{n=1}^{\infty} P_n^{(1)}(0)\, z^n$$

$$P_2(x,z) = \sum_{n=1}^{\infty} P_n^{(2)}(x)\, z^n\ ;\ P_2(0,z) = \sum_{n=1}^{\infty} P_n^{(2)}(0)\, z^n$$

Solving equations (22)-(24) and using equations (25), (26), and (27), we get the probability distribution function of the queue size distribution

$$\varphi(z) = \frac{(1-\rho)(z-1)[(q+B_2^*(a(z)))B_1^*(a(z))]}{[z-\{q+B_2^*(a(z)))B_1^*(a(z))\}]} \tag{28}$$

where $a(z) = \lambda(1-z)$ and $\rho = \lambda(E(B_1) + p\,E(B_2))$

Using equation (28), we can evaluate the mean queue size and average waiting time.

II. Multi-optional phase service

In this section, we outline a brief procedure for analyzing a single-arrival queueing system having multiple optional service phases. We assume that the first phase of service is mandatory for all customers, while the subsequent phases are optional. Customers have the flexibility to choose additional service phases based on their individual requirements.

Consider a single-arrival queueing system where units arrive according to a Poisson process with an arrival rate λ. The service process follows a general distribution and is structured into m phases. The initial phase of service is mandatory for all arriving units, while the remaining stages are optional and provided based on demand. After completing the first mandatory service, a unit may request the second service with probability p_1 or exit from the system with probability $q_1 = 1 - p_1$. If the unit opts for the second service, it may then request the third optional service with probability p_2 or leave with probability $q_2 = 1 - p_2$. This pattern continues, where after completing the $(m-1)^{th}$ optional service, the unit may either request the m^{th} optional service with probability p_{m-1} or leave the system with probability $q_{m-1} = 1 - p_{m-1}$. Let the service times for all phases follow a general distribution, with the distribution function for the ith phase given by $B_i(x)$. The corresponding Laplace transform is denoted as $B_i^*(s)$, and the first moment is $E(B_i)$ for $i = 1,2,\dots,m$. Let $N(t)$ represent the unit count in the system at time t, and let $B^i(t)$ be a supplementary variable representing the elapsed service time for the i^{th} optional service phase, where optional phases for $i = 1,2,..,m$. Now, we define a random variable as follows:

$$\xi(t) = \begin{cases} 0 & : \text{if the server is idle} \\ 1 & : \text{if the server is busy with service} \\ 2 & : \text{if the server is busy with second service} \\ . & \\ . & \\ . & \\ m & : \text{if the server is busy with mth service} \end{cases}$$

The bivariate random variable $\{N(t), \sigma(t)\}$ form Markov process, where $\sigma(t) = 0, if\ \xi(t) = 0, \sigma(t) = B^i(t)\ if\ \xi(t) = i$, for $i = 1,2, \dots m$. The governing equations for the present case can be derived using probabilistic reasoning, as outlined by Cox [1}, as follows:

$$P_0(t) = Prob.\{\,N(t) = 0, \sigma(t) = 0\} \tag{29}$$

$$P_n^{(i)}(t,x) = Prob.\{\,N(t) = n, \sigma(t) = B^i(t), x < B^i(t) \le x + dx\,\} \quad \text{for } x > 0 \quad i = 1,2,\dots,m \tag{30}$$

The hazard rate for essential and optional service rates are given by

$$\mu_i(x)dx = \frac{dB_i(x)}{1-B_i(x)}, \text{for } i = 1,2,\dots,m \tag{31}$$

The equations that govern the system are given by

$$\lambda P_0 = \int_0^\infty \mu_m(x) P_1^{(m)}(x)dx + \sum_{i=1}^{m-1} q_i \int_0^\infty \mu_i(x) P_1^{(i)}(x)dx \tag{32}$$

$$\frac{dP_n^{(i)}(x)}{dx} = -(\lambda + \mu_i(x))P_n^{(i)}(x) + \lambda P_{n-1}^{(i)}(x) \quad \text{for } n \ge 1, x > 0, i = 1,2,\dots,m \tag{33}$$

The boundary conditions are given by

$$P_n^{(1)}(0) = \lambda \delta_{n,1} P_0 + \int_0^\infty \mu_m(x) P_{n+1}^{(m)}(x)dx + \sum_{i=1}^{m-1} q_i \int_0^\infty \mu_i(x) P_{n+1}^{(i)}(x)dx \tag{34}$$

$$P_n^{(i)}(0) = p_{i-1} \int_0^\infty \mu_{i-1}(x) P_n^{(i-1)}(x)dx \tag{35}$$

The normalization condition for the present model is given by

$$P_0 + \sum_{i=1}^{m} \sum_{n=1}^{\infty} \int_0^{\infty} P_n^{(i)}(x)\, dx = 1 \tag{36}$$

The probability generating functions used for the present system are given by

$$P_i(x,z) = \sum_{n=1}^{\infty} P_n^{(i)}(x)\, z^n \quad \text{and} \quad P_i(0,z) = \sum_{n=1}^{\infty} P_n^{(i)}(0)\, z^n \tag{37}$$

Multiply governing equation (23) by a proper power of z and using (35), we have

$$P_i(x,z) = P_i(0,z)\exp\{-a(z)x\}\,[1 - B_i(x)],\ \ for\ i = 1,2,\dots,m \tag{38}$$

On boundary condition (34), apply the same treatment and using (32), we have

$$zP_1(0,z) = \lambda P_0 z(z-1) + \sum_{i=1}^{m} q_i P_i(0,z) B_i^*(0,z) \quad \text{with } q_m = 1 \tag{39}$$

$$P_i(0,z) = P_1(0,z) \prod_{j=1}^{i-1} p_j B_j^*\big(a(z)\big), \qquad i = 2,3,\dots,m \tag{40}$$

From (39) and (40), we have

$$P_1(0,z) = \frac{\lambda P_0 z(z-1)}{[z - \sum_{i=1}^{m} q_i \{\prod_{j=1}^{i-1} p_j B_j^*(a(z))\} B_i^*(a(z))]} \tag{41}$$

$$P_i(0,z) = \frac{\lambda P_0 z(z-1) \prod_{j=1}^{i-1} p_j B_j^*(a(z))}{[z - \sum_{i=1}^{m} q_i \{\prod_{j=1}^{i-1} p_j B_j^*(a(z))\} B_i^*(a(z))]}, \qquad i = 2,3,\dots,m \tag{42}$$

Integrating equations (38) with limit 0 to ∞ and using (41)-(42), we have

$$P_1(z) = \frac{P_0 z[1 - B_1^*(a(z))]}{[\sum_{i=1}^{m} q_i \{\prod_{j=1}^{i-1} p_j B_j^*(a(z))\} B_i^*(a(z)\} - z]} \tag{43}$$

$$P_i(z) = \frac{P_0 z \prod_{j=1}^{i-1} p_j B_j^*(a(z))\}[1 - B_i^*(a(z)]}{[\sum_{i=1}^{m} q_i \{\prod_{j=1}^{i-1} p_j B_j^*(a(z))\} B_i^*(a(z)\} - z]} \tag{44}$$

Taking limit $z \to 1$ in (41)-(44) and using equation (36), we have
$P_0 = 1 - \rho$
where $\rho = \lambda[E(B_1) + p_1 E(B_2) + p_1 p_2 E(B_3) + \dots + p_1 p_2 \dots p_{m-1} E(B_m)]$
The PGF for the queue size distribution is given by

$$\varphi(z) = \frac{(1-\rho)[\sum_{i=1}^{m} q_i \{\prod_{j=1}^{i-1} p_j B_j^*(a(z))\} B_i^*(a(z)\} - z + z\sum_{i=1}^{m} \{\prod_{j=1}^{i-1} p_j B_j^*(a(z))\}(1 - B_i^*(a(z))]}{[\sum_{i=1}^{m} q_i \{\prod_{j=1}^{i-1} p_j B_j^*(a(z))\} B_i^*(a(z)\} - z]} \tag{45}$$

III. Single phase multiple options

In this section, we present a procedure for analyzing a queueing system where units arrive following a Poisson process and receive service in a single phase with multiple options. Each unit selects one service from m available options and exits the system upon service completion.
Consider M/G/1 queueing model in which a unit joins the system at a Poisson arrival rate λ. Here, each arriving unit chooses the i[th] optional service with probability p_i, and after completion, that service unit leaves the system. Let the service times for all phases follow a general distribution, with the distribution function for the ith phase given by $B_i(x)$. The corresponding Laplace transform is denoted as $B_i^*(s)$, and the first moment is $E(B_i)$ for $i = 1,2,\dots,m$,
Let us define a random variable as follows

$$\xi(t) = \begin{cases} 0 & : if\ the\ server\ is\ idle \\ 1 & : if\ the\ server\ is\ busy\ with\ first\ service \\ 2 & : if\ the\ server\ is\ busy\ with\ second\ service \\ . & \\ . & \\ . & \\ m & : if\ the\ server\ is\ busy\ with\ mth\ service \end{cases}$$

