Queueing Models and
Service Management
Vol. 3, No. 2, page 203-234, 2020

QMSM


# Performance Analysis of a Discrete-Time Queue with Versatile Batch Transmission Rule Under Batch Size Sensitive Policy

Arunava Maity, U. C. Gupta*, and Nitin Kumar

Department of Mathematics
Indian Institute of Technology Kharagpur
Kharagpur-721302, India
*(Received October 2019 ; accepted May 2020)*

**Abstract:** This paper considers a single channel, infinite buffer, batch transmission system in a slotted time set up. Packets arrive according to the Bernoulli process and the transmission time is arbitrarily distributed and depends on the number of packets undergoing transmission. The service is provided according to versatile service policy i.e., the server can decide the number of packets (threshold bound) to be transmitted on beforehand with certain probability. Present study may greatly help in understanding the related performances of batch transmission channels which allows packets of variable size with the transmission time depending on the size of the packets undergoing transmission. We obtain the distribution of the number of packets in the queue and also with the server at post transmission epochs in a tractable and presentable form. We use probability generating function approach for our analysis. The joint distribution of the state probabilities essentially provides the necessary information about the busy states of the system. Furthermore, we establish the relationships between arbitrary and post transmission epoch state probabilities. Next, we also obtain some essential marginal distributions, performance characteristics and develop a related cost structure for the present model, which may be very useful to the vendors for optimal utilization of the present system by possibly controlling only the sensitive parameters in pre-implementation stage. We also present some illustrative numerical examples to accomplish our study.

**Keywords:** Batch service, cost analysis, discrete-time channel, embedded Markov chain, generating function, optimal control.

## 1. Introduction

In recent years, discrete-time service systems have taken abound attention of the researchers mainly due to its compatibility to model the digital communication system rather than continuous-time counterpart. Also, due to the slotted nature of the time axis, performances and related behaviors (quality of service (QoS) etc.) of packet switching or circuit switching networks, hybrid multiplexing, cellular base stations (femtocell etc.), internet protocol or ethernet that use variable sized packets and frames protocols, ATM multiplexer in the broadband integrated services digital network (B-ISDN), circuit-switched time-division multiple access (TDMA) systems etc., are better described by the modeling and analysis of an analogous queueing or service systems in discrete-time set up. More application of such

*Corresponding author
Email : umesh@maths.iitkgp.ac.in

queueing or service systems can be found in [5], [2], [22], [18].

The modern telecommunication/wireless networks (5th generation mobile networks(5G), 4G/ LTE, 3GPP LTE-A etc.) are primarily intended for transferring information of some kinds e.g., voice, video, data, signals, writings, images, sounds or intelligence of any nature etc., which we usually refer as messages. Before the transmission begin, the messages/data are broken into small, manageable packets, each bundled up with the necessary information (cells) and then are transmitted to the proper destination over the network consisting of several nodes by establishing the virtual connection. In queueing analogy, each node behaves like a single/ multiple server queue with finite/infinite waiting space/buffer.

An individual node, in such facsimile systems behaves like a multiserver or as a sole server offering batch transmission. Usually batch transmission in such systems occurs due to the simultaneous transfer of multiple packets (may be with different bandwidth demands). As an example, Point-to-multi point (P2M) connections, which are most frequently used in wireless communications, are possible over multi drop links e.g., a mainframe and its terminals. The device which is responsible for multi point connection is usually an intelligent controller that manages the flow of information to attached devices. It is generally seen that these P2M connections require lower investment cost as compared to point-to-point (P2P) communications. Also, in modern mobile telecommunication 5G, numerous simultaneous connections (within some threshold limits) is to be supported for massive sensor deployments. Also it offers huge data transfer rates up to 1 gigabit per second to multiple clients depending on the number of clients at a time with a improved latency as compared to LTE. Another popular example of batch service in telecommunications, is femtocell (a subset of small cell) is a small, low-power, user installed, cellular base station, generally designed for indoor coverage or small enterprizes. It connects to the service providers network via broadband (such as DSL or cable). It currently support four to eight active mobile phones in a residential setting, and eight to sixteen active mobile phones in enterprize settings. Femtocells have the ability to provide services to huge number of mobile devices (within certain threshold limits) and make transmission depending on the size of the transmitting batch. Femtocells benefits both the mobile operator and the consumer especially in terms of coverage and quality. So, the present telecommunication systems discussed here can be modeled and analyzed by an analogous discrete-time batch service queue. It is quite perceptible that the transmission of packets in batches with varying size enhances the efficiency of the system as well as it is cost-effective too. Furthermore, there are other telecommunication systems which behaves like a batch service queue e.g., IEEE 802.16, IEEE 802.11n etc. These systems include a Base Station (BS) and one or more Subscriber Stations (SS). Depending on the bandwidth demand BS allocates variable number of physical slots to each SS. Now data transfer takes place from BS to SS and SS to BS using Time Division Multiple Access (TDMA) within certain threshold of size of the batches and depending on the size of transmission batches. So, from the above discussion, it may be concluded that the several important telecommunication systems needs to be modeled as discrete-time batch service queue with batch size sensitive transmission rates for efficient performance analysis.

In Fig 1 we display a virtual diagram of a general P2M architecture. Thus it is ob-

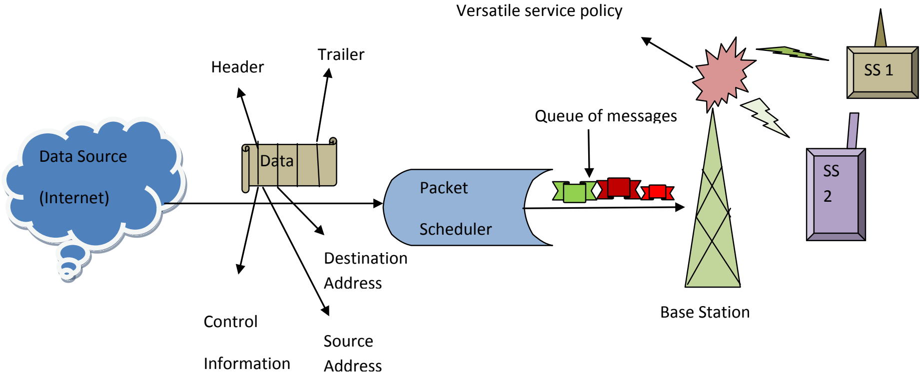


Figure 1. Point to multi point architecture for base station.

served that many telecommunication systems are based on batch processing control (i.e., here the server processes in batch rather individually) as it improves quality of service, as well as, adds flexibility to the system. However, it has been a major concern to the vendors that how to achieve the most cost-effective combination of performance or reliability of the system output. So, there is a need to model an acceptable standard representation of such batch processing system in slotted time setup. Also, it has to be kept in mind that, in modern age the server utilization and processing rate should also be met together for maximum response. This motivates us to model an infinite-buffer discrete-time batch processing transmission channel where the transmission rates depend on the number of packets undergoing transmission.

We choose Bernoulli arrival process and arbitrarily distributed service time for the batches in the present study. Also, we use the versatile batch service rule to set up the service policy. In this rule, the server is given a chance in beforehand to select the number of packets randomly for transmission following certain distribution. According to this policy, on completion of a batch transmission if the server finds $m \geq 0$ packets waiting in the queue, then (**i**) When $m < a$, the server remains '$idle$' until at least number '$a$' packets are accumulated, and the server takes the batch of size '$a$' packets for the transmission, according to the discrete random variable $Y$ with support $[a, B]$. (**ii**) When $m \geq a$, the server takes a batch of size $i(= min\{m, Y\})$ packets for the transmission, according to the discrete random variable $Y$ with support $[a, B]$.

Since in the most of the application the server is finitely capacitated, so we use the service batch-size rule with a finite support. This rule is also called the '$(a, Y)$' rule. The reasons that drive us toward present assumptions are as follows:

- In modern day communication system, large number of data in the form of packets are processed on a regular basis so, the waiting space needs to be large enough. So, we opt for infinite waiting space despite of its analysis being not quite easy.
- Bernoulli input process is quite tractable and well suited for the arrivals admitting no fractal characteristics like self similarity etc.

- In general, the processing times of batches do not follow certain well known distribution so the general service time distribution is assumed here.
- '$(a, Y)$' rule is one of the most general and promising batch service policy available in the literature.
- The present model unifies almost all the results of the earlier authors who used Bernoulli input with batch transmission.
- We use batch size dependent transmission rates which essentially reduces congestion and improves productivity of the system.

Keeping in view the myriad of literature on the discrete-time service systems with batch processing, one can categorize the literature in the following way.

1. (a) Late Arrival Delayed Access (LAS-DA) systems (b) Early Arrival Systems (EAS)
2. (a) With finite-buffer (b) with infinite-buffer
3. (a) With batch size dependent service (b) Without batch size dependent service
4. (a) Geometric service time (b) non-Geometric service time.
5. (a) Versatile batch service rule (with thresholds, minimum threshold preferably being different from 1) (b) other batch service rule e.g., 'General batch service rule' or '$(L, K)$'rule, 'Fixed batch service rule' etc.

It is easily noticeable that the most of the literature deals with (1a), (2a), (3b), (4a), (5b) while, there are very few papers that includes (2b), (3a) and (5a), for some recent references see [3], [6], [7], [9], [10], [11], [12], [13], [14], [16], [15], [17], [19], [20], [21], [23], [10], [4]. Chaudhury and Chang [8] has analyzed $Geo/G^Y/1/N + B$ system, wherein they obtained post transmission epoch state probabilities using embedded Markov chain technique and established a relationship amongst the random and pre-arrival epoch probabilities using renewal theoretic arguments. Recently, [23] has considered $Geo/G^{(a,Y)}/1/K$ system and obtained post transmission state epoch probabilities using embedded Markov chain technique and established the relationship with the random epoch state probabilities using "rate-in-rate out" principle. They both have considered the systems in finite buffer along with unbiased service rates. In this paper, we consider the $Geo/G_n^{(a,Y)}/1$ system which essentially includes the infinite-buffer case and batch size dependent service rates. Also, since the detailed description of the busy states are very essential for the vendors, we have successfully obtained the joint distribution of both server content and the queue content in a quite tractable and presentable form, through the inversion of the the bi-variate probability generating function (b.p.g.f.) associated with it, despite it needs an extensive mathematical involvement. Some useful performance measures like 'mean number of packets awaiting for transmission', 'average number of packets with the server', 'mean waiting time', 'proportion of the packets that have immediate access to transmission channel at post transmission epochs' etc., are also obtained. Further, we develop a cost structure for the present system, which may help the vendors to optimize the system cost according to his will on beforehand by adjusting the system parameters. It may be also admitted that, the procedures adopted here may be further utilized to explore more complex facsimile systems. Throughout the paper the term 'service' is used to refer the transmission through the channel by the server.

This paper is organized as follows: Section 2 provides the model description; in section 3 joint distribution of number of packets in the queue and with the server is obtained, marginal distributions and some important performance measures are presented in section 4 and in section 5 a cost model is developed. Numerical results, optimality analysis and conclusion are presented in section 6, 7 and 8, respectively, followed by references.