The bivariate random variable $\{N(t), \sigma(t)\}$ form a Markov process, where
$\sigma(t) = 0, if\ \xi(t) = 0,\ \sigma(t) = B^i(t)\ if\ \xi(t) = i,$ for $i = 1,2,\dots m$.
The system state governing equation for the present system is given by

$$P_0(t) = Prob.\{\,N(t) = 0, \sigma(t) = 0\} \tag{46}$$

$$P_n^{(i)}(t,x) = Prob.\{\,N(t) = n, \sigma(t) = B^i(t), x < B^i(t) \le x + dx\,\} \quad ; x > 0,\quad i = 1,2,\dots,m \tag{47}$$

The hazard rate for the i[th] service is given by

$$\mu_i(x)dx = \frac{dB_i(x)}{1-B_i(x)}\text{ , for } i = 1,2,\dots,m \tag{48}$$

The equations that govern the system are given by

$$\lambda P_0 = \sum_{i=1}^{m}\int_0^\infty \mu_i(x)P_1^{(i)}(x)dx \tag{49}$$

$$\frac{dP_n^{(i)}(x)}{dx} = -\big(\lambda+\mu_i(x)\big)P_n^{(i)}(x) + \lambda P_{n-1}^{(i)}(x) \quad \text{for n} \geq 1, \text{x} > 0, \text{i} = 1,2,\dots,\text{m} \tag{50}$$

The boundary conditions are given by

$$P_n^{(j)}(0) = \lambda p_j\delta_{n,1}P_0 + p_j\sum_{i=1}^{m}\int_0^\infty \mu_i(x)P_{n+1}^{(i)}(x)dx\,,\ \ n\geq 1, j = 1,2,..,m \tag{51}$$

The normalization condition for the present model is as follows

$$P_0 + \sum_{i=1}^{m}\sum_{n=1}^{\infty}\int_0^\infty P_n^{(i)}(x)\,dx = 1 \tag{52}$$

The PGF used for the present system is as follows

$$P_i(x,z) = \sum_{n=1}^{\infty}P_n^{(i)}(x)\,z^n \qquad \text{and} \quad P_i(0,z) = \sum_{n=1}^{\infty}P_n^{(i)}(0)\,z^n \tag{53}$$

Multiply governing equation (50) by the proper exponent of z, and using (53), we have

$$P_j(x,z) = P_j(0,z)\exp\{-a(z)x\}\big[1-B_j(x)\big],\ \text{for } j = 1,2,\dots,m \tag{54}$$

By multiplying equation (51) by a proper exponent of z and summing over all terms, we obtain

$$zP_j(0,z)[z-p_jB_j^*(a(z))] = p_j\lambda P_0z(z-1) + p_j\sum_{\substack{i=1\\ i\neq j}}^{m}P_i(0,z)B_i^*\big(a(z)\big), j = 1,2,\dots,m \tag{55}$$

Equation (55) gives m simultaneous equation in variables $P_1(0,z), P_2(0,z),\dots,P_m(0,z)$, on solving this we get

$$P_j(0,z = \frac{p_j\lambda P_0z(z-1)}{\sum_{i=1}^{m}[p_iB_i^*(a(z))-1]},\ j = 1,2,\dots,m \tag{56}$$

Solving the equation (54) and using equation (56), we have

$$P_j(z) = \frac{p_jP_0z(1-B_j^*(a(z)))}{[\sum_{i=1}^{m}p_iB_i^*(a(z))-z]},\ for\ j = 1,2,\dots,m \tag{57}$$

Now applying the normalizing condition $P_0 + \sum_{i=1}^{m}P_i(1) = 1$, we have

$$P_0 = (1-\rho) \tag{58}$$

Where $\rho = \lambda[p_1E(B_1) + p_2E(B_2) + p_3E(B_3) + \cdots + p_mE(B_m)]$,

The queue size distribution is characterized by its probability generating function, expressed as

$$\varphi(z) = \frac{(1-\rho)(1-z)[\sum_{i=1}^{m}p_iB_i^*(a(z)\}]}{[\sum_{i=1}^{m}p_iB_i^*(a(z)\}-z]} \tag{59}$$

IV. Two-phase multi-optional service

In this section, we outline a concise procedure for solving a queueing model with a two-phase service structure. The first phase is mandatory for all customers, whereas the second phase offers m optional services. Customers can select one of these m available service options based on their preferences.

Consider $M/G/1$ queueing situation where units join the system in a Poisson arrival rate λ and service is provided in two phases. The first phase is essential to all, while the second is optional. The customer may choose i^{th} optional service with probability p_i after completion of the essential service. Let both phases' service follow general distribution with distribution functions respectively $B_e(x), B_i(x)$, with Laplace transforms $B_e^*(s)$, $B_i^*(s)$, and first moments $E(B_e), E(B_i)i = 1,2,..,m$. Let $N(t)$ denote the number of units in the system at time t, and $B^e(t)$, $B^i(t)$ be supplementary variables that denote the elapsed service time for the first phase and second phase, respectively for $i = 1,2,..,m$. Let us define a random variable as follows

$$\xi(t) = \begin{cases} 0 & : if\ the\ server\ is\ idle \\ 1 & : if\ the\ server\ is\ busy\ in\ the\ first\ phase\ service \\ i+1 & : if\ the\ server\ is\ busy\ in\ rendering\ ith\ optional\ service \end{cases}$$

The bivariate random variable $\{N(t),\sigma(t)\}$ form Markov process, where $v(t) = 0, if\ \xi(t) = 0, \sigma(t) = B^e(t)\ if\ \xi(t) = 1, \sigma(t) = B^i(t)\ if\ \xi(t) = i+1$.

The governing equation for the present system is given by

$$P_0(t) = Prob.\{\,N(t) = 0, \sigma(t) = 0\} \tag{60}$$

$$P_n^{(e)}(t,x) = Prob.\{\,N(t) = n, \sigma(t) = B^e(t), x < B^e(t) \leq x+dx\,\}\ \ ; x > 0 \tag{61}$$

$$P_n^{(i)}(t,x) = Prob.\{ N(t) = n, \sigma(t) = B^i(t), x < B^i(t) \leq x + dx \} \;\; ; x > 0 \tag{62}$$

And the steady state probability is given by

$$P_0 = \lim_{t\to\infty} P_0(t)\,, P_n^{(e)}(x) = \lim_{t\to\infty} P_n^{(e)}(t,x)\,, P_n^{(i)}(x) = \lim_{t\to\infty} P_n^{(i)}(t,x) \tag{63}$$

The hazard rates for both the essential and optional service phases are given by

$$\mu_e(x)dx = \frac{dB_e(x)}{1-B_e(x)},\; \mu_i(x)dx = \frac{dB_i(x)}{1-B_i(x)} \tag{64}$$

The governing equation of the present system is given by

$$\lambda P_0 = \sum_{i=1}^{m} \int_0^\infty \mu_i(x) P_1^{(i)}(x)dx + p_0 \int_0^\infty \mu_1(x) P_1^{(1)}(x)dx \tag{65}$$

$$\frac{dP_n^{(e)}(x)}{dx} = -\big(\lambda + \mu_e(x)\big)P_n^{(e)}(x) + \lambda P_{n-1}^{(e)}(x) \quad \text{for } n \geq 1, x > 0 \tag{66}$$

$$\frac{dP_n^{(i)}(x)}{dx} = -\big(\lambda + \mu_i(x)\big)P_n^{(i)}(x) + \lambda P_{n-1}^{(i)}(x) \quad \text{for } n \geq 1, x > 0 \tag{67}$$

Equations (66) and (67) are solved under conditions at x = 0

$$P_n^{(e)}(0) = \lambda \delta_{n,1} P_0 + p_0 \int_0^\infty \mu_e(x) P_{n+1}^{(e)}(x)dx + \sum_{i=1}^{m} \int_0^\infty \mu_i(x) P_{n+1}^{(i)}(x)dx \tag{68}$$

$$P_n^{(i)}(0) = p_i \int_0^\infty \mu_e(x) P_n^{(e)}(x)dx \;\; , n \geq 1 \tag{69}$$

The normalization condition for the present case is given by

$$P_0 + \sum_{n=1}^{\infty} \int_0^\infty P_n^{(e)}(x) + \sum_{i=1}^{m} \sum_{n=1}^{\infty} \int_0^\infty P_n^{(i)}(x)dx = 1 \tag{70}$$

The probability generating functions used in this part are

$$P_e(x,z) = \sum_{n=1}^{\infty} P_n^{(e)}(x)\, z^n \qquad \text{and} \quad P_e(0,z) = \sum_{n=1}^{\infty} P_n^{(e)}(0)\, z^n$$

$$P_i(x,z) = \sum_{n=1}^{\infty} P_n^{(i)}(x)\, z^n \qquad \text{and} \quad P_i(0,z) = \sum_{n=1}^{\infty} P_n^{(i)}(0)\, z^n$$

Solving equation (66) in the usual manner and using (64), we have

$$P_e(x,z) = P_e(0,z) \exp\{-a(z)x\}\, [1 - B_e(x)], \tag{71}$$

$$P_i(x,z) = P_i(0,z) \exp\{-a(z)x\}\, [1 - B_i(x)],\; for\ i = 1,2,\dots,m \tag{72}$$

Solving equations (68) and (69) in the usual manner, we have

$$zP_e(0,z) = \lambda P_0 z(z-1) + p_0 P_e(0,z) B_e^*(a(z)\,) + \sum_{i=1}^{m} p_i P_i(0,z) B_i^*(a(z)\,) \tag{73}$$

$$P_i(0,z) = P_e(0,z) p_i B_e^*\big(a(z)\big),\;\; i = 2,3,\dots,m \tag{74}$$

From equations (73) and (74) we have

$$P_e(0,z) = \frac{\lambda P_0 z(z-1)}{[z - \{p_0 + \sum_{i=1}^{m} p_i B_i^*(a(z))\} B_e^*(a(z))]} \tag{75}$$

$$P_i(0,z) = \frac{\lambda P_0 z(z-1) p_i B_e^*\big(a(z)\big)}{[z - \{p_0 + \sum_{i=1}^{m} p_i B_i^*(a(z))\} B_e^*(a(z))]}, \qquad i = 2,3,\dots,m \tag{76}$$

Solving equations (71) and (72) using equations (75) and (76), we have

$$P_e(z) = \frac{P_0 z\, [1 - B_e^*(a(z))]}{[\{p_0 + \sum_{i=1}^{m} p_i B_i^*(a(z))\} B_e^*(a(z)) - z]}, \tag{77}$$

$$P_i(z) = \frac{P_0 z p_i B_e^*(a(z))[1 - B_i^*(a(z))]}{[\{p_0 + \sum_{i=1}^{m} p_i B_i^*(a(z))\} B_e^*(a(z)) - z]}, for\ i = 1,2,\dots,m \tag{78}$$

Now, applying the normalizing condition (70), we have

$$P_0 = (1-\rho) \tag{79}$$

Where $\rho = \lambda[E(B_e) + \sum_{i=1}^{m} p_i E(B_i)]$

The PGF for the queue size distribution is given by

$$\varphi(z) = \frac{(1-\rho)(1-z)[\{p_0 + \sum_{i=1}^{m} p_i B_i^*(a(z))\} B_e^*(a(z))]}{[\{p_0 + \sum_{i=1}^{m} p_i B_i^*(a(z))\} B_e^*(a(z)) - z]} \tag{80}$$

IV. Review of Literature

I. Literature review(Up to 2013)

Queueing models incorporating phase-type services have seen widespread application across a variety of domains, including industrial operations, telecommunications, and computer networks. A significant body of research has focused on systems in which the service process is completed in multiple phases—either two or more—with some phases being essential, optional, or a mix of both.

This section reviews key contributions to the study of queueing models with phase-type services prior to 2014. The literature can broadly be classified based on whether the service phases are optional, essential, or multi-optional. Table 1 summarizes these studies, highlighting the nature of the service phases, the methodological tools employed (such as Laplace-Stieltjes Transforms (LST), Probability Generating Functions (PGF), and Supplementary Variable Techniques (SVT)), and the key features of each model. Notable early contributions include the work of Madan [4], Wang [5], Medhi [6], Choudhury et al. [7], Singh et al. [8], Choudhury and Deka [9], among others. These studies have addressed various complexities, such as server breakdowns, feedback mechanisms, retrial phenomena, batch arrivals, and different vacation policies. A detailed classification and comparison of these studies are provided in Table 1.

Table 1: *Literature on phase service up to 2013.*

Type of service	Author	Key features	Methodologies used
Second phase optional service	Madan [4]	The first essential is arbitrarily distributed but the optional service is exponential	LST, PGF, SVT
	Wang [5]	M/G/1 queue, server Breakdown, phase-dependent service	LST, PGF, SVT
	Medhi [6]	Essential as well as optional is generally distributed	LST, PGF, SVT
	Madan et al. [10]	Single server, optional service, Bernoulli vacation schedule	LST, PGF, SVT
	Madan and Baklizi [12]	Single arrival, optional additional service, optional re-service	LST, PGF, SVT
	Choudhury [12]	Single arrival, two-phase general distributed service	PGF, SVT.
	Jaraha and Madan [13]	$M/G/1$ model, optional service, general distributed service	LST, PGF, SVT
	Choudhury and Paul [14]	Two-phase queueing system, Bernoulli feedback to the tail of the queue	LST, PGF, SVT
	Atencia & Moreno [15]	$Geo/G/1$ model with retrial queue, second optional service	LST, PGF,SVT
	Choudhury[16]	Single arrival, Linear retrial policy, Bernoulli schedule vacation.	LST, PGF, SVT
	Choudhury et al.[17]	N-policy, two-phase service, delayed repair, linear cost structure.	LST, PGF, SVT
	Choudhury and Tadj [18]	Single arrival, generally distributed service, random breakdown, and delay repair	LST, PGF, SVT
	Ke and Chu [19]	M/G/1 Queueing system,<p, T> policy, second optional service, cost analysis to find optimal value of p & T.	PGV, SVT, LST
	Wang and Li [20]	Single arrival, general retrial time, two-phase service, server breakdown, and repair	PGF, SVT.
	Zadeh [21]	$Mx/(G_1, G_2)/1/G(BS)/V$, admissibility restricted.	PGV, SVT, LST
	Choudhury et al.[22]	Retrial queue, Bulk arrival, optional service, random failure, delay repair.	LST, SVT
	Jain and Agarwal [23]	Decrete time $Geo^X/G/1$, retrial queue, starting failure	LST, PGF, SVT
	Kumar and Arumuganathan [24]	Bulk arrival, retrial queue, two types of repair, Active breakdown	PGF, SVT.

	Maraghi et al.[25]	Bulk queue, second stage elective in nature, Bernoulli vacations, and random server failure	LST, SVT
	Ramnath and Kalidas [26]	Two-phase service, vacation, retrial policy, non-persistent customers.	LST,PGF,SVT
	Wang et al.[27]	Single arrival, second optional service, Comparison of randomized policy for (T,p) and (p, N) policy, and general start-up time using cost function.	PGF, SVT
	Singh et al.[28]	Bulk arrival, state-dependent arrival, and general distributed service.	LST,PGF,SVT
	Jain et al.[29]	Batch arrival, General distributed service, retrial, Bernoulli vacation, and server breakdown.	PGV, SVT, LST
	Jailaxmi et al.[30]	Batch arrival, Two Stage service, active server failure, and retrial	LST, PGF, SVT
Both phase essential	Kumar et al.[31]	Two-phase service, preemptive resume service, and retrial	PGF, SVT
	Madan and Swami [32]	M/D/1 queue with two-stage service and Heterogeneous vacation.	PGV,SVT, LST
	Choudhury and Paul [33]	Batch arrival, additional phase service, and N-policy	PGF, SVT
	Choudhury and Madan [34]	Two-phase, batch arrival, Bernoulli vacation	PGV, SVT, LST
	Choudhury and Madan [35]	Bulk queue, Modified Bernoulli vacation schedule, N-policy.	PGF, SVT
	Madan et al.[36]	Mx/(G_1G_2)/1/G(BS)/Vs, Two phase service and Vacation	PGV, SVT, LST
	Madan and Choudhury [37]	Single-server, two-phase heterogeneous service, Bernoulli schedule, general vacation time.	PGV, SVT, LST
	Choudhury and Paul [38]	Two-phase service, Bernoulli vacation, multiple vacation policy	PGV, SVT, LST
	Madan and Choudhury [39]	Mx/(G_1, G_2)/1 queue with restricted admissibility and random setup time	PGV, SVT, LST
	Choudhury et al [40]	Bulk queue, Modified Bernoulli vacation schedule	PGV, SVT, LST
	Kumar and Chandan [41]	N-policy, Batch service without gating, Single vacation, Cost analysis for optimal value of N.	PGF, SVT
	Dimitriou and Langaris [42]	Retrial queue, single vacation	PGF, SVT
	Choudhury and Kalita [43]	Two Phase with repeated attempts, Bernoulli Vacation	PGV, SVT
	Dimitriou and Langaris [44]	Two-phase service, retrial and start-up time, Random failure, Single vacation.	PGF, SVT
	Park et al.[45]	Two-phase service with fixed batch size policy, Linear cost analysis.	PGF, Rouché's theorem,
	Thangraj and Vanitha [46]	Two-phase service, compulsory vacation, and random failure	PGF, SVT
	Choudhury and Tadj [47]	Unreliable server, Bernoulli vacation, N-policy, delay repair. Linear cost structure.	PGV, SVT, LST
	Choudhury and Deka [48]	Single Poisson arrival, Bernoulli vacation, Unreliable server	PGV, SVT, LST

	Tadj et al.[49]	Two-phase, setup time, N-policy, Unreliable server, delay repair.	PGV, SVT, LST
	Singh et al.[50]	Bulk arrival, Two-phase service, m-phase of repair.	PGV, SVT, LST
	Vemuri et al. [51]	$M^X/M/1$ system, N policy, Unreliable server	PGF
	Salah [52]	Finite queueing system, Two-stage service	PGF
	Redenko [53]	Two-phase, finite buffer between phases, unreliable server	PGF, SVT
	Choudhury and Deka [54]	Bulk arrival, linear retrial policy, Bernoulli vacation, control admission policy.	PGF, SVT
	Dudina et al. [55]	Multi-server, retrial customer, Phase type service	PGF
Multi optional	Jianghua and Jinting [56]	Single arrival, general distributed service, feedback, balking, and unreliable server	PGF ,SVT
	Salehirad and Badamchizadeh [57]	Single server queue, multiphase service, general distributed service, and feedback to the tail of the queue	PGF, SVT
	Jain and Upadhyaya [58]	Optimal repairable, batch arrival, multi-optional service, Bernoulli vacation.	LST, PGF, SVT
	Wang and Li [59]	Bulk arrival, second service is multi-optional, unreliable server	PGF, SVT
	Jain et al.[60]	Unreliable server, second service, and vacation are multi-optional in nature.	LST, PGF, SVT
	Bhagat and Jain [61]	Unreliable server, Balking, Bulk arrival, Multi-optional second service	LST, PGF, SVT

II. Review of Literature (2014-2024)

This section reviews significant contributions to the study of queueing models with phase-type services published since 2014. The existing literature can be broadly categorized based on whether the service phases are optional, essential, or multi-optional. Research on queueing models with optional services has explored various complexities, including vacation policies, retrial mechanisms, and server reliability.