## 2. Model Description

For the present system, we assume that the packets arrive according to a Bernoulli process with parameter $\lambda$, are transmitted in groups through a single channel, according to the versatile batch-service rule or '$(a, Y)$' rule, where $Y$ is the random variable corresponding to the serving batch size with the following distribution,

$$P(Y = i) = \begin{cases} y_i, & i = a, a+1, a+2, \ldots B; \\ 0, & \text{elsewhere}, \end{cases} \tag{1}$$

where $y_i$, is the probability that the server takes $i$ packets for transmission at the beginning of service and $B$ is the maximum serving capacity of the server, and mean batch size ($\overline{y}$) is given by $\overline{y} = \sum_{i=a}^{B} i y_i$.

Further we assume that, the service time of the batch is arbitrarily distributed and is dependent on the number of packets in the batch. Let $S_r$, $a \leq r \leq B$ denotes the random variables for the service time corresponding to the batch of size $r$ packets, which are assumed to be independent and its probability mass function (p.m.f.) is given by, for $a \leq r \leq B$, $s_r(n) = P(S_r = n)$, $n = 1, 2, \ldots$. We then have $\sum\limits_{n=1}^{\infty} s_r(n) = 1$. The corresponding generating function for the service time of a batch of size $r$ ($a \leq r \leq B$) is given by $s_r^*(z) = \sum\limits_{n=1}^{\infty} s_r(n) z^n$. Let $s_r = \frac{1}{\mu_r} = s_r^{*(1)}(1)$ is the mean service time of a batch of size $r$, where $s_r^{*(1)}(1)$ is the first order derivative of $s_r^*(z)$ evaluated at $z = 1$. Arriving packets are not allowed to join the ongoing transmission even if there is an unused service capacity of the server. It is assumed that, the successive service capacities, service times and the inter arrival times are mutually independent sequences of random variables. For sake of convenience, we denote our model mathematically by $Geo/G_n^{(a,Y)}/1$, where $n$ indicates that the service time of the batch depends on the number of packets in the batch.

For the service system in discrete-time set up, the time axis is divided into a sequence of equal intervals of unit length, called slots, and it is assumed that inter-arrival and service times are integral multiples of the slots. Under this discrete-time setting, an arrival and a transmission may take place simultaneously at a slot boundary. For the present model we assume late arrival delayed access system (LAS-DA), where the packets are assumed to arrive late during a slot, just prior to the end of a slot, while services are assumed to begin and end only at the slot boundaries i.e., if the time axis be marked by $0, 1, 2, ..., k, ....,$ an arrival occurs in $(k-, k)$ and a transmission takes place in $(k, k+)$ (see Fig 2).

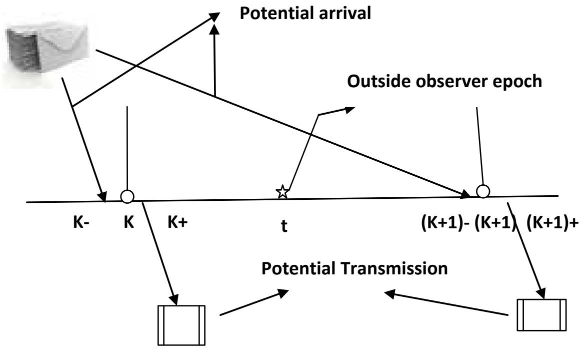


Figure 2. Various epochs for LAS-DA Systems.

Although, in information transmission systems, actual arrivals may take place at random time during a slot but usually, these arrivals are measured at slot boundaries, while the service attunes to slot boundaries. Evidently, our present model complies well and good with such kind of communication systems.

The present model, from a possible application point of view may be well understood from the Fig 3. Here we pictorially describe how the packets arrive at a service station from the sources and are transmitted through the server to the destined recipients with the present service policy.

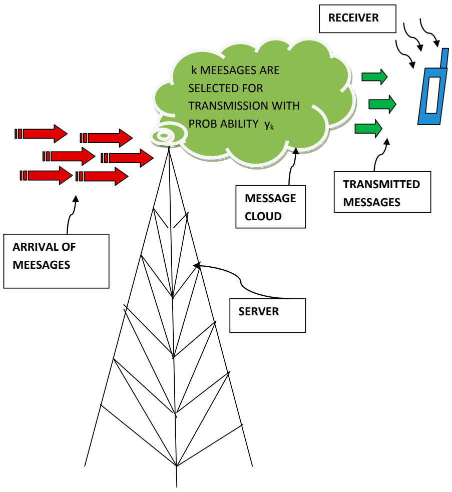


Figure 3. Sample description of the present model.

## 3. Joint Distribution of the Number of Packets in the Queue and with the Server at Post Transmission Epoch

In this section, we obtain the joint distribution of the number of packets in the queue and with the server at post transmission epoch using generating function approach (under the stability condition of the system).

**Theorem 1.** *The present queueing system is stable iff the traffic intensity* $\rho = \dfrac{\lambda \sum_{i=a}^{B} \dfrac{y_i}{\mu_i}}{\bar{y}} < 1.$

*Proof.* The proof is very straight forward and follows from the definition of the traffic intensity. Also the stability condition can be perceived from Theorems 3 and 4 of the Appendix.

**Remark 1.** *The stability of the queue can be proved using the result given in Abolnikov and Dukhovny [1]. If we consider the TPM of only queue length distribution at departure epoch of a batch, the TPM will be* $\Delta_{m,n}$ *type matrix. Then using the stability condition given in [1], we conclude that the corresponding Markov chain is ergodic iff* $K'(1) < B$. *Here* $K(x) = \sum_{j=a}^{B} y_j s_j^*(\overline{\lambda} + \lambda x) x^{B-j}$. *Now to show the stability of the queueing system, it is sufficient to show that* $K'(1) < B \Leftrightarrow \rho < 1$, *which is given in Theorem 3 in Appendix.*

Purpose of this section is to provide a clear insight and description of the busy and idle states of the system. It is perceptible from the following analysis that the construction of the b.p.g.f. (as the current state vector is two-dimensional) and extraction of the state probabilities in a presentable form, through the inversion of the b.p.g.f. is quite mathematically involving.

For the present system (with LAS-DA), we observe the state of the system just prior to the beginning of a slot boundary. Let us define the following random variables just prior to the potential arrival (i.e. at $k-$):

- $N(k-)$:= Number of the packets in the queue waiting to be transmitted.
- $J(k-)$:= Number of packets with the server.
- $U(k-)$:= Remaining service time of the batch in service (if the server is busy).

Then the joint probabilities can be defined as,

$$p_{n,0}(k-) = Pr\{N(k-) = n, J(k-) = 0\}, \qquad 0 \le n \le a-1, \tag{2}$$

$$p_{n,r}(u, k-) = Pr\{N(k-) = n, J(k-) = r, U(k-) = u\}, \quad u \ge 1,\ n \ge 0,\ a \le r \le B. \tag{3}$$

In steady-state, we write

$$\begin{aligned} p_{n,0} &= \lim_{k- \to \infty} p_{n,0}(k-), \qquad 0 \le n \le a-1, \\ p_{n,r}(u) &= \lim_{k- \to \infty} p_{n,r}(u, k-), \qquad u \ge 1,\ n \ge 0,\ a \le r \le B. \end{aligned} \tag{4}$$

### *3.1. Governing equations*

Now observing the states of the system just prior to the beginning of the slot boundary, i.e., at the time epochs $k-$ and $(k+1)-$, we write the following difference equations in steady-state (for $u \geq 1$).

$$p_{0,0} = \overline{\lambda} p_{0,0} + \overline{\lambda} \sum_{i=a}^{B} p_{0,i}(1), \tag{5}$$

$$p_{n,0} = \overline{\lambda} p_{n,0} + \overline{\lambda} \sum_{r=a}^{B} p_{n,r}(1) + \lambda p_{n-1,0} + \lambda \sum_{r=a}^{B} p_{n-1,r}(1), \quad 1 \leq n \leq a-1, \tag{6}$$

$$p_{0,a}(u) = \overline{\lambda} p_{0,a}(u+1) + \overline{\lambda} s_a(u) \sum_{r=a}^{B} p_{a,r}(1) \sum_{j=a}^{B} y_j + \lambda s_a(u) \sum_{r=a}^{B} p_{a-1,r}(1) \sum_{j=a}^{B} y_j + \lambda p_{a-1,0} s_a(u), \tag{7}$$

$$p_{0,r}(u) = \overline{\lambda} p_{0,r}(u+1) + \overline{\lambda} \sum_{i=a}^{B} p_{r,i}(1) \sum_{j=r}^{B} y_j s_r(u) + \lambda \sum_{i=a}^{B} p_{r-1,i}(1) \sum_{j=r}^{B} y_j s_r(u), \quad a+1 \leq r \leq B, \tag{8}$$

$$p_{n,r}(u) = \overline{\lambda} p_{n,r}(u+1) + \overline{\lambda} \sum_{i=a}^{B} p_{n+r,i}(1) y_r s_r(u) + \lambda \sum_{i=a}^{B} p_{n+r-1,i}(1) y_r s_r(u) + \lambda p_{n-1,r}(u+1), \quad n \geq 1, a \leq r \leq B, \tag{9}$$

where $\overline{\lambda} = 1 - \lambda$.
Next we define the probability generating function of $p_{n,r}(u)$ as

$$p^*_{n,r}(z) = \sum_{u=1}^{\infty} p_{n,r}(u) z^u, \quad |z| \leq 1, \text{ and } p_{n,r} := p^*_{n,r}(1). \tag{10}$$

The normalizing condition is given by

$$\sum_{n=0}^{a-1} p_{n,0} + \sum_{n=0}^{\infty} \sum_{r=a}^{B} p_{n,r} = 1. \tag{11}$$

In order to find the joint distribution of the packets in the queue as well as in the transmission channel (i.e., with the server) at post transmission epoch and random epoch, we adopt the following notations.

- $p^+_{n,r}$ := probability that there are $n$ packets in the queue at post transmission epoch of a batch and $r$ packets with the departing batch, $n \geq 0$, $a \leq r \leq B$,

- $p_n^+$ : =probability that the queue contains $n$ packets at post transmission epoch of a batch=$\sum_{r=a}^{B} p_{n,r}^+,\ n \geq 0,$
- $q_k^+$ :=probability that there are $k$ packets in the transmission channel at post transmission epoch of a batch=$\sum_{n=0}^{\infty} p_{n,k}^+,\ a \leq k \leq B.$

Now we obtain $p_{n,0}$, $(0 \leq n \leq a-1)$, $p_{n,r}$, $(a \leq r \leq B, n \geq 0)$ and $p_{n,r}^+$, $(a \leq r \leq B, n \geq 0)$. For this first we prove the following lemmas.

**Lemma 1.** *Probabilities* $p_{n,0}$, $0 \leq n \leq a-1$ *and* $p_{n,r}(1)$, $a \leq r \leq B, n \geq 0$ *are related by the following relation*

$$\begin{aligned}1-\sum_{n=0}^{a-1} p_{n,0} &= s_a \sum_{n=0}^{a-1}\sum_{r=a}^{B} p_{n,r}(1) + \sum_{n=B}^{\infty}\sum_{i=a}^{B} p_{n,i}(1)\sum_{r=a}^{B} s_r y_r \\ &+ \sum_{n=a}^{B-1}\sum_{r=a}^{B} p_{n,r}(1)\left[\overline{\lambda}\{\sum_{j=n}^{B} y_j s_n + \sum_{k=min(a,n-1)}^{n-1} y_k s_k\}\right] \\ &+ \sum_{n=a}^{B-1}\sum_{r=a}^{B} p_{n,r}(1)\left[\lambda\{\sum_{j=n+1}^{B} y_j s_{n+1} + \sum_{k=a}^{n} y_k s_k\}\right]. \qquad (12)\end{aligned}$$

*Proof.* Multiplying the equations (7) to (9) by $z^u$ and adding over $u$ from 1 to $\infty$ we get,

$$\begin{aligned}\left(\frac{z-\overline{\lambda}}{z}\right) p_{0,a}^*(z) &= \lambda p_{a-1,0} s_a^*(z)\sum_{i=a}^{B} y_i + \overline{\lambda} s_a^*(z)\sum_{i=a}^{B} p_{a,i}(1)\sum_{i=a}^{B} y_i \\ &+ \lambda s_a^*(z)\sum_{i=a}^{B} p_{a-1,i}(1)\sum_{i=a}^{B} y_i - \overline{\lambda} p_{0,a}(1), \qquad (13)\\ \left(\frac{z-\overline{\lambda}}{z}\right) p_{0,r}^*(z) &= \overline{\lambda}\sum_{i=a}^{B} p_{r,i}(1)\sum_{j=r}^{B} y_j s_r^*(z) + \lambda\sum_{i=a}^{B} p_{r-1,i}(1)\sum_{j=r}^{B} y_j s_r^*(z) \\ &- \overline{\lambda} p_{0,r}(1), \qquad a+1 \leq r \leq B, \qquad (14)\\ \left(\frac{z-\overline{\lambda}}{z}\right) p_{n,r}^*(z) &= \overline{\lambda}\sum_{i=a}^{B} p_{n+r,i}(1) y_r s_r^*(z) - \lambda p_{n-1,r}(1) + \lambda\sum_{i=a}^{B} p_{n+r-1,i}(1) y_r s_r^*(z) \\ &+ \frac{\lambda}{z} p_{n-1,r}^*(z) - \overline{\lambda} p_{n,r}(1),\ \ n \geq 1,\ a \leq r \leq B. \qquad (15)\end{aligned}$$

Now, using the equations (5) and (6) we obtain the following relation

$$\lambda p_{n,0} = \sum_{k=0}^{n-1}\sum_{i=a}^{B} p_{k,i}(1) + \overline{\lambda}\sum_{i=a}^{B} p_{n,i}(1),\ 0 \leq n \leq a-1. \qquad (16)$$

Now using the relation (16) in (13) and then adding the equations (13) to (15), we get for $|z| < 1$

$$\frac{1}{z}\sum_{n=0}^{\infty}\sum_{r=a}^{B} p^{*}_{n,r}(z) = \frac{s^{*}_{a}(z)-1}{z-1}\sum_{n=0}^{a-1}\sum_{r=a}^{B} p_{n,r}(1) + \sum_{n=a}^{B-1}\sum_{r=a}^{B} p_{n,r}(1)\frac{\left[\overline{\lambda}\{\sum_{j=n}^{B} y_j s^{*}_{n}(z) + \sum_{k=min(a,n-1)}^{n-1} y_k s^{*}_{k}(z)\}\right]}{z-1} + \sum_{n=a}^{B-1}\sum_{r=a}^{B} p_{n,r}(1)\frac{\left[\lambda\{\sum_{j=n+1}^{B} y_j s^{*}_{n+1}(z) + \sum_{k=a}^{n} y_k s^{*}_{k}(z)\}\right]-1}{z-1} + \frac{\sum_{r=a}^{B} s^{*}_{r}(z)y_r - 1}{z-1}\sum_{n=B}^{\infty}\sum_{i=a}^{B} p_{n,i}(1). \tag{17}$$

Now, taking limit $z \to 1$ in the above expression and using L' Hôpital's rule we get the desired result.

In the following lemma we establish the relationship between post transmission and random epoch state probabilities.

**Lemma 2.** *The probabilities* $p^{+}_{n,r}$ *and* $p_{n,r}(1)$, $n \geq 0$, $a \leq r \leq B$ *are connected by the following relations*

$$\begin{aligned} p^{+}_{0,r} &= \kappa\overline{\lambda}p_{0,r}(1), \;\; a \leq r \leq B, \\ p^{+}_{n,r} &= \kappa(\overline{\lambda}p_{n,r}(1) + \lambda p_{n-1,r}(1)), \;\; a \leq r \leq B, n \geq 1, \end{aligned} \tag{18}$$

*where* $\kappa^{-1} = \sum_{n=a}^{\infty}\sum_{r=a}^{B} p_{n,r}(1)$.

*Proof.* It may be noted that $p_{n,r}(1)$ denotes the probability that there are $n$ packets in the queue and $r$ with the server and the remaining service time is just one slot, whereas $p^{+}_{n,r}$ denotes the probability that there are $n$ packets in the queue and $r$ packets with with the departing batch. So it can be perceived that

$$\begin{aligned} p^{+}_{0,r} &\propto \overline{\lambda}p_{0,r}(1), \;\; a \leq r \leq B, \\ p^{+}_{n,r} &\propto (\overline{\lambda}p_{n,r}(1) + \lambda p_{n-1,r}(1)), \;\;\; a \leq r \leq B, n \geq 1. \end{aligned}$$

Let $\kappa$ be the proportionality constant and its value can be easily obtained by the relation $\sum_{n=0}^{\infty}\sum_{r=a}^{B} p^{+}_{n,r} = 1$. Hence the lemma is proved.

The following lemma gives an native expression for $\kappa$, which will be used in a later stage.

**Lemma 3.** *The value of $\kappa$ is given by*

$$\kappa^{-1} \;=\; \psi^{-1}\left(1-\sum_{n=0}^{a-1} p_{n,0}\right), \tag{19}$$

*where*

$$\psi \;=\; s_a \sum_{n=0}^{a} p_n^+ + \sum_{n=a+1}^{B} p_n^+ \left[\sum_{j=n}^{B} y_j s_n + \sum_{j=a}^{n-1} y_j s_j\right] + \sum_{n=B+1}^{\infty} p_n^+ \sum_{r=a}^{B} y_r s_r.$$

*Proof.* The proof is easy and follows from Lemma 1 after some algebraic manipulations.

### *3.2. Construction of b.p.g.f. for queue and server content at post transmission epoch*

The aim of this section is to derive the b.p.g.f. corresponding to the queue and server content at post transmission epoch and extract the probabilities $p^+_{n,r}$, $n \geq 0$, $a \leq r \leq B$ from it through the inversion of the b.p.g.f. For this purpose let us define the following p.g.f.s

$$\widetilde{P}(z,x,\xi) = \sum_{n=0}^{\infty}\sum_{r=a}^{B} p^*_{n,r}(z)x^n\xi^r,\ |x| \leq 1,\ |\xi| \leq 1. \tag{20}$$

$$P^+(x) = \sum_{n=0}^{\infty} p_n^+ x^n,\ |x| \leq 1. \tag{21}$$

$$P^+(x,\xi) = \sum_{n=0}^{\infty}\sum_{r=a}^{B} p^+_{n,r}x^n\xi^r,\ |x| \leq 1,\ |\xi| \leq 1. \tag{22}$$

$$P^+(x,1) = \sum_{n=0}^{\infty}\sum_{r=a}^{B} p^+_{n,r}x^n = \sum_{n=0}^{\infty} p_n^+ x^n = P^+(x),\ |x| \leq 1. \tag{23}$$

Now multiplying (13) to (15) by the appropriate powers of $x$ and $\xi$ and summing over $n$ from 0 to $\infty$ and $r$ from $a$ to $B$, we get

$$\begin{aligned}
\left[\frac{z-(\overline{\lambda}+\lambda x)}{z}\right]\widetilde{P}(z,x,\xi) &= s_a^*(z)\sum_{n=0}^{a-1}\sum_{r=a}^{B}\left[\overline{\lambda}p_{n,r}(1)+\lambda p_{n-1,r}(1)\right]\xi^a \\
&+ \sum_{n=a}^{B}\sum_{r=a}^{B}\left[\overline{\lambda}p_{n,r}(1)+\lambda p_{n-1,r}(1)\right]\sum_{j=n}^{B} y_j\xi^n s_n^*(z) \\
&+ \sum_{n=1}^{\infty}\sum_{r=a}^{B}\sum_{i=a}^{B}\left[\overline{\lambda}p_{n+r,i}(1)+\lambda p_{n+r-1,i}(1)\right]\left[y_r\widehat{\Gamma}^{(r)}(x)\xi^r x^n\right] \\
&- \sum_{n=0}^{\infty}\sum_{r=a}^{B}\left[\overline{\lambda}p_{n,r}(1)+\lambda p_{n-1,r}(1)\right]x^n\xi^r.
\end{aligned} \tag{24}$$

Now substituting $z = \overline{\lambda} + \lambda x$ and using $\widehat{\Gamma}^{(r)}(x) = s_r^*(\overline{\lambda} + \lambda x), \ \ (a \le r \le B)$ in (24), and multiplying by $\kappa$ on both sides, then after some algebraic manipulations we get

$$P^+(x,\xi) = \widehat{\Gamma}^{(a)}(x) \sum_{n=0}^{a-1} p_n^+ \xi^a + \sum_{n=a}^{B} p_n^+ \left[ \sum_{j=n}^{B} y_j \xi^n \widehat{\Gamma}^{(n)}(x) \right] + \sum_{n=1}^{\infty} \sum_{r=a}^{B} p_{n+r}^+ \left[ y_r \widehat{\Gamma}^{(r)}(x) \xi^r x^n \right]. \tag{25}$$

Next we substitute $\xi = 1$ in (25) and after using some simple algebra, we get

$$P^+(x) = \frac{N(x)}{\chi(x)}, \tag{26}$$

where

$$N(x) = \sum_{n=0}^{a} p_n^+ \left[ \widehat{\Gamma}^{(a)}(x) x^B - x^n \sum_{j=a}^{B} y_j \widehat{\Gamma}^{(j)}(x) x^{B-j} \right] + \sum_{n=a+1}^{B-1} p_n^+ \left[ x^B \widehat{\Gamma}^{(n)}(x) \sum_{j=n}^{B} y_j - x^n \sum_{j=n}^{B} y_j \widehat{\Gamma}^{(j)}(x) x^{B-j} \right], \tag{27}$$

$$\chi(x) = x^B - \sum_{j=a}^{B} y_j \widehat{\Gamma}^{(j)}(x) x^{B-j}. \tag{28}$$