Two-phase optional service: Gao et al. [62] employed C0-semigroup theory to study the asymptotic stability of a queueing model by introducing an additional optional service. Rajadurai et al. [63] investigated a queueing model featuring customer retrials, group arrivals, and balking behavior, incorporating dual service choices and the possibility of re-service under an adapted vacation policy framework. Saravanarajan and Chandrasekaran [64] analyzed an M/G/1 queue featuring two types of service, Bernoulli vacations, feedback, and random server failure. The inclusion of random breakdowns makes the model applicable to manufacturing and cloud-based services. Kalidass and Kasturi [65] introduced an M/G/1 queueing model incorporating Bernoulli feedback and a dual-phase service structure, in which the initial phase is mandatory, whereas the subsequent phase is discretionary, involving a choice among various available service alternatives. Jiang et al. [66] examined a $M/G/1$ queue in which the server is subject to random failures during any service phase, and after a repair period that follows an exponential distribution, it resumes operation in phase i with probability q_i. Chakravarthy [67] further applied the matrix analytic method to a single-server queueing system in which customers, with certain predefined probabilities, could opt for either individual service or cooperative service. Arivudainambi and Godhandaraman [68] investigated a queueing model incorporating retrial customers, balking, and an elective second service, in which the server initiates a vacation each time the system becomes idle. Jain et al. [69] applied maximum entropy analysis to investigate a system characterized as a retrial queue with bulk arrivals, an optional second service, and a Bernoulli schedule vacation mechanism, thereby contributing to a more generalized approach to performance evaluation.

Bhagat and Jain [70], in their study of a G-queue with bulk arrivals, retrials, impatient customers, and multiple service choices, considered a server governed by the N-policy and subject to random failure. Singh et al. [71] introduced a queue model with bulk arrival, phased service, and variable arrival rates, assuming repairs occur in multiple optional phases. Vadivu and Arumuganathan [72] examined a queueing system featuring two service phases and offering the flexibility of both single as well as multiple vacation policies. By incorporating the Markovian Arrival Process (MAP), the model achieves greater adaptability in capturing diverse arrival patterns compared to traditional Poisson-based models. Wang et al. [73] conducted a systematic study of retrial queues with a two-phase service structure and a finite source population. Their analysis focused on the effects of customer retrials, highlighting the model's applicability to systems with a limited number of users, such as sensor networks and computing clusters.

Arivudainambi and Gowsalya [74] analyzed a single arrival retrial queue incorporating vacations under a Bernoulli schedule, dual service types, and server startup failures. The inclusion of startup failures makes the model applicable to unreliable manufacturing systems and cloud computing networks. Singh et al. [75] studied a batch arrival queue with state-dependent arrival rates, incorporating the concept of elective second service and vacation under a randomised policy. Singh et al. [76] extended this framework by developing a bulk queue model featuring two-phase service, customer balking, and an unreliable server, where repairs progress through m distinct phases.

Bagyam and Udya [77] examined a bulk queue model with state-dependent arrivals, two-phase service, and a retrial queue, where customer admission depends on the server's state. Ayyappan and Deepa [78] analyzed a queueing system in which both arrival as well as service are done in bulk, incorporating the features of optional services, multiple vacation policies, and setup times to improve system flexibility. Singh et al. [79] studied a bulk queue system incorporating optional services, vacation policies, and an unreliable server. Singh et al. [80] analyzed a G-queue with retrials, bulk arrivals, optional services, server unreliability, and delayed repairs, highlighting the impact of maintenance scheduling on queue performance.

Ayyappan and Nirmala [81] analyzed a bulk arrival queueing model incorporating an unreliable server, optional second-phase service, setup time, multiple vacation policies, and an N-policy. The model also features a closedown mechanism that allows the system to shut down during periods of low demand, enhancing its applicability to power-saving environments. Pikkala and Edadasari [82] investigated a Markovian queueing model that integrates variant working vacations, an optional second service phase, server failures occurring at random, and a mechanism for customer retention following reneging behavior. Madheswari and Suganthi [83], who examined a single-server retrial queue featuring two service phases, considered a model with an optional second service, server starting failures, and K types of Bernoulli vacation policies.

Abdollahi and Salehi Rad [84], who investigated a retrial queueing system with bulk arrivals, dual-phase service, and customer feedback, incorporated a server-dependent admission control mechanism in their model. Ayyappan and Thilagavathy [85] applied the matrix analytic method to investigate an advanced queueing model integrating a Markovian arrival process, standby servers, optional services, breakdowns, and customer impatience. Deepa and Azhagappan [86], who studied a bulk arrival queueing model with arrival rates dependent on the system state and incorporating multiple vacation policies, considered an optional second service and re-service mechanisms to improve the analysis of system performance. Laxmi and George [87], who analyzed a non-Markovian queueing model with batch service and an optional second service phase, examined the system's performance under both transient and steady-state conditions. More recently, Ayyappan and Gurulakshmi [88], who examined a Markovian arrival queueing model featuring distinct vacation types, interruptions governed by the N-policy, optional service phases, server breakdowns and repairs, setup delays, and customer discouragement behavior, provided a comprehensive analysis of system dynamics.

Two-phase essential service: Queueing models with two-phase service have been extensively studied, with researchers incorporating various complexities to enhance their applicability. Rajadurai et al. [89] investigated a queueing model featuring retrial, two-phase service, a revised vacation policy, and optional re-service, eliminating the need for orbiting. Saravanarajan and Chandrasekaran [90] expanded this framework by incorporating feedback from customers, where the server undergoes a mandatory vacation after completing the second-phase service for each customer. Ramaya et al. [91] investigated a queueing model featuring mandatory two-phase service under a gated service vacation policy. In their model, the first phase involves processing customers in batches, while the second phase delivers individual service to each customer within the batch. Van Do [92] obtained explicit solutions for the steady-state probabilities of a tollbooth queue with two distinct servers by applying eigenvalue analysis to the characteristic polynomial of the corresponding Markov chain. Rajadurai et al. [93] extended the study of retrial models by examining a batch arrival system with feedback in the retrial queue, incorporating negative customers, a Bernoulli vacation policy, and server breakdowns. In another study, they investigated a bulk queueing system featuring retrials and two-stage service, incorporating Bernoulli vacation policies and random service interruptions. Choudhury and Deka [94] conducted a study on a single Poisson arrival queueing system characterized by server unreliability, service in two stages, and a Bernoulli vacation scheme governed by a randomized vacation policy. In another work, they proposed a batch arrival model with an unreliable server, incorporating a Bernoulli vacation mechanism, two-phase service, and general retrial times to assess system performance. Building on this line of research, Choudhury and Deka [95] analyzed a queueing system with Poisson input in bulk, featuring two-stage service, random server failures, and a randomised vacation policy. Baumann and Sandmann [96] explored a tandem queue with multiple servers that deliver services in two stages under the assumption of finite buffer space, where blocking and loss can occur.

Arivudainambi and Gowsalya [97] introduced a non-Markovian retrial queueing model with a single server, featuring two distinct service types and a Bernoulli vacation policy. Later, Ayyappan et al. [98] investigated a bulk queue system with two-phase distinct service and a standby server, where the standby server activates only when the primary server is under repair. Sivasamy and Peter [99] investigated a multi-server queue with Poisson arrival, where customers become impatient due to a leisurely service rate. Dahmane and Aissani [100] utilized Markov decision theory to enhance routing management within a queueing system featuring retrial customers and a two-stage service process. Meanwhile, Liu et al. [101] developed a matrix-geometric method to examine a stochastic model with three-dimensional continuous time, incorporating multiple simultaneous working vacations for selected servers. Wu and Yang [102] used the Matrix Geometric Method to optimize costs in a Markovian queueing system with two essential stages of service. Kumar & Jain [103] used the same approach to study an unreliable queueing system with two stages of service, considering both regular and working vacations. Radha et al. [104] adapted queueing models for cloud computing by introducing bulk arrival in a non-Markovian queueing system, multiple service stages, and a structured vacation mechanism. More recently, Raghavendran and Vidhya [105] examined queueing models featuring phase-type distributions and modulated arrival rates, assuming that fluctuations in arrival rate modulation impact the server's service rate. Expanding on this, Ayyappan and Thilagavathy [106] applied matrix-geometric methods to analyze a Queueing model with a Markovian arrival process, incorporating various factors such as heterogeneous service modes, standby servers, customer impatience, optional services, startup time, system breakdowns, and phase-type repairs.