Now using the expressions of $P^+(x)$ in (25) and after simplification, we finally get

$$P^+(x,\xi) = \frac{\phi_1(x,\xi)}{\chi(x)}, \tag{29}$$

where

$$\begin{aligned} \phi_1(x,\xi) \ &= \ \sum_{n=0}^{a} p_n^+ \Bigg\{ \chi(x) \left( \widehat{\Gamma}^{(a)}(x) \xi^a - \sum_{r=a}^{B} y_r \xi^r \widehat{\Gamma}^{(r)}(x) x^{n-r} \right) \\ &+ \ \left[ \widehat{\Gamma}^{(a)}(x) x^B - x^n \sum_{r=a}^{B} y_r \widehat{\Gamma}^{(r)}(x) x^{B-r} \right] \sum_{r=a}^{B} y_r \xi^r \widehat{\Gamma}^{(r)}(x) x^{-r} \Bigg\} \\ &+ \ \sum_{n=a+1}^{B-1} p_n^+ \Bigg\{ \chi(x) \left( \sum_{r=n}^{B} y_r \xi^n \widehat{\Gamma}^{(n)}(x) - \sum_{r=n}^{B} y_r \xi^r \widehat{\Gamma}^{(r)}(x) x^{n-r} \right) \\ &+ \ \left[ x^B \widehat{\Gamma}^{(n)}(x) \sum_{r=n}^{B} y_r - x^n \sum_{r=n}^{B} y_r \widehat{\Gamma}^{(r)}(x) x^{B-r} \right] \sum_{r=a}^{B} y_r \xi^r \widehat{\Gamma}^{(r)}(x) x^{-r} \Bigg\}. \end{aligned} \tag{30}$$

Now for the inversion of the b.p.g.f. $P^+(x,\xi)$, we need to evaluate all the $B$ unknowns $p_n^+$, $0 \le n \le B-1$, which are placed in $\phi_1(x,\xi)$. Using the Rouché's theorem (see, Theorem

4 in Appendix), it is proved that $\chi(x)$ admits $B$ zeros viz., $x_0 = 1,\ x_1, \ldots,\ x_{B-1}$ in the closed unit disk $\{x \in C : |x| \leq 1\}$. Since, under stability condition $P^+(x)$ is analytic in this closed unit disk. So, $P^+(x)$ admits no singularities inside this domain. Hence the zeros of $\chi(x)$ must also be the zeros of $N(x)$ inside this domain. Now these $B$ unknowns can be found using the following steps.

Step 1. First obtain the $B-1$ linear equations using the fact that $P^+(x)$ has zeros $x_i$, $1 \leq i \leq B-1$ inside the unit disk. Which are as follows

$$0 = \sum_{n=0}^{a} p_n^+ \left[ \widehat{\Gamma}^{(a)}(x_i) {x_i}^B - {x_i}^n \sum_{j=a}^{B} y_j \widehat{\Gamma}^{(j)}(x_i) {x_i}^{B-j} \right] + \sum_{n=a+1}^{B-1} p_n^+ \left[ {x_i}^B \widehat{\Gamma}^{(n)}(x_i) \sum_{j=n}^{B} y_j - {x_i}^n \sum_{j=n}^{B} y_j \widehat{\Gamma}^{(j)}(x_i) {x_i}^{B-j} \right]. \tag{31}$$

Step 2. Next using the normalizing condition $P^+(1) = 1$ (after using L' Hôpital's rule), we get another equation.

Thus in aggregate we obtain $B$ independent equations, which are sufficient for the evaluation of the unknowns $p_n^+$, $0 \leq n \leq B-1$.

### *3.3. Extraction of the joint state probabilities $\{p_{n,r}^+\}$ from the b.p.g.f. $P^+(x, \xi)$*

Since $P^+(x, \xi)$ is completely known to us after evaluation of the unknown probabilities $p_n^+$, $0 \leq n \leq B-1$, our task reduces to the extraction of the state probabilities $\{p_{n,r}^+\}$. The entire procedure adopted for extraction of the state probabilities are described below in a stepwise manner.

Step 1. First we collect the coefficients of $\xi^a$ from both sides of (29). Which gives,

$$\sum_{n=0}^{\infty} p_{n,a}^+ x^n = \frac{N_a(x)}{\chi(x)}, \tag{32}$$

where

$$\begin{aligned} N_a(x) &= \sum_{n=0}^{a} p_n^+ \Bigg\{ \chi(x) \left( \widehat{\Gamma}^{(a)}(x) - y_a \widehat{\Gamma}^{(a)}(x) x^{n-a} \right) \\ &\quad + \left[ \widehat{\Gamma}^{(a)}(x) x^B - x^n \sum_{r=a}^{B} y_r \widehat{\Gamma}^{(r)}(x) x^{B-r} \right] y_a \widehat{\Gamma}^{(a)}(x) x^{-a} \Bigg\} \\ &+ \sum_{n=a+1}^{B-1} p_n^+ \Bigg\{ \left[ x^B \widehat{\Gamma}^{(n)}(x) \sum_{r=n}^{B} y_r - x^n \sum_{r=n}^{B} y_r \widehat{\Gamma}^{(r)}(x) x^{B-r} \right] y_a \widehat{\Gamma}^{(a)}(x) x^{-a} \Bigg\}. \end{aligned}$$

Step 2. Next, we collect the coefficients of $\xi^j$, $a+1 \le j \le B-1$ from both sides of (29). Which gives,

$$\sum_{n=0}^{\infty} p_{n,j}^{+} x^n = \frac{N_j(x)}{\chi(x)}, \tag{33}$$

where

$$\begin{aligned} N_j(x) &= \sum_{n=0}^{a} p_n^{+} \Bigg\{ \chi(x) \Bigg( - y_j \widehat{\Gamma}^{(j)}(x) x^{n-j} \Bigg) \\ &+ \Bigg[ \widehat{\Gamma}^{(a)}(x) x^B - x^n \sum_{r=a}^{B} y_r \widehat{\Gamma}^{(r)}(x) x^{B-r} \Bigg] y_j \widehat{\Gamma}^{(j)}(x) x^{-j} \Bigg\} \\ &+ p_j^{+} \chi(x) \sum_{r=j+1}^{B} y_r \widehat{\Gamma}^{(j)}(x) - \chi(x) \sum_{n=a+1}^{j-1} p_n^{+} y_j \widehat{\Gamma}^{(j)}(x) x^{n-j} \\ &+ \sum_{n=a+1}^{B-1} p_n^{+} \Bigg\{ \Bigg[ x^B \widehat{\Gamma}^{(n)}(x) \sum_{r=n}^{B} y_r - x^n \sum_{r=n}^{B} y_r \widehat{\Gamma}^{(r)}(x) x^{B-r} \Bigg] y_j \widehat{\Gamma}^{(j)}(x) x^{-j} \Bigg\}. \end{aligned}$$

Step 3. Next, we collect the coefficients of $\xi^B$ from both sides of (29), which results in

$$\sum_{n=0}^{\infty} p_{n,B}^{+} x^n = \frac{N_B(x)}{\chi(x)}, \tag{34}$$

where

$$\begin{aligned} N_B(x) &= \sum_{n=0}^{a} p_n^{+} \Bigg\{ \chi(x) \Bigg( - y_B \widehat{\Gamma}^{(B)}(x) x^{n-B} \Bigg) \\ &+ \Bigg[ \widehat{\Gamma}^{(a)}(x) x^B - x^n \sum_{r=a}^{B} y_r \widehat{\Gamma}^{(r)}(x) x^{B-r} \Bigg] y_B \widehat{\Gamma}^{(B)}(x) x^{-B} \Bigg\} \\ &+ \sum_{n=a+1}^{B-1} p_n^{+} \chi(x) \Bigg( - y_B \widehat{\Gamma}^{(B)}(x) x^{n-B} \Bigg) \\ &+ \sum_{n=a+1}^{B-1} p_n^{+} \Bigg\{ \Bigg[ x^B \widehat{\Gamma}^{(n)}(x) \sum_{r=n}^{B} y_r - x^n \sum_{r=n}^{B} y_r \widehat{\Gamma}^{(r)}(x) x^{B-r} \Bigg] y_B \widehat{\Gamma}^{(B)}(x) x^{-B} \Bigg\}. \end{aligned}$$

Step 4. Next step is to get the state probabilities $\{p_{n,r}^{+}\}$ through the inversion of the generating functions (32)-(34), which are completely known functions for a specific service time distribution. Since in most of the cases (32)-(34) generating functions are rational functions, we can proceed for the partial fraction technique on the generating functions to extract the state probabilities. We explain the entire procedure in the following.

One can see that $\widehat{\Gamma}^{(j)}(x),\ (a \le j \le B)$ is a rational function for most of the distributions (e.g., geometric, negative binomial, etc.) and hence $N_j(x)$ and $\chi(x)$ are also rational functions. So after substituting $N_j(x)$ and $\chi(x)$ in (32)-(34), we get rational functions whose numerators and denominators are different from $N_j(x)$ and $\chi(x)$. For this we denote the numerators of the generating functions (32)-(34) as $\widetilde{N_j}(x)$, $a \le j \le B$ and denominators as $\widetilde{D_j}(x)$, $a \le j \le B$, respectively, to maintain the continuity of our analysis. Clearly, the degrees of $\widetilde{N_j}(x)$ and $\widetilde{D_j}(x)$ will vary on the selection of different service time and service batch-size rule. Let us assume that the degrees of $\widetilde{N_j}(x)$ and $\widetilde{D_j}(x)$ be $F_j$ and $G_j$ respectively. Now we obtain $\{p^+_{n,r}\}$, $n \ge 0$, $a \le r \le B$ in terms of the roots of $\widetilde{D_j}(x) = 0$. Let us assume that $\alpha_{1,j}, \alpha_{2,j}, \ldots, \alpha_{G_j,j}$ be the the distinct roots of $\widetilde{D_j}(x) = 0$, $a \le j \le B$. As the degree of $\widetilde{D_j}(x)$ is $G_j$, and $B$ inside roots are canceled out with the roots of the numerator, so there are $G_j - B$ remaining roots (say, $\alpha_{1,j}, \alpha_{2,j}, \ldots, \alpha_{G_j-B,j}$) that lies in $|x| > 1$. Now the following cases arise:

Case 1. $F_j \ge G_j$

Now applying the partial fraction expansion, the rational function $\sum_{n=0}^{\infty} p^+_{n,j} x^n$, $a \le j \le B$, can be uniquely expressed as,

$$\sum_{n=0}^{\infty} p^+_{n,j} x^n = \sum_{i=0}^{F_j-G_j} \omega_{i,j} x^i + \sum_{k=1}^{G_j-B} \frac{\vartheta_{k,j}}{\alpha_{k,j} - x}. \tag{35}$$

Now the constants $\omega_{i,j}$ and $\vartheta_{k,j}$ are to be determined. $\omega_{i,j}$ can easily be obtained through the division of the polynomials $\widetilde{N_j}(x)$ and $\widetilde{D_j}(x)$ and collecting the corresponding coefficients of the quotient. Next, using residue theorem, we have

$$\vartheta_{k,j} = -\frac{\widetilde{N_j}(\alpha_{k,j})}{\widetilde{D_j}'(\alpha_{k,j})}, k = 1, 2, \ldots, G_j - B,$$

where $\widetilde{D_j}'(\alpha_{k,j})$ is the first order derivative of $\widetilde{D_j}(x)$ evaluated at $x = \alpha_{k,j}$. Now, collecting the coefficients of $x^n$ from both sides of (35), we have for $a \le j \le B$

$$p^+_{n,j} = \begin{cases} \omega_{n,j} + \sum_{k=1}^{G_j-B} \frac{\vartheta_{k,j}}{\alpha_{k,j}^{n+1}}, & 0 \le n \le F_j - G_j, \\ \sum_{k=1}^{G_j-B} \frac{\vartheta_{k,j}}{\alpha_{k,j}^{n+1}}, & n > F_j - G_j. \end{cases} \tag{36}$$

Case 2. $F_j < G_j$

In this case using partial fraction expansion, the rational function $\sum_{n=0}^{\infty} p_{n,j}^{+} x^n$, $a \leq j \leq B$, can be uniquely expressed as,

$$\sum_{n=0}^{\infty} p_{n,j}^{+} x^n = \sum_{k=1}^{G_j - B} \frac{\vartheta_{k,j}}{\alpha_{k,j} - x}. \tag{37}$$

Where $\vartheta_{k,j}$ are determined by $\vartheta_{k,j} = -\frac{\widetilde{N_j}(\alpha_{k,j})}{\widetilde{D_j}'(\alpha_{k,j})}, k = 1, 2, \ldots, G_j - B.$

Now, collecting the coefficients of $x^n$ from both sides of (37), we have

$$p_{n,j}^{+} = \sum_{k=1}^{G_j - B} \frac{\vartheta_{k,j}}{\alpha_{k,j}^{n+1}}, n \geq 0, a \leq j \leq B. \tag{38}$$

**Remark 1.** *We have discussed here only the case when $\widetilde{D_j}(x) = 0$ admits distinct roots. If some of the roots are repeated then this partial fraction technique can easily be modified to get the desired result.*

After determination of post transmission epoch state probabilities, we obtain the random epoch state probabilities by establishing a relationship between these two.

### *3.4. Relationship between random epoch and post transmission epoch state probabilities*

It is always desirable for the vendors to know the number of packets in the transmission channel as well as in the queue waiting for transmission at a any random epoch. We now establish a relationship among the arbitrary and departure epoch state probabilities.

**Theorem 2.** *The state probabilities $\{p_{n,0}, p_{n,r}\}$ and $\{p_{n,r}^{+}, p_n^{+}\}$ are connected by the following relations*

$$p_{n,0} = \frac{1}{\widetilde{\widetilde{E}}} \sum_{j=0}^{n} p_j^{+}, \qquad 0 \leq n \leq a-1, \tag{39}$$

$$p_{0,a} = \frac{1}{\widetilde{\widetilde{E}}} \left( \sum_{n=0}^{a} p_n^{+} - p_{0,a}^{+} \right), \tag{40}$$

$$p_{0,r} = \frac{1}{\widetilde{\widetilde{E}}} \left( p_r^{+} \sum_{j=r}^{B} y_j - p_{0,r}^{+} \right), \qquad a+1 \leq r \leq B, \tag{41}$$

$$p_{n,r} = p_{n-1,r} + \frac{1}{\widetilde{\widetilde{E}}} \left( p_{n+r}^{+} y_r - p_{n,r}^{+} \right), \qquad n \geq 1,\ a \leq r \leq B, \tag{42}$$

*where*

$$\widetilde{E} = \lambda\psi + \sum_{j=0}^{a-1} (a-j)p_j^{+}. \tag{43}$$

*Proof.* The proof is quite straight forward and can be easily done assuming $z \to 1$ in equations (13)-(15), and simple algebraic manipulations with the equations (5)-(9) and (18)-(19).

**Remark 2.** *The present analysis can be utilized to obtain the results for some important queueing models, listed below.*

1. *The $Geo/G/1$ model*
   *This is the most basic model in queueing theory. The server transmits packets one by one and it cannot be idle as long as the queue is non-empty. So, this model can be obtained if we set $a = B = 1$, $y_a = y_B = 1$ in our case.*

2. *The $Geo/G^{[k]}/1$ model*
   *In this queue, the server transmits only a fixed number (say, $k$) of packets. The server remains idle until there are '$k$' packets in the queue. If number of packets in the queue is larger than '$k$' then he takes only '$k$' packets for transmission. (this type of service rule is known as 'fixed batch size rule). This model can be obtained if we set $a = B = k$, $y_k = 1$ in our present set up.*

3. *The $Geo/G_n^{(l)}/1$ model*
   *Here the server is allowed to take up to $l$ number of packets. The server cannot be idle if at least one packets is waiting for transmission. If the queue is non-empty then he takes $min(l,\ whole\ queue)$ packets into his service. This model can be obtained if we take $a = 1, B = l$ and $y_l = 1$ in our model.*

4. *The $Geo/G_n^{[Y]}/1$ model*
   *Here the server transmits packets with random-batch-size $Y$ with the p.m.f. $P(Y = i) = y_i, i = 1, 2, \ldots, B$ with finite mean $E(Y)$, where $B$ is the maximum transmission capacity of the server. If at the beginning of the transmission, the server has a serving capacity $i$, then the server takes $min(i,\ the\ whole\ queue)$ packets for service with the probability $y_i$ (i.e., with random batch-size rule with a maximum threshold $B$). To best of authors' knowledge, no such results are available in the literature for this model. This can be obtained from our model, if we take $a = 1$.*

5. *The $Geo/G_n^{[a,b]}/1$ model*
   *This model is known as 'General Batch Service rule' model can be obtained from our model if we take $B = b$ and $y_b = 1$. In this rule, the server remains idle until there are '$a$' packets in the queue. If number of packets in the queue is larger than '$a$' then he takes $min(b,\ the\ whole\ queue)$ packets for transmission.*

# 4. Marginal Distributions and Performance Measures

In the following we present some important marginal distributions, as well as, performance measures which may be useful to the vendors to determine the service policy beforehand.

### 4.1. Marginal distributions

1. Distribution of the number of packets in the system ($P^n_{system},\ n \geq 0$), where

$$P^n_{system} = \begin{cases} p_{n,0}, & n \leq a-1, \\ \sum_{r=a}^{n} p_{n-r,r}, & a \leq n \leq B, \\ \sum_{r=a}^{B} p_{n-r,r}, & n \geq B+1. \end{cases}$$

2. Distribution of the number of packets in the channel awaiting transmission ($P^n_{queue},\ n \geq 0$), where

$$P^n_{queue} = \begin{cases} p_{n,0} + \sum_{r=a}^{B} p_{n,r}, & n \leq a-1, \\ \sum_{r=a}^{B} p_{n,r}, & n \geq a. \end{cases}$$

3. Distribution of the number of packets with the server ($P^n_{server},\ n \geq 0$), where

$$p^{server}_r = \begin{cases} \displaystyle\sum_{n=0}^{a-1} p_{n,0}, & \text{when } r = 0; \\ \displaystyle\sum_{n=0}^{\infty} p_{n,r}, & \text{when } a \leq r \leq B. \end{cases}$$

### 4.2. Performance characteristics

1. Probability that the server is idle ($p_{idle}$) $= \sum_{n=0}^{a-1} p_{n,0}$.

2. Probability that the server is busy ($p_{busy}$) $= 1 - \sum_{n=0}^{a-1} p_{n,0}$.

3. Mean number of packets awaiting transmission $(L_{queue}) = \sum_{n=0}^{\infty} np^{queue}_n$.

4. Mean waiting time $(W) = \frac{L_{queue}}{\lambda}$.

5. Proportion of the packets that have immediate access to transmission channel at post transmission epochs $P_{ims} = \sum_{n=a}^{B} p^+_n$.

6. Average number of packets with the server $(L_{server})$ =$\sum_{r=a}^{B} rp_r^{server}$.

7. Average number of packets in the system $(L_{system})$=$\sum_{n=0}^{\infty} np_n^{system}$.

## 5. System Outlay Model

It is very often important to the vendors to minimize the system cost as much as possible in a pre-operating phase. The objective of the present section is to demonstrate the effects of the system parameters viz., $a,\ B$ and service batch-size rule $Y$ on the system cost, which may help the vendors to design the service system accordingly. For the present model, we associate some level dependent cost to the system characteristics viz.,
$(i)$ $C_w$:= waiting cost per packet per unit time in the system during transmission phase.
$(ii)$ $Idleness\ cost\ (C_{idle})$ := cost per packet per unit time either due to empty queue or due to less packets (dependent on the lower threshold of the serving capacity)
$(ii)$ $C_{loss}$:= cost per unit time due to incapability of the server to take the packets beyond the maximum serving capacity.
Hence the total system cost is given by,

$$\begin{aligned} S(a,B,Y) &= C_w \sum_{n=0}^{\infty} n \left[ p_{a-1,0} + \frac{1}{\widetilde{\widetilde{E}}} \left( \sum_{i=a}^{B} \sum_{j=i}^{B} p_i^+ y_j - p_0^+ \right) + \frac{1}{\widetilde{\widetilde{E}}} \sum_{j=1}^{n} \left( \sum_{i=a}^{B} p_{i+j}^+ y_i - p_j^+ \right) \right] \\ &+ C_{idle} \sum_{n=0}^{a-1} n \left( \frac{1}{\widetilde{\widetilde{E}}} \sum_{i=0}^{n} p_i^+ \right) + C_{loss} \left[ \sum_{n=B+1}^{\infty} np_n^+ - B \right]. \end{aligned} \tag{44}$$

Our objective is to minimize the system cost by controlling the parameters $a,\ B$ for a specified service batch-size rule i.e., $Y$. Hence, our optimization problem is as follows,

$$Minimize \quad S(a,B,Y) \tag{45}$$
$$a, B \in N, a \leq B \tag{46}$$

**Remark 3.** *It is evident from the above complicated expression of the cost function that, it is very hard to obtain the optimized value of $S(a,B,Y)$ as a function of $a$ and $B$ through conventional optimization procedures e.g., 'Direct Search Method' etc. So, in this case some numerical procedures can be opted for optimization purpose. We have carried out an extensive number of numerical examples to observe the effects of the system parameters and as well as service batch-size rule on the system cost. Due to lack of space we present a few of them in the following sections.*

## 6. Numerical Results

The purpose of this section is to validate our analysis as well as to provide an insight to the readers from the possible application point of view. Following we present some examples where service time distribution is taken as geometric, deterministic, and the discrete PH-type.

### *6.1. The* $Geo/Geo_n{}^{(a,Y)}/1$ *system*

Example 1: Here we consider service time distribution as geometric ($Geo$) and demonstrate some of the numerical results.

Input parameters

- $Threshold\ value$: We have taken $a = 2$ and $B = 7$.
- $Arrival\ Process$: Packets are arriving singly according to Bernoulli arrival process with the parameter $\lambda = 0.7$.
- $Service\ batch - size\ rule$: We have chosen arbitrary service batch-size rule with $y_2 = \frac{1}{3}, y_3 = \frac{1}{6}, y_4 = \frac{1}{6}, y_5 = \frac{1}{9}, y_6 = \frac{1}{9}, y_7 = \frac{1}{9}$.
- $Service\ Time\ Distribution$: Transmission time follows geometric distribution and depends on the service batch-size rule with parameter $\mu_r = \frac{1}{r+1},\ (a \le r \le B)$.

Output

In Table 1 and 2, we present the distribution of the stationary probabilities at post transmission and random epochs, respectively. Some important performance measures related to the model are also given in the last row of Table 2.