Multiphase essential service: The queueing model with multiphase essential service has garnered significant attention from researchers due to its wide applicability in various expanding domains of queueing congestion. This model is particularly relevant in fields such as telecommunications, healthcare, manufacturing systems, and transportation, where efficient service mechanisms are crucial for minimizing delays and optimizing resource utilization. Lakshmi and Ramnath [107] applied the SVT to investigate a retrial queueing model with K choices in the second phase of essential service. Vignesh et al. [108] explored a bulk arrival with a general distributed service, featuring multistage service, constrained admissibility, response services, and three optional vacation types, with applications in production and manufacturing systems. Bhagat and Jain [109] analyzed a retrial queueing model with batch arrival and an unreliable server, integrating multiphase essential service and repair. The model differentiates between two customer types: priority and ordinary. Ahuja et al. [110] conducted a transient analysis of a queueing model with M phases of essential service under the assumption that the server is subject to random failure using the Runge-Kutta method. Their study assumes that the server is susceptible to failure during both vacation and busy periods.

Multi optional service: One of the key features of queueing models with phase service is the concept of multi-optional service**.** In this system, customers have the flexibility to choose from multiple service options within the same service center, allowing for a more dynamic and customized service experience. This feature enhances the efficiency of service systems by accommodating diverse customer needs and optimizing resource allocation. The queueing model with multi-optional service has emerged as a significant area of interest among researchers due to its broad applicability in various real-world scenarios. Zadeh [111] examined a queueing model in which units arrive in groups of random size and service is done in k phases and allowing feedback to the tail of the queue at each phase. Varalakshmi et al. [112] studied a queueing model featuring two service phases, retrial, feedback follows a Bernoulli distribution, exhaustive vacation periods, and starting letdowns. Sundari & Srinivasan [113] investigated optional service mechanisms in a queueing system with general distributed service time, where unit demand for the first optional or second optional service occurs after the completion of the essential one.

Singh & Kumar [114] explored a bulk arrival queueing model with Bernoulli vacations, finite optional services, and N-policy. The inclusion of multiple optional services enhances flexibility in customer service-oriented applications. Li et al. [115] utilized the supplementary variable technique to analyze a general queue with multi-stage operations, in which the server may take multiple vacations if the system is empty. Rajadurai et al. [116] explored cost optimization in general queue models incorporating retrials, multiple service phases, vacations, and server breakdowns. Kumar [117] presented a general queueing system with bulk arrival and an unreliable server that integrates optional services, vacation under a Bernoulli schedule, and customer impatient behavior. Vignesh et al. [118] explored a general queueing system with bulk arrival under the assumption of mandatory three-stage services, an optional fourth stage, service disruptions, and deterministic server vacations. Jain and Kaur [119] studied a bulk queue model with multi-phase essential service and feedback under the (p, N) policy. The server is prone to random failure during any service phase, and it goes on vacation as the system becomes empty. Abdollahi et al. [120] studied the sensitivity of a queueing system retrial queue with k-1 optional service phases following an essential service. Unsatisfied customers may request reservice at each phase, and the server can take a Bernoulli vacation after the k^{th} phase. Jain et al. [121] utilized the supplementary variable technique to analyze a batch arrival queue model with an unpredictable server and a randomized vacation policy. The model assumes that unsatisfied customers may rejoin the queue for service. Sangeetha et al. [122] investigated a queueing system with a retrial customer having S phases of essential service, each offering different optional services. The server experiences a breakdown immediately upon the arrival of a customer, which is negative in nature. Li and Liu [123] used the matrix-geometric method to analyze the performance of a general queueing model with a working vacation and multi-phase service, incorporating a Bernoullischeduled vacation interruption. Kumar [124] examined an unreliable queue model with delayed repair and a single vacation, assuming that the server provides service through different optional phases.

V. Conclusion

In the present study, we have provided a comprehensive review of queueing systems with phase service, highlighting their significance in modeling and analyzing real-world congestion scenarios across various industries. By classifying and examining different queueing models with phase service, we have proposed a unified framework for their analysis using a queue-theoretic approach. These models serve as valuable tools for system analysts, engineers, and managers in mitigating congestion, reducing blocking and delays, and enhancing overall system efficiency.

Our review specifically focuses on the supplementary variable technique (SVT) for solving non-Markovian single-server queueing models with different phase service structures. We have categorized and discussed key contributions in this domain, including M/G/1 queueing models with essential and optional multistage services, as well as models addressing service interruptions. Despite significant progress in recent years, there remains a need for further advancements in queueing models with phase service to better reflect real-world complexities, such as priority queues, server vacations, breakdowns, and working vacations. Future research should integrate these factors to develop more robust and fault-tolerant systems. We hope that this review provides valuable insights for both researchers and practitioners, facilitating the development of more efficient and cost-effective service and manufacturing systems where congestion management is a critical concern.

Acknowledgements

The corresponding author gratefully acknowledges the esteemed referee and the chief editor for their thoughtful feedback and valuable suggestions, which have greatly enhanced the quality of this work. Furthermore, the first author sincerely appreciates the Research and Development Cell of Patna University, Patna, for providing financial support through the Minor Research Project (Project No. RDC/MRP/26).

References

[1] Cox, D. R. (1955). The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Mathematical Proceedings of the Cambridge Philosophical Society, 51:* 433–441. https://doi.org/10.1017/S0305004100030437

[2] Chaudhry, M. L., & Templeton, J. G. C. (1983). A first course on bulk queues. John Wiley & Sons Inc.

[3] Borthakur, A., & Medhi, J. (1974). A queueing system with arrival and service in batches of variable size. *Cahiers du CERO, 16*: 117–126.

[4] Madan, K. C. (2000). An M/G/1 queue with second optional service. *Queueing Systems, 34*: 37–46.

[5] Wang, J. (2004). An M/G/1 queue with second optional service and server breakdowns. *Computers and Mathematics with Applications, 47*(10–11): 1713–1723.

[6] Medhi, J. (2002). A single server Poisson arrival queue with a second optional channel. *Queueing Systems, 42*(3): 239–242.

[7] Choudhury, G., Ke, J. C., & Tadj, L. (2009). The N-policy for an unreliable server with delaying repair and two phases of service. *Journal of Computational and Applied Mathematics, 231*(1): 349–364. https://doi.org/10.1016/j.cam.2009.02.101

[8] Singh, C. J., Jain, M., & Kumar, B. (2011). Queueing model with state-dependent bulk arrival and second optional service. *International Journal of Mathematics in Operational Research, 3*(3): 322–340. https://doi.org/10.1504/IJMOR.2011.040029

[9] Choudhury, G., & Deka, M. (2012). A single server queueing system with two phases of service subject to server breakdown and Bernoulli vacation. *Applied Mathematical Modelling, 36*(12): 6050–6060. https://doi.org/10.1016/j.apm.2012.01.047

[10] Madan, K. C., Abu-Dayyeh, W., & Saleh, M. F. (2002). A queue with second optional service and Bernoulli schedule server vacations. *Systems Science, 28*(3): 51–62.

[11] Madan, K., & Baklizi, A. (2002). An M/G/1 queue with additional second-stage service and optional re-service. *International Journal of Information and Management Sciences, 13*: 13–31.

[12] Choudhury, G. (2003). Some aspects of an M/G/1 queueing system with optional second service. *Top, 11*: 141–150. https://doi.org/10.1007/BF02578955

[13] Jaraha, J.A., & Madan, K.C. (2003). An M/G/1 Queue with Second Optional Service with General Service Time Distribution. *Information and Management System,*14(2):47-56.

[14] Choudhury, G., & Paul, M. (2005). A two-phase queueing system with Bernoulli feedback. *Journal of Information and Management Sciences, 16*(1): 35–52.

[15] Atencia, I., & Moreno, P. (2006). Geo/G/1 retrial queue with second optional service. *International Journal of Operational Research, 1*(4): 340–362. https://doi.org/10.1504/IJOR.2006.010209

[16] Choudhury, G. (2008). Steady-state analysis of an M/G/1 queue with linear retrial policy and two-phase service under Bernoulli vacation schedule. *Applied Mathematical Modelling, 32:* 2480–2489.

[17] Choudhury, G., Ke, J. C., & Tadj, L. (2009). The N-policy for an unreliable server with delaying repair and two phases of service. *Journal of Computational and Applied Mathematics, 231*(1): 349–364.

[18] Choudhury, G., & Tadj, L. (2009). A queue with two phases of service subject to the server breakdown and delayed repair. *Applied Mathematical Modelling, 33*(6): 2699–2709.

[19] Ke, J.-C., & Chu, Y.K. (2009). Notes of M/G/1 system under the <p, T> policy with second optional service. *Central European Journal of Operations Research, 17*(4): 425–431.

[20] Wang, J., & Li, J. (2009). A single server retrial queue with general retrial times and two-phase service. *Journal of Systems Science and Complexity, 22*: 291–302. https://doi.org/10.1007/s11424-009-9164-8

[21] Zadeh, A. B. (2009). An Mx/(G_1, G_2)/1/G(BS)/Vs with optional second service and admissibility restricted. *International Journal of Information Management,* 20: 305–316.