Table 1. Stationary distribution of the state probabilities at post transmission-epoch for geometric service time distribution.

| $n$ | $p^+_{n,2}$ | $p^+_{n,3}$ | $p^+_{n,4}$ | $p^+_{n,5}$ | $p^+_{n,6}$ | $p^+_{n,7}$ | $\sum_{r=a}^{B} p^+_{n,r}$ | $p^+_n$ |
|---|---|---|---|---|---|---|---|---|
| 0 | 0.03723881 | 0.00526434 | 0.00253747 | 0.00115897 | 0.00055722 | 0.00020949 | 0.04696631 | 0.04696631 |
| 1 | 0.11201309 | 0.01688644 | 0.00847671 | 0.00392763 | 0.00198794 | 0.00084507 | 0.14413690 | 0.14413690 |
| 2 | 0.07595248 | 0.01469963 | 0.00841909 | 0.00408071 | 0.00236707 | 0.00128832 | 0.10680731 | 0.10680731 |
| 3 | 0.05272838 | 0.01262161 | 0.00802584 | 0.00405375 | 0.00257730 | 0.00159040 | 0.08159727 | 0.08159727 |
| 4 | 0.03763969 | 0.01078390 | 0.00747670 | 0.00392187 | 0.00267216 | 0.00178829 | 0.06428260 | 0.06428260 |
| 5 | 0.02772708 | 0.00922107 | 0.00687504 | 0.00373269 | 0.00268866 | 0.00190906 | 0.05215360 | 0.05215360 |
| 6 | 0.02112339 | 0.00792086 | 0.00627788 | 0.00351608 | 0.00265254 | 0.00197271 | 0.04346346 | 0.04346346 |
| 7 | 0.01664732 | 0.00685149 | 0.00571447 | 0.00329048 | 0.00258178 | 0.00199410 | 0.03707964 | 0.03707964 |
| 8 | 0.01354912 | 0.00597583 | 0.00519785 | 0.00306694 | 0.00248892 | 0.00198441 | 0.03226307 | 0.03226307 |
| 9 | 0.01135110 | 0.00525824 | 0.00473179 | 0.00285182 | 0.00238277 | 0.00195206 | 0.02852779 | 0.02852779 |
| 10 | 0.00974756 | 0.00466746 | 0.00431511 | 0.00264851 | 0.00226946 | 0.00190344 | 0.02555154 | 0.02555154 |
| . | . | . | . | . | . | . | . | . |
| 20 | 0.00392998 | 0.00195586 | 0.00194739 | 0.00128669 | 0.00124678 | 0.00118491 | 0.01155163 | 0.01155163 |
| . | . | . | . | . | . | . | . | . |
| 50 | 0.00056461 | 0.00027904 | 0.00027661 | 0.00018340 | 0.00018303 | 0.00018319 | 0.00166989 | 0.00166989 |
| . | . | . | . | . | . | . | . | . |
| 100 | 0.00002337 | 0.00001155 | 0.00001145 | 0.00000759 | 0.00000758 | 0.00000759 | 0.00006914 | 0.00006914 |
| . | . | . | . | . | . | . | . | . |
| 200 | 0.00000004 | 0.00000002 | 0.00000002 | 0.00000001 | 0.00000001 | 0.00000001 | 0.00000012 | 0.00000012 |
| . | . | . | . | . | . | . | . | . |
| $\geq$250 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 |
| Total | 0.53194034 | 0.15781355 | 0.12484290 | 0.07339076 | 0.06083550 | 0.05117695 | 1.00000000 | 1.00000000 |

Table 2. Stationary distribution of the state probabilities at random-epoch for geometric service time distribution.

| | Idle | Busy | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $n$ | $p_{n,0}$ | $p_{n,2}$ | $p_{n,3}$ | $p_{n,4}$ | $p_{n,5}$ | $p_{n,6}$ | $p_{n,7}$ | $p_n^{queue}$ |
| 0 | 0.01502009 | 0.08336430 | 0.01571328 | 0.00946747 | 0.00518903 | 0.00291065 | 0.00125059 | 0.13291543 |
| 1 | 0.06111590 | 0.05624030 | 0.01373923 | 0.00953641 | 0.00547738 | 0.00359249 | 0.00212677 | 0.15182846 |
| 2 | | 0.03880291 | 0.01181804 | 0.00916058 | 0.00548993 | 0.00398192 | 0.00272846 | 0.07198183 |
| 3 | | 0.02749974 | 0.01009822 | 0.00857025 | 0.00533995 | 0.00417139 | 0.00312779 | 0.05880734 |
| 4 | | 0.02009564 | 0.00862585 | 0.00789881 | 0.00509942 | 0.00422476 | 0.00337738 | 0.04932187 |
| 5 | | 0.01518112 | 0.00739655 | 0.00722069 | 0.00481363 | 0.00418641 | 0.00351600 | 0.04231441 |
| 6 | | 0.01186505 | 0.00638397 | 0.00657491 | 0.00451066 | 0.00408726 | 0.00357248 | 0.03699434 |
| 7 | | 0.00958225 | 0.00555475 | 0.00597963 | 0.00420750 | 0.00394895 | 0.00356841 | 0.03284150 |
| 8 | | 0.00797301 | 0.00487589 | 0.00544105 | 0.00391404 | 0.00378664 | 0.00352006 | 0.02951068 |
| 9 | | 0.00680734 | 0.00431800 | 0.00495884 | 0.00363566 | 0.00361089 | 0.00343976 | 0.02677050 |
| 10 | | 0.00593747 | 0.00385636 | 0.00452932 | 0.00337493 | 0.00342909 | 0.00333689 | 0.02446406 |
| . | | . | . | . | . | . | . | . |
| 20 | | 0.00251307 | 0.00166611 | 0.00207389 | 0.00164632 | 0.00186775 | 0.00203776 | 0.01180490 |
| . | | . | . | . | . | . | . | . |
| 50 | | 0.00036249 | 0.00023887 | 0.00029598 | 0.00023549 | 0.00027419 | 0.00031366 | 0.00172068 |
| . | | . | . | . | . | . | . | . |
| 100 | | 0.00001501 | 0.00000989 | 0.00001225 | 0.00000975 | 0.00001135 | 0.00001300 | 0.00007124 |
| . | | . | . | . | . | . | . | . |
| 200 | | 0.00000003 | 0.00000002 | 0.00000002 | 0.00000002 | 0.00000002 | 0.00000002 | 0.00000012 |
| . | | . | . | . | . | . | . | . |
| ≥251 | | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 |
| Total | 0.07613599 ($p_0^{server}$) | 0.35724689 ($p_2^{server}$) | 0.14131510 ($p_3^{server}$) | 0.13973916 ($p_4^{server}$) | 0.09857729 ($p_5^{server}$) | 0.09533215 ($p_6^{server}$) | 0.09165342 ($p_7^{server}$) | 1.00000000 |
| Performance Measures | $L_{queue} = 10.93024048$, $L_{server} = 3.40384898$, $W = 15.61462925$, $P_{ims} = 0.38538388$ | | | | | | | |

### 6.2. The $Geo/D_n^{(a,Y)}/1$ system

Example 2: Here we consider deterministic ($D$) service time distribution and demonstrate some of the numerical results.

<u>Input parameters</u>

- $Threshold\ value$: We have taken $a = 3$ and $B = 7$.
- $Arrival\ Process$: Packets are arriving singly according to Bernoulli arrival process with the parameter $\lambda = 0.6$.
- $Service\ batch-size\ rule$: We have chosen service batch-size rule as
  $$y_i := \begin{cases} \frac{2i}{(B+a)(B-a+1)}, & i = a \dots B; \\ 0, & \text{elsewhere.} \end{cases}$$
- $Service\ Time\ Distribution$: Transmission time follows deterministic distribution and depends on the service batch-size rule with parameter $s_r = r + 2,\ (a \leq r \leq B)$.

<u>Output</u>

In Table 3 and 4, we present the distribution of the stationary probabilities at post transmission and random epochs, respectively, along with some important performance measures related to the model.

Table 3. Stationary distribution of the state probabilities at post transmission-epoch for deterministic service time distribution.

| $n$ | $p^+_{n,3}$ | $p^+_{n,4}$ | $p^+_{n,5}$ | $p^+_{n,6}$ | $p^+_{n,7}$ | $\sum_{r=a}^{B} p^+_{n,r}$ | $p^+_n$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.00497948 | 0.00096463 | 0.00018246 | 0.00002039 | 0.00000155 | 0.00614851 | 0.00614851 |
| 1 | 0.03767495 | 0.00878301 | 0.00193540 | 0.00024801 | 0.00002149 | 0.04866285 | 0.04866285 |
| 2 | 0.11469473 | 0.03350765 | 0.00883386 | 0.00132560 | 0.00013296 | 0.15849479 | 0.15849479 |
| 3 | 0.17695553 | 0.06890014 | 0.02255414 | 0.00407696 | 0.00048441 | 0.27297117 | 0.27297117 |
| 4 | 0.14199551 | 0.08154627 | 0.03499744 | 0.00792859 | 0.00115104 | 0.26761884 | 0.26761884 |
| 5 | 0.05440932 | 0.05480588 | 0.03350569 | 0.01008265 | 0.00186662 | 0.15467015 | 0.15467015 |
| 6 | 0.01044190 | 0.01968880 | 0.01921243 | 0.00838494 | 0.00210281 | 0.05983087 | 0.05983087 |
| 7 | 0.00439751 | 0.00435937 | 0.00629444 | 0.00447012 | 0.00164989 | 0.02117135 | 0.02117135 |
| 8 | 0.00157756 | 0.00154686 | 0.00142080 | 0.00153337 | 0.00090318 | 0.00698178 | 0.00698178 |
| 9 | 0.00053403 | 0.00052131 | 0.00047710 | 0.00041919 | 0.00035807 | 0.00230971 | 0.00230971 |
| 10 | 0.00017642 | 0.00017224 | 0.00015765 | 0.00013852 | 0.00011833 | 0.00076316 | 0.00076316 |
| . | . | . | . | . | . | . | . |
| 15 | 0.00000070 | 0.00000068 | 0.00000062 | 0.00000055 | 0.00000047 | 0.00000301 | 0.00000301 |
| . | . | . | . | . | . | . | . |
| 20 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000001 | 0.00000001 |
| . | . | . | . | . | . | . | . |
| $\geq 25$ | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 |
| Total | 0.54792404 | 0.27488120 | 0.12964925 | 0.03869673 | 0.00884879 | 1.00000000 | 1.00000000 |

Table 4. Stationary distribution of the state probabilities at random-epoch for deterministic service time distribution.

| | Idle | Busy | | | | | |
|---|---|---|---|---|---|---|---|
| $n$ | $p_{n,0}$ | $p_{n,3}$ | $p_{n,4}$ | $p_{n,5}$ | $p_{n,6}$ | $p_{n,7}$ | $p_n^{queue}$ |
| 0 | 0.00166822 | 0.13058643 | 0.06363572 | 0.03016553 | 0.00843584 | 0.00160797 | 0.23609971 |
| 1 | 0.01487150 | 0.12907770 | 0.06796716 | 0.03288710 | 0.00974716 | 0.00213254 | 0.25668316 |
| 2 | 0.05787453 | 0.10299440 | 0.06147316 | 0.03163913 | 0.00984213 | 0.00227193 | 0.26609529 |
| 3 | | 0.05693057 | 0.04369815 | 0.02589857 | 0.00888637 | 0.00219848 | 0.13761214 |
| 4 | | 0.01909345 | 0.02187599 | 0.01652835 | 0.00678487 | 0.00190534 | 0.06618800 |
| 5 | | 0.00455835 | 0.00710624 | 0.00747895 | 0.00406565 | 0.00140523 | 0.02461443 |
| 6 | | 0.00180044 | 0.00179738 | 0.00227990 | 0.00179607 | 0.00083678 | 0.00851057 |
| 7 | | 0.00063215 | 0.00062554 | 0.00057660 | 0.00058502 | 0.00038982 | 0.00280914 |
| 8 | | 0.00021234 | 0.00020946 | 0.00019260 | 0.00016958 | 0.00014500 | 0.00092899 |
| 9 | | 0.00007016 | 0.00006922 | 0.00006365 | 0.00005604 | 0.00004792 | 0.00030699 |
| 10 | | 0.00002319 | 0.00002288 | 0.00002104 | 0.00001852 | 0.00001584 | 0.00010148 |
| . | | . | . | . | . | . | . |
| 15 | | 0.00000009 | 0.00000009 | 0.00000008 | 0.00000007 | 0.00000006 | 0.00000040 |
| . | | . | . | . | . | . | . |
| $\geq 20$ | | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 |
| Total | 0.07441425 ($p_0^{server}$) | 0.44599065 ($p_3^{server}$) | 0.26849220 ($p_4^{server}$) | 0.14774181 ($p_5^{server}$) | 0.05039641 ($p_6^{server}$) | 0.01296467 ($p_7^{server}$) | 1.00000000 |
| Performance Measures | $L_{queue} = 1.6720472$, $L_{server} = 3.543781$, $W = 2.7867452698$, $P_{ims} = 0.7762624$ | | | | | | |

### *6.3. The* $Geo/DPH_n{}^{(a,Y)}/1$ *system*

Example 3: Here we consider discrete PH-type ($DPH$) service time distribution and demonstrate some of the numerical results.

Input parameters

- $Threshold\ value$: We have taken $a = 2$ and $B = 6$.
- $Arrival\ Process$: Packets are arriving singly according to Bernoulli arrival process with the parameter $\lambda = 0.07$.
- $Service\ batch - size\ rule$: We have taken arbitrary service batch-size rule with $y_2 = 0.2, y_3 = 0.3, y_4 = 0.15, y_5 = 0.25, y_6 = 0.10$.
- $Service\ Time\ Distribution$: Transmission time follows Discrete Phase type distribution ($DPH(\boldsymbol{\tau}, \mathbf{T})$). Recall that $DPH(\boldsymbol{\tau}, \mathbf{T})$ considers the absorption time into the state 0 in the discrete-time Markov Chain with initial probability vector $(\tau_0, \boldsymbol{\tau})$ and the transition probability matrix

$$P := \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{t} & \mathbf{T} \end{pmatrix},$$

where $T_{ij} \geq 0$ and $t_i \geq 0$ for $1 \leq i, j \leq n$ and that $\mathbf{t} + \mathbf{T1} = \mathbf{1}$ where $\mathbf{1}$ is a column vector of ones with suitable dimension. It is to be noted that, $DPH(\boldsymbol{\tau}, \mathbf{T})$ distribution includes discrete analogues of exponential, hyper exponential, Erlang, geometric, mixture of geometric, and negative binomial distributions etc. So, it is evident from the modeling point of view that discrete PH-type distribution is quite appealing.

For our example we took matrices $\mathbf{T}_r$ $(a \leq r \leq B)$ of order 3 as:

$$\mathbf{T}_r = \begin{pmatrix} 1-\theta_r & \theta_r & 0 \\ 0 & 1-\theta_r & \theta_r \\ 0 & 0 & 1-\theta_r \end{pmatrix}$$

where $\theta_r = \frac{0.2}{r+1}$ and $\boldsymbol{\tau} = (\ 0.2\ 0.5\ 0.3\ )$. The mean service time $(s_r)$ of a batch size $r$ $(a \leq r \leq B)$ is obtained by $s_r = \boldsymbol{\tau}(\mathbf{I}_3 - \mathbf{T}_r)^{-1}\mathbf{1}$, and is given as $s_2 = 28.5, s_3 = 38, s_4 = 47.5, s_5 = 57, s_6 = 66.5$.

Output

In Table 5 and 6, we present the distribution of the stationary probabilities at post transmission and random epochs, respectively. Some important performance measures related to the model are presented in the last row of Table 6.

**Remark 4.** *Although we have carried out extensive numerical works, due to space constraints we have presented few of them.*

Table 5. Stationary distribution of the state probabilities at post transmission-epoch for DPH service time distribution.

| $n$ | $p^+_{n,2}$ | $p^+_{n,3}$ | $p^+_{n,4}$ | $p^+_{n,5}$ | $p^+_{n,6}$ | $\sum_{r=a}^{B} p^+_{n,r}$ | $p^+_n$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.09220916 | 0.01462051 | 0.00614407 | 0.00299845 | 0.00061876 | 0.11659096 | 0.11659096 |
| 1 | 0.08660721 | 0.01865469 | 0.00776647 | 0.00493917 | 0.00118860 | 0.11915614 | 0.11915614 |
| 2 | 0.06759672 | 0.01958908 | 0.00824514 | 0.00626272 | 0.00163685 | 0.10333051 | 0.10333051 |
| 3 | 0.04987541 | 0.01905077 | 0.00817818 | 0.00709047 | 0.00196335 | 0.08615819 | 0.08615819 |
| 4 | 0.03594316 | 0.01773784 | 0.00778953 | 0.00751456 | 0.00217972 | 0.07116481 | 0.07116481 |
| 5 | 0.02583453 | 0.01609763 | 0.00723273 | 0.00762937 | 0.00230261 | 0.05909687 | 0.05909687 |
| 6 | 0.01880259 | 0.01439513 | 0.00660746 | 0.00751936 | 0.00234985 | 0.04967439 | 0.04967439 |
| 7 | 0.01400824 | 0.01277294 | 0.00597485 | 0.00725448 | 0.00233833 | 0.04234884 | 0.04234884 |
| 8 | 0.01075344 | 0.01129697 | 0.00536968 | 0.00688983 | 0.00228295 | 0.03659287 | 0.03659287 |
| 9 | 0.00852434 | 0.00998826 | 0.00480959 | 0.00646703 | 0.00219628 | 0.03198550 | 0.03198550 |
| 10 | 0.00696679 | 0.00884369 | 0.00430154 | 0.00601653 | 0.00208856 | 0.02821711 | 0.02821711 |
| 20 | 0.00201114 | 0.00296910 | 0.00147110 | 0.00239910 | 0.00093273 | 0.00978317 | 0.00978317 |
| 40 | 0.00027354 | 0.00040317 | 0.00019897 | 0.00032892 | 0.00013116 | 0.00133576 | 0.00133576 |
| 60 | 0.00003746 | 0.00005521 | 0.00002724 | 0.00004504 | 0.00001797 | 0.00018291 | 0.00018291 |
| 80 | 0.00000513 | 0.00000756 | 0.00000373 | 0.00000617 | 0.00000246 | 0.00002505 | 0.00002505 |
| 100 | 0.00000070 | 0.00000104 | 0.00000051 | 0.00000084 | 0.00000034 | 0.00000343 | 0.00000343 |
| . | . | . | . | . | . | . | . |
| 150 | 0.00000000 | 0.00000001 | 0.00000000 | 0.00000001 | 0.00000000 | 0.00000002 | 0.00000002 |
| . | . | . | . | . | . | . | . |
| 170 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 |
| Total | 0.47126209 | 0.24135581 | 0.11112231 | 0.13180953 | 0.04445025 | 1.00000000 | 1.00000000 |

Table 6. Stationary distribution of the state probabilities at random-epoch for DPH service time distribution.

| | Idle | Busy | | | | | |
|---|---|---|---|---|---|---|---|
| $n$ | $p_{n,0}$ | $p_{n,2}$ | $p_{n,3}$ | $p_{n,4}$ | $p_{n,5}$ | $p_{n,6}$ | $p_n^{queue}$ |
| 0 | 0.03839232 | 0.08129149 | 0.01788248 | 0.00969377 | 0.00582366 | 0.00143198 | 0.15451570 |
| 1 | 0.07762934 | 0.05844676 | 0.01876983 | 0.01005535 | 0.00828656 | 0.00243509 | 0.17562294 |
| 2 | | 0.04087457 | 0.01815734 | 0.00979390 | 0.00971058 | 0.00310106 | 0.08163745 |
| 3 | | 0.02834307 | 0.01679130 | 0.00919266 | 0.01038817 | 0.00350780 | 0.06822300 |
| 4 | | 0.01977878 | 0.01513391 | 0.00843510 | 0.01054683 | 0.00371921 | 0.05761383 |
| 5 | | 0.01406073 | 0.01344802 | 0.00763331 | 0.01035746 | 0.00378648 | 0.04928599 |
| 6 | | 0.01027915 | 0.01186760 | 0.00685127 | 0.00994515 | 0.00374995 | 0.04269312 |
| 7 | | 0.00777287 | 0.01044908 | 0.00612206 | 0.00939946 | 0.00364093 | 0.03738440 |
| 8 | | 0.00609019 | 0.00920558 | 0.00545976 | 0.00878313 | 0.00348339 | 0.03302206 |
| 9 | | 0.00493420 | 0.00812831 | 0.00486746 | 0.00813914 | 0.00329544 | 0.02936456 |
| 10 | | 0.00411461 | 0.00719909 | 0.00434233 | 0.00749611 | 0.00309054 | 0.02624268 |
| 20 | | 0.00126170 | 0.00244911 | 0.00149393 | 0.00289451 | 0.00130311 | 0.00940235 |
| 40 | | 0.00017187 | 0.00033324 | 0.00020269 | 0.00039622 | 0.00018153 | 0.00128556 |
| 60 | | 0.00002354 | 0.00004563 | 0.00002776 | 0.00005425 | 0.00002486 | 0.00017604 |
| 80 | | 0.00000322 | 0.00000625 | 0.00000380 | 0.00000743 | 0.00000340 | 0.00002411 |
| 100 | | 0.00000044 | 0.00000086 | 0.00000052 | 0.00000102 | 0.00000047 | 0.00000330 |
| . | | | | | | | |
| 150 | | 0.00000000 | 0.00000001 | 0.00000000 | 0.00000001 | 0.00000000 | 0.00000002 |
| . | | | | | | | |
| 170 | | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.00000000 |
| Total | 0.11602166 ($p_0^{server}$) | 0.30958857 ($p_2^{server}$) | 0.21140678 ($p_3^{server}$) | 0.12166690 ($p_4^{server}$) | 0.17318059 ($p_5^{server}$) | 0.06813550 ($p_6^{server}$) | 1.00000000 |
| Performance Measures | $L_{queue} = 7.3430458367$, $L_{server} = 3.0147810191$, $W = 104.8507725325$, $P_{ims} = 0.3694247678$ | | | | | | |

# 7. Optimality Analysis

Our objective in this section is to provide a decision rule for the system designer or vendor, about how to minimize the system cost by controlling the system parameters $a$ and $B$ or the service batch-size rule $Y$.