[22] Choudhury, G., Tadj, L., & Deka, K. (2010). A batch arrival retrial queueing system with two phases of service and service interruption. *Computers & Mathematics with Applications, 59*(1): 437–450. https://doi.org/10.1016/j.camwa.2009.06.021

[23] Jain, M., & Agarwal, S. (2010). A discrete-time retrial queueing system with starting failures and optional service. *International Journal of Operational Research, 8*(4): 428–457. https://doi.org/10.1504/IJOR.2010.034068

[24] Kumar, M. S., & Arumuganathan, R. (2010). An MX/G/1 retrial queue with two-phase service subject to active server breakdowns and two types of repair. *International Journal of Operational Research, 8(*3): 261–291. https://doi.org/10.1504/IJOR.2010.033540

[25] Maraghi, F. A., Madan, K. C., & Darby-Dowman, K. (2010). Batch arrival vacation queue with second optional service and random system breakdowns. *Journal of Statistical Theory and Practice, 4*: 137–153. https://doi.org/10.1080/15598608.2010.10411977

[26] Ramanath, K., & Kalidass, K. (2010). A two-phase service M/G/1 vacation queue with general retrial times and non-persistent customers. *International Journal of Open Problems in Computer Science and Mathematics, 3*(2): 175–185.

[27] Wang, K. H., Yang, D. Y., & Pearn, W. L. (2010). Comparison of two randomized policy M/G/1 queues with second optional service, server breakdown, and startup. *Journal of Computational and Applied Mathematics, 234*(3): 812–824. https://doi.org/10.1016/j.cam.2010.01.045

[28] Singh, C. J., Jain, M., & Kumar, B. (2011). Queueing model with state-dependent bulk arrival and second optional service. *International Journal of Mathematics in Operational Research, 3*(3): 322–340. https://doi.org/10.1504/IJMOR.2011.040029

[29] Jain, M., Sharma, G. C., & Sharma, R. (2012). A batch arrival retrial queuing system for essential and optional services with server breakdown and Bernoulli vacation. *International Journal of Internet and Enterprise Management, 8*(1): 16–45. https://doi.org/10.1504/IJIEM.2012.049868

[30] Jailaxmi, V., Arumuganathan, R., & Rathinasamy, A. (2013). An Mx/G/1 retrial queue with two-phase service under active server breakdowns, two types of repair and multiple vacations with N-policy. *International Journal of Operational Research, 18*(1): 35–61. https://doi.org/10.1504/IJOR.2013.055532

[31] Kumar, B. K., Vijayakumar, A., & Arivudainambi, D. (2002). An M/G/1 retrial queueing system with two-phase service and preemptive resume. *Annals of Operations Research, 113*: 61–79. https://doi.org/10.1023/A:1020901710087

[32] Madan, K. C., & Swami, H. M. (2002). Steady-state analysis of an M/D/1 queue with two stages of heterogeneous server vacations (M/D/G1, G1/1 queue). *Revista Investigación Operacional, 23*(2): 150–163.

[33] Choudhury, G., & Paul, M. (2004). A batch arrival queue with an additional service channel under N-policy. *Applied Mathematics and Computation, 156*(1): 115–130.

[34] Choudhury, G., & Madan, K. C. (2004). A two-phase batch arrival queueing system with a vacation time under Bernoulli schedule. *Applied Mathematics and Computation, 149:* 337–349.

[35] Choudhury, G., & Madan, K. C. (2005). A two-stage batch arrival queueing system with a modified Bernoulli schedule vacation under N-policy. *Mathematical and Computer Modelling, 42*(1–2): 71–85.

[36] Madan, K. C., Al-Rawi, Z. R., & Al-Nasser, A. D. (2005). On M^X/(G1,G2)/1/G(BS)/Vs vacation queue with two types of general heterogeneous service. *Journal of Applied Mathematics and Decision Sciences*, 2005(3): 123–135.

[37] Madan, K. C., & Choudhury, G. (2005). A single-server queue with two phases of heterogeneous service under Bernoulli schedule and a general vacation time. *International Journal of Information and Management Sciences,* 16(2): 1–16.

[38] Choudhury, G., & Paul, M. (2006). A two-phase queueing system with Bernoulli vacation schedule under multiple vacation policy. *Statistical Methodology, 3:* 174–185.

[39] Madan, K. C., & Choudhury, G. (2006). Steady-state analysis of an M^X/(G1, G2)/1 queue with restricted admissibility and random setup time. *International Journal of Information and Management Sciences, 17*(2): 33–56.

[40] Choudhury, G., Tadj, L., & Paul, M. (2007). Steady-state analysis of an MX/G/1 queue with two phases of service and Bernoulli vacation schedule under multiple vacation policy. *Applied Mathematical Modelling, 31*: 1079–1091.

[41] Kumar, V. V., & Chandan, K. (2008). Cost analysis of a two-phase MX/EK/1 queueing system with N-policy. *OPSEARCH, 45:* 155–174. https://doi.org/10.1007/BF03398811

[42] Dimitriou, I., & Langaris, C. (2008). Analysis of a retrial queue with two-phase service and server vacations. *Queueing Systems, 60*: 111–129. https://doi.org/10.1007/s11134-008-9089-2

[43] Choudhury, G., & Kalita, S. (2009). A two-phase queueing system with repeated attempts and Bernoulli vacation schedule. *International Journal of Operational Research, 5*(4): 392–407. https://doi.org/10.1504/IJOR.2009.0257

[44] Dimitriou, L., & Langaris, C. (2010). A repairable queueing model with two-phase service, start-up times, and retrial customers. *Computers & Operations Research, 37(7)*: 1181–1190. https://doi.org/10.1016/j.cor.2009.03.003

[45] Park, H. M., Kim, T. S., & Chae, K. C. (2010). Analysis of a two-phase queueing system with a fixed-size batch policy. *European Journal of Operational Research, 206*(1): 118–122. https://doi.org/10.1016/j.ejor.2010.02.005

[46] Thangaraj, V., & Vanitha, S. (2010). A queue with two-stage heterogeneous service, compulsory server vacation, and random breakdowns. *International Journal of Contemporary Mathematical Sciences, 5*(7): 307–322.

[47] Choudhury, G., & Tadj, L. (2011). The optimal control of an $M^X/G/1$ unreliable server queue with two phases of service and Bernoulli vacation schedule. *Mathematical and Computer Modelling, 54*(1–2): 673–688. https://doi.org/10.1016/j.mcm.2011.03.010

[48] Choudhury, G., & Deka, M. (2012). A single-server queueing system with two phases of service subject to server breakdown and Bernoulli vacation. *Applied Mathematical Modelling, 36*(6): 6050–6060. https://doi.org/10.1016/j.apm.2012.01.047

[49] Tadj, L., Choudhury, G., & Rekab, K. (2012). A two-phase quorum queueing system with Bernoulli vacation schedule, setup, and N-policy for an unreliable server with delaying repair. *International Journal of Services and Operations Management, 12(2)*: 139–164.https://doi.org/10.1504/IJSOM.2012.047103

[50] Singh, C. J., Jain, M., & Kumar, B. (2013). Analysis of queue with two phases of service and m phases of repair for server breakdown under N-policy. *International Journal of Services and Operational Management, 16*(3): 373–406. https://doi.org/10.1504/IJSOM.2013.056769

[51] Vemuri, V. K., Hari Prasad Boppana, V. S. N., & Kotagiri, C. (2011). Optimal strategy analysis of an N-policy two-phase $M^X/M/1$ queueing system with server startup and breakdowns. *OPSEARCH, 48*: 109–122. https://doi.org/10.1007/s12597-011-0046-1

[52] Salah, K. (2011). Analysis of a two-stage network server. *Applied Mathematics and Computation, 217*(23): 9635–9645. https://doi.org/10.1016/j.amc.2011.04.048

[53] Rudenko, I. V. (2012). Two-phase queuing system with unreliable servers. *Moscow University Mathematics Bulletin, 67*: 145–150. https://doi.org/10.3103/S002713221204002X

[54] Choudhury, G., & Deka, K. (2013). A batch arrival retrial queue with two phases of service and Bernoulli vacation schedule. *Acta Mathematica Applicatae Sinica, English Series, 29*: 15–34. https://doi.org/10.1007/s10255-007-7083-9

[55] Dudina, O., Kim, C., & Dudin, S. (2013). Retrial queuing system with Markovian arrival flow and phase-type service time distribution. *Computers & Industrial Engineering, 66*(2): 360–373. https://doi.org/10.1016/j.cie.2013.06.020

[56] Jianghua, L., & Jinting, W. (2006). An M/G/1 retrial queue with second multi-optional service, feedback, and unreliable server. *Applied Mathematics: A Journal of Chinese Universities, 21*: 252–262. https://doi.org/10.1007/s11766-003-0002-9

[57] Salehirad, M. R., & Badamchizadeh, A. (2009). On the multi-phase M/G/1 queueing system with random feedback. *Central European Journal of Operations Research, 17*: 131–139. https://doi.org/10.1007/s10100-008-0079-6

[58] Jain, M., & Upadhyaya, S. (2010). Optimal repairable Mx/G/1 queue with multi-optional services and Bernoulli vacation. *International Journal of Operational Research, 7*(1): 109–132. https://doi.org/10.1504/IJOR.2010.029520

[59] Wang, J. T., & Li, J. H. (2010). Analysis of the M[X]/G/1 queues with second multi-optional service and unreliable server. *Acta Mathematica Applicatae Sinica, English Series, 26*: 353–368. https://doi.org/10.1007/s10255-010-0001-6

[60] Jain, M., Sharma, G. C., & Sharma, R. (2013). Unreliable server M/G/1 queue with multi-optional services and multi-optional vacations. *International Journal of Mathematics in Operational Research, 5*(2): 145–169. https://doi.org/10.1504/IJMOR.2013.052458

[61] Bhagat, A., & Jain, M. (2013). Unreliable M[X]/G/1 retrial queue with multi-optional services and impatient customers. *International Journal of Operational Research, 17*(2): 248–273. https://doi.org/10.1504/IJOR.2013.053614

[62] Gao, C., Chen, X.-M., Zheng, F., & Zhu, G. (2014). Asymptotic stability of the M/G/1 queueing system with optional second service. *Applied Mathematical Modelling, 38*(19–20): 4705–4716. https://doi.org/10.1016/j.apm.2014.03.032

[63] Rajadurai, P., Saravanarajan, M. C., & Chandrasekaran, V. M. (2014). Analysis of an M[X]/(G1,G2)/1 retrial queueing system with balking, optional re-service under modified vacation policy and service interruption. *Ain Shams Engineering Journal, 5*(3): 935–950. https://doi.org/10.1016/j.asej.2014.02.003

[64] Saravanarajan, M. C., & Chandrasekaran, V. M. (2014). Analysis of M/G/1 feedback queue with two types of services, Bernoulli vacations and random breakdowns. *International Journal of Mathematics in Operational Research, 6*(5): 567–588. https://doi.org/10.1504/IJMOR.2014.064839

[65] Kalidass, K., & Kasturi, R. (2014). A two-phase service M/G/1 queue with a finite number of immediate Bernoulli feedbacks. *OPSEARCH, 51*: 201–218. https://doi.org/10.1007/s12597-013-0136-3

[66] Jiang, T., Liu, L., & Li, J. (2015). Analysis of the M/G/1 queue in multi-phase random environment with disasters. *Journal of Mathematical Analysis and Applications, 430*(2): 857–873. https://doi.org/10.1016/j.jmaa.2015.05.028

[67] Chakravarthy, S. R. (2016). Queueing models with optional cooperative services. *European Journal of Operational Research, 248*(3): 997-1008. https://doi.org/10.1016/j.ejor.2015.08.015

[68] Arivudainambi, D., & Godhandaraman, P. (2015). Retrial queueing system with balking, optional service, and vacation. *Annals of Operations Research, 229*: 67–84. https://doi.org/10.1007/s10479-014-1765-5

[69] Jain, M., Sharma, R., & Sharma, G. C. (2015). Maximum entropy analysis of bulk arrival retrial queue with second optional service and Bernoulli vacation. *International Journal of Industrial and Systems Engineering, 20*(3): 369–396. https://doi.org/10.1504/IJISE.2015.069918

[70] Bhagat, A., & Jain, M. (2016). N-policy for M^X/G/1 unreliable retrial G-queue with preemptive resume and multi-services. *Journal of the Operations Research Society of China, 4*: 437–459. https://doi.org/10.1007/s40305-016-0128-0

[71] Singh, C. J., Jain, M., & Kumar, B. (2016). Unreliable retrial queue with option of additional services and Bernoulli vacation. *Ain Shams Engineering Journal, 7:* 415–429. https://doi.org/10.1016/j.asej.2015.05.006

[72] Vadivu, A. S., & Arumuganathan, R. (2016). Analysis of MAP/G/1/N queue with two phases of service under single (multiple) vacation(s). *International Journal of Operational Research, 25*(1): 47–76. https://doi.org/10.1504/IJOR.2016.073251

[73] Wang, J., Wang, F., Sztrik, J., & Kuki, A. (2017). Finite source retrial queues with two phase service. *International Journal of Operational Research, 30(4)*: 421–440. https://doi.org/10.1504/IJOR.2017.087824

[74] Arivudainambi, D., & Gowsalya, M. (2017). Analysis of an M/G/1 retrial queue with Bernoulli vacation, two types of service and starting failure. *International Journal of Artificial Intelligence and Soft Computing, 6*(3): 222–249. https://doi.org/10.1504/IJAISC.2017.088889

[75] Singh, C. J., Kumar, B., & Kaur, S. (2018). MX/G/1 state dependent arrival queue with optional service and vacation under randomised policy. *International Journal of Industrial and Systems Engineering, 29*(2): 252–272. https://doi.org/10.1504/IJISE.2018.091902

[76] Singh, C. J., Jain, M., & Kaur, S. (2018). Performance analysis of bulk arrival queue with balking, optional service, delayed repair, and multi-phase repair. *Ain Shams Engineering Journal, 9(4*): 2067–2077. https://doi.org/10.1016/j.asej.2016.08.025

[77] Bagyam, E. A., & Chandrika, K. U. (2019). Fuzzy analysis of bulk arrival two-phase retrial queue with vacation and admission control. *Journal of Analysis, 27*: 209–232. https://doi.org/10.1007/s41478-018-0118-1

[78] Ayyappan, G., & Deepa, T. (2019). Analysis of batch arrival bulk service queue with additional optional service multiple vacation and setup time. *International Journal of Mathematics in Operational Research, 15*(1): 1–25. https://doi.org/10.1504/IJMOR.2019.101609

[79] Singh, C. J., Kaur, S., & Jain, M. (2019). Analysis of bulk queue with additional optional service, vacation and unreliable server. *International Journal of Mathematics in Operational Research, 14*(4): 517–540. https://doi.org/10.1504/IJMOR.2019.100738

[80] Singh, C. J., Kaur, S., & Jain, M. (2020). Unreliable server retrial G-queue with bulk arrival, optional additional service and delayed repair. *International Journal of Operational Research, 38*(1): 82–111. https://doi.org/10.1504/IJOR.2020.106362

[81] 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 Mathematics in Operational Research, 16*(1): 53–81. https://doi.org/10.1504/IJMOR.2020.104686

[82] Pikkala, V. L., & Edadasari, G. B. (2021). Variant working vacation Markovian queue with second optional service, unreliable server and retention of reneged customers. *International Journal of Mathematics in Operational Research, 19*(1): 45–64. https://doi.org/10.1504/IJMOR.2021.115447

[83] Madheswari, S. P., & Suganthi, P. (2021). Retrial queue with unreliable server and second optional service under K types of Bernoulli vacations. *International Journal of Services and Operations Management, 40(*4): 502–533. https://doi.org/10.1504/IJSOM.2021.120058

[84] Abdollahi, S., & Salehi Rad, M. R. (2022). Analysis of a batch arrival retrial queue with two-phase services, feedback, and admission. *Bulletin of the Iranian Mathematical Society, 48*: 791–804. https://doi.org/10.1007/s41980-021-00547-1

[85] Ayyappan, G., & Thilagavathy, K. (2022). Analysis of MAP/PH1E, PH2O/1 queue with standby server, second optional service, variant arrival rate, Bernoulli schedule vacation, impatient behaviour of customers, breakdown, essential and optional repair. *International Journal of Mathematics in Operational Research, 23*(2): 147–214. https://doi.org/10.1504/IJMOR.2022.127054

[86] Deepa, T., & Azhagappan, A. (2022). Analysis of state dependent M[X]/G(a, b)/1 queue with multiple vacation second optional service and optional re-service. *International Journal of Operational Research, 44*(2): 254–278. https://doi.org/10.1504/IJOR.2022.123393

[87] Laxmi, P. V., & George, A. A. (2023). Analysis of non-Markovian batch service queue with second optional service under transient and steady state domain. *International Journal of Mathematics in Operational Research, 24*(3): 387–407. https://doi.org/10.1504/IJMOR.2023.129496

[88] Ayyappan, G., & Gurulakshmi, G. A. A. (2024). Analysis of MAP/PH/1 queue with differentiated vacation, vacation interruption under N-policy, optional service, breakdown, repair, setup and discouragement of customers. *International Journal of Mathematics in Operational Research, 27*(4): 415–457. https://doi.org/10.1504/IJMOR.2024.138463

[89] Rajadurai, P., Saravanarajan, M. C., & Chandrasekaran, V. M. (2014). Analysis of an M[X]/(G_1, G_2)/1 retrial queueing system with balking, optional re-service under modified vacation policy and service interruption. *Ain Shams Engineering Journal, 5*(3): 935–950. https://doi.org/10.1016/j.asej.2014.02.003

[90] Saravanarajan, M. C., & Chandrasekaran, V. M. (2014). Analysis of MX_XX/G/1 feedback queue with two-phase service, compulsory server vacation, and random breakdowns. *OPSEARCH, 51*: 235–256. https://doi.org/10.1007/s12597-013-0141-6

[91] Ramaya, K., Palaniammal, S., & Vijyalakshmi, C. (2015). Design and analysis of a vacation model for two-phase queueing system with gated service. *Procedia Computer Science, 50*: 301–306. https://doi.org/10.1016/j.procs.2015.04.044

[92] Van Do, T. (2015). A closed-form solution for a tollbooth tandem queue with two heterogeneous servers and exponential service times. *European Journal of Operational Research, 247(2*): 672–675. https://doi.org/10.1016/j.ejor.2015.05.044

[93] Rajadurai, P., Chandrasekaran, V. M., & Saravanarajan, M. C. (2015). Steady-state analysis of batch arrival feedback retrial queue with two phases of service, negative customers, Bernoulli vacation, and server breakdown. *International Journal of Mathematics in Operational Research, 7*(5): 519–546. https://doi.org/10.1504/IJMOR.2015.071276

[94] Choudhury, G., & Deka, M. (2015). An M/G/1 unreliable server queue with two phases of service and Bernoulli vacation schedule under randomised vacation policy. *International Journal of Applied Management Science, 7*(4): 309–337. https://doi.org/10.1504/IJAMS.2015.073198

[95] Choudhury, G., & Deka, M. (2016). A batch arrival unreliable server queue with two phases of service and Bernoulli vacation schedule under randomised vacation policy. *International Journal of Services and Operations Management, 24*(1): 33–72. https://doi.org/10.1504/IJSOM.2016.075763

[96] Baumann, H., & Sandmann, W. (2017). Multi-server tandem queue with Markovian arrival process, phase-type service times, and finite buffers. *European Journal of Operational Research, 256*: 187–195. https://doi.org/10.1016/j.ejor.2016.07.035

[97] Arivudainambi, D., & Gowsalya, M. (2018). A single server non-Markovian retrial queue with two types of service and Bernoulli vacation. *International Journal of Operational Research, 33*(1): 55–81. https://doi.org/10.1504/IJOR.2018.094231

[98] Ayyappan, G., Nirmala, M., & Karpagam, S. (2020). Analysis of repairable single server bulk queue with standby server, two-phase heterogeneous service, starting failure, and multiple vacation. *International Journal of Applied and Computational Mathematics, 6*: 52. https://doi.org/10.1007/s40819-020-00805-6

[99] Sivasamy, R., & Peter, P. O. (2021). A c-server Poisson queue with customer impatience due to a slow-phase service. *International Journal of Mathematics in Operational Research, 20*(1): 85–98. https://doi.org/10.1504/IJMOR.2021.117632

[100] Dahmane, Z., & Aissani, A. (2021). Optimal routing control of a retrial queue with two-phase service. *International Journal of Mathematics in Operational Research, 20*(1): 20–40. https://doi.org/10.1504/IJMOR.2021.117631

[101] Liu, F., Ma, Z., Si, Q., & Yan, M. (2021). Performance analysis of peer-to-peer networks based on two-phase service queuing theory. *International Journal of Communication Networks and Distributed Systems, 27*(4): 349–365. https://doi.org/10.1504/IJCNDS.2021.119215

[102] Wu, C. H., & Yang, D. Y. (2021). Bi-objective optimization of a queueing model with two-phase heterogeneous service. *Computers & Operations Research, 130*: 105230. https://doi.org/10.1016/j.cor.2021.105230

[103] Kumar, A., & Jain, M. (2022). Cost optimization of an unreliable server queue with two-stage service process under hybrid vacation policy. *Mathematics and Computer Simulations, 204*: 259–281. https://doi.org/10.1016/j.matcom.2022.08.007

[104] Radha, S., Maragathasundari, S., & Swedheetha, C. (2023). Analysis on a non-Markovian batch arrival queuing model with phases of service and multi vacations in cloud computing services. *International Journal of Mathematics in Operational Research, 24*(3): 425–449. https://doi.org/10.1504/IJMOR.2023.129489

[105] Raghavendran, R., & Vidhya, V. (2024). Single server queueing model of Markov modulated bursty arrival with service following phase type distributions for underlay and overlay customers. *Ain Shams Engineering Journal, 15*(8): 102825. https://doi.org/10.1016/j.asej.2024.102825

[106] Ayyappan, G., & Thilagavathy, K. (2024). Analysis of MAP/PH(1), PH(2), PH(3)/1 queueing system with two modes of heterogeneous service, standby server, vacation, impatient behaviour of customers, additional service, start-up time, breakdown and phase type repairs. *International Journal of Operational Research, 49*(1): 85–130. https://doi.org/10.1504/IJOR.2024.136008

[107] Lakshmi, K., & Ramanath, K. (2014). An M/G/1 two-phase multi-optional retrial queue with Bernoulli feedback, non-persistent customers, and breakdown and repair. *International Journal of Operational Research, 19*(1): 78–95. https://doi.org/10.1504/IJOR.2014.057844

[108] Vignesh, P., Srinivasan, S., & Sundari, S. M. (2017). Analysis of non-Markovian batch arrival queueing model with multi-stages of service of restricted admissibility, feedback service, and three optional vacations in production and manufacturing. *International Journal of Mathematics in Operational Research, 11*(3): 285–309. https://doi.org/10.1504/IJMOR.2017.087199

[109] Bhagat, A., & Jain, M. (2020). Retrial queue with multiple repairs, multiple services, and non-preemptive priority. *OPSEARCH, 57*: 787–814. https://doi.org/10.1007/s12597-020-00443-y

[110] Ahuja, A., Jain, A., & Jain, M. (2022). Transient analysis and ANFIS computing of unreliable single server queueing model with multiple stage service and functioning vacation. *Mathematics and Computers in Simulation, 192*: 464–490. https://doi.org/10.1016/j.matcom.2021.09.011

[111] Zadeh, A. B. (2015). A batch arrival multi-phase queueing system with random feedback in service and single vacation policy. *OPSEARCH, 52*: 617–630. https://doi.org/10.1007/s12597-015-0206-9

[112] Varalakshmi, M., Rajadurai, P., Saravanarajan, M. C., & Chandrasekaran, V. M. (2016). An M/G/1 retrial queueing system with two phases of service, immediate Bernoulli feedbacks, single vacation, and starting failures. *International Journal of Mathematics in Operational Research, 9*(3): 302–328. https://doi.org/10.1504/IJMOR.2016.078823

[113] Sundari, S. M., & Srinivasan, S. (2017). Optional services in a non-Markovian queue. *International Journal of Operational Research, 28*(1): 1–17. https://doi.org/10.1504/IJOR.2017.080593

[114] Singh, C. J., & Kumar, B. (2017). Bulk queue with Bernoulli vacation and m-optional services under N-policy. *International Journal of Operational Research, 30*(4): 460–483. https://doi.org/10.1504/IJOR.2017.087826

[115] Li, J., Liu, L., & Jiang, T. (2018). Analysis of an M/G/1 queue with vacations and multiple phases of operation. *Mathematical Methods of Operations Research, 87*: 51–72. https://doi.org/10.1007/s00186-017-0606-0

[116] Rajadurai, P., Saravanarajan, M. C., & Chandrasekaran, V. M. (2018). Cost optimization analysis of retrial queue with K optional phases of service under multiple working vacations and random breakdowns. *International Journal of Industrial and Systems Engineering, 29*(2): 193–222. https://doi.org/10.1504/IJISE.2018.091900

[117] Kumar, B. (2018). Unreliable bulk queueing model with optional services, Bernoulli vacation schedule, and balking. *International Journal of Mathematics in Operational Research, 12*(3): 293–316. https://doi.org/10.1504/IJMOR.2018.090799

[118] Vignesh, P., Srinivasan, S., & Sundari, S. M. (2019). Analysis of a non-Markovian single-server batch arrival queueing system of compulsory three stages of services with fourth optional stage service, service interruptions, and deterministic server vacations. *International Journal of Operational Research, 34*(1):28–53. https://doi.org/10.1504/IJOR.2019.096937

[119] Jain, M., & Kaur, S. (2020). (p,N)-Policy for unreliable server bulk queue with Bernoulli feedback. *International Journal of Applied and Computational Mathematics, 6:* 170. https://doi.org/10.1007/s40819-020-00912-4

[120] Abdollahi, S., Rad, M. R. S., & Farsi, M. A. (2021). Reliability and sensitivity analysis of retrial queue with optional k-phase services, vacation, and feedback. *Iranian Journal of Science and Technology, Transactions of Science, 45:* 1361–1374. https://doi.org/10.1007/s40995-021-01101-8

[121] Jain, M., Kaur, S., & Singh, M. (2021). Unreliable server queue with balking, optional service, Bernoulli feedback, and vacation under randomized policy. In P. K. Kapur, G. Singh, & S. Panwar (Eds.), *Advances in interdisciplinary research in engineering and business management*. Asset Analytics. Springer, Singapore. https://doi.org/10.1007/978-981-16-0037-1_24

[122] Sangeetha, N., Ebenesar Anna Bagyam, J., & Udayachandrika, K. (2022). Performance analysis of retrial queueing system in wireless local area network. In D. Giri, J. K. Mandal, K. Sakurai, & D. De (Eds.), *Proceedings of International Conference on Network Security and Blockchain Technology (ICNSBT 2021)* (Vol. 481, *Lecture Notes in Networks and Systems*). Springer, Singapore. https://doi.org/10.1007/978-981-19-3182-6_10

[123] Li, J. J., & Liu, L. W. (2023). The GI/M/1 queue in a multi-phase service environment with working vacations and Bernoulli vacation interruption. *Journal of the Operations Research Society of China, 11*: 627–656. https://doi.org/10.1007/s40305-021-00371-8

[124] Kumar, B. (2024). Unreliable queue model with multi-phases of services, delay repair, and single vacation. *International Journal of Industrial and Systems Engineering, 48*(1): 100–124. https://doi.org/10.1504/IJISE.2024.140682