A1. Here first we fix the range of the server (i.e., highest capacity of the server ($B$)-lowest capacity of the server ($a$)) and investigate the effect of the lower threshold value i.e.,'$a$' on the total system cost $S(a, B, Y)$ for different service batch-size rule ($Y$) (see below).

a. Increasing batch-size distribution (IBSD):
Here we take $y_i := \begin{cases} \frac{2i}{(B+a)(B-a+1)}, & i = a..B; \\ 0, & \text{elsewhere.} \end{cases}$

b. Decreasing batch-size distribution (DBSD):
Here we take $y_i := \begin{cases} \frac{2(B+a-i)}{(B+a)(B-a+1)}, & i = a..B; \\ 0, & \text{elsewhere.} \end{cases}$

c. Uniform batch-size distribution (UBSD):
Here we take $y_i := \begin{cases} \frac{1}{(B-a+1)}, & i = a..B; \\ 0, & \text{elsewhere.} \end{cases}$

d. Truncated Geometric batch-size distribution (TGBSD):
Here $y_i := \begin{cases} \frac{\tau(1-\tau)^{i-a}}{1-(1-\tau)^{B-a+1}}, & i = a..B; \\ 0, & \text{elsewhere.} \end{cases}$, where $\tau$ $(0 \leq 1)$ is the parameter of the TGBSD.

In Fig 4, we observe the effect of '$a$' on $S(a, B, Y)$. For this, we consider the deterministic service-time distribution with $s_r = r + 2,$ $(a \leq r \leq B)$ and fix the range of the server's capacity $B = 7$ and $\lambda = 0.4$. Also for this example we took $C_w = 15$ per unit message, $C_{idle} = 20$ per unit time, $C_{loss} = .01$ per unit time to form the cost function $S(a, B, Y)$ for our system.

It is observed from the Fig 4 that, the cost function grows almost linearly with the increasing value of the parameter '$a$'. Also we see the significant impact of the service batch-size rule $Y$ for lower values of '$a$'. On the basis of the above figure one may be able to suggest that as '$a$' reaches to '$B$', the effect of service batch-size rule does not play a major role. Here, for $a = 1$ IBSD gives minimum system cost. It may be hereby admitted that, however here we have considered a moderate value for the different costs and may vary from system to system, the lower threshold value may play significant role for different service batch-size rule $Y$.

A2. Now, we fix the lower threshold value '$a$' and change the upper threshold value '$B$' and look for the optimal '$B$' for different service batch-size rule $Y$. After executing a number of examples it is found that, for a fixed lower threshold value '$a$', the larger

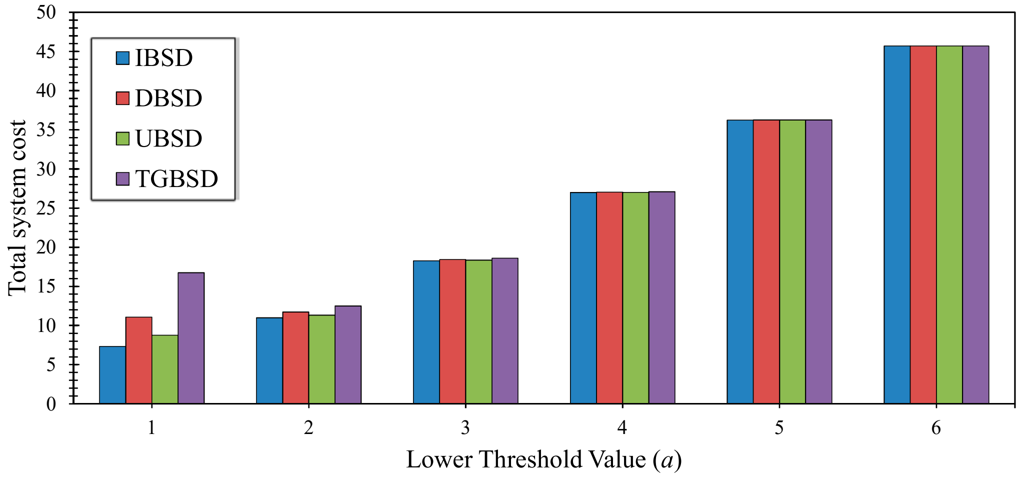


Figure 4. Effect of '$a$' on $S(a, B, Y)$.

'$B$' gives more efficient system performance, irrespective of different batch-size rule $Y$. Based on our computational experience, it may be suggested that one should try to increase the range of the server's capacity as much as possible to get a better performance of the system.

A3. The purpose of the example is to compare the batch size independent service (NBSD) with batch size dependent service (BSD) on the basis of its cost effectiveness. Since the cost function consists of some key performance measures, so overall performance of these two service assumptions will be reflected therein. Now, for this purpose, we consider the $Geo/Geo_n^{(a,Y)}/1$ queue, with $\lambda = 0.45$, minimum threshold value $a = 2$ and service batch-size rule as UBSD, with

$$y_i := \begin{cases} \frac{1}{(B-a+1)}, & i = a..B; \\ 0, & \text{elsewhere.} \end{cases}$$

Cost components are assumed to be the same as in example A1. Now we consider the following two cases,

**Case 1**. For BSD we assume that $\mu_r = \frac{1.5}{r+1}$, $(2 \leq r \leq B)$, see column-2 of Table 7.

**Case 2**. For NBSD, the same service rates for all the batches are the weighted average of the service rates of case 1. The weighted average is calculated by the formula $\frac{\sum_{r=a}^{B} r\mu_r}{\sum_{r=a}^{B} r}$ and is given in column-3 of Table 7.

The total system cost for various values of $B$ is given in columns 4 and 5 of Table 7. For a better representation, we have displayed these cost in Fig 5 with varying $B$ from 4 to

Table 7. Service rates and total system cost for Case-1 and Case-2 for different $B$.

| | Service rates | | Total system cost | |
|---|---|---|---|---|
| $B$ | For Case-1 | For Case-2 $(a \leq r \leq B)$ | For Case-1 | For Case-2 |
| 3 | $\mu_2 = 0.5000, \mu_3 = 0.3750$ | $\mu_r = 0.4250$ | 11.0541 | 11.9826 |
| 4 | $\mu_2 = 0.5000, \mu_3 = 0.3750, \mu_4 = 0.3000$ | $\mu_r = 0.3694$ | 10.8230 | 13.8701 |
| 5 | $\mu_2 = 0.5000, \mu_3 = 0.3750, \mu_4 = 0.3000, \mu_5 = 0.2500$ | $\mu_r = 0.3268$ | 10.6400 | 15.2635 |
| 6 | $\mu_2 = 0.5000, \mu_3 = 0.3750, \mu_4 = 0.3000, \mu_5 = 0.2500, \mu_6 = 0.2143$ | $\mu_r = 0.2930$ | 10.5004 | 17.0004 |
| 7 | $\mu_2 = 0.5000, \mu_3 = 0.3750, \mu_4 = 0.3000, \mu_5 = 0.2500, \mu_6 = 0.2143, \mu_7 = 0.1875$ | $\mu_r = 0.2657$ | 10.3687 | 19.6653 |
| 8 | $\mu_2 = 0.5000, \mu_3 = 0.3750, \mu_4 = 0.3000, \mu_5 = 0.2500, \mu_6 = 0.2143, \mu_7 = 0.1875, \mu_8 = 0.1667$ | $\mu_r = 0.2430$ | 10.2462 | 21.9449 |
| 9 | $\mu_2 = 0.5000, \mu_3 = 0.3750, \mu_4 = 0.3000, \mu_5 = 0.2500, \mu_6 = 0.2143, \mu_7 = 0.1875, \mu_8 = 0.1667, \mu_9 = 0.1500$ | $\mu_r = 0.2240$ | 10.1295 | 24.6258 |
| 10 | $\mu_2 = 0.5000, \mu_3 = 0.3750, \mu_4 = 0.3000, \mu_5 = 0.2500, \mu_6 = 0.2143, \mu_7 = 0.1875, \mu_8 = 0.1667, \mu_9 = 0.1500, \mu_{10} = 0.1364$ | $\mu_r = 0.2078$ | 10.0166 | 26.8252 |

10. One can observe from Fig 5 that for a fixed value of $B$, total system cost is always lower in case 1 as compared to the case 2. It may also be noted here that increase of $B$ produces steady decrease of total system cost for case 1 whereas steady increase of total system cost for case 2. These outcomes suggest that BSD produces better performance of the system as compared to NBSD.

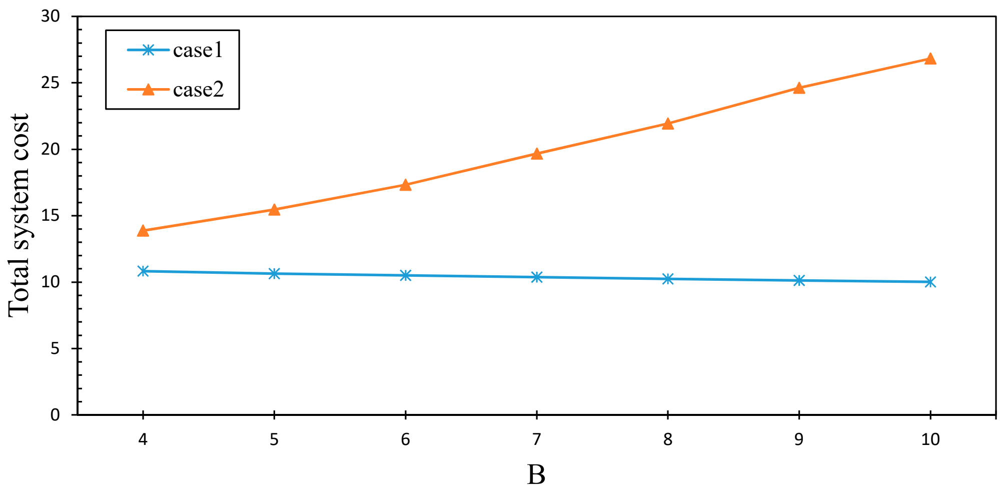


Figure 5. Effect of '$B$' on $S(a, B, Y)$ for dependent (Case-1) and independent (Case-2) batch size distribution.

## 8. Conclusion

So far, in this paper we have considered an infinite buffer information transmission channel with Bernoulli input, arbitrarily distributed service time with load dependent versatile service policy and obtained stationary probabilities for both the queue content and server content (busy states of the server) at both random and post transmission epochs. The state probabilities are presented in a tractable and simplified manner, which greatly helps to understand the busy states in details and it is hoped that this would be quite beneficial to the vendors. Further we have also obtained important performance characteristics related to the present model and developed a cost model to provide the essence of the usefulness of the present system to the vendors. Also we have shown the importance of inclusion of load sensitive service rates in real system to enhance its efficiency and cost effectiveness. However, it may be concluded here that, the present analysis can be further applied to analyze more complex systems such as queues with correlated arrival processes that is considering the arrival process as discrete-time Markovian arrival process (D-MAP).

## Acknowledgements

The authors would like to thank the referees for their valuable comments and suggestions which led to improvements in the paper.

## References

[1] Abolnikov, L. & Dukhovny, A. (1991). Markov chains with transition delta-matrix: ergodicity conditions, invariant probability measures and applications. *International Journal of Stochastic Analysis*, 4(4), 333–355.

[2] Alfa, A. (2010). *Queueing Theory for Telecommunications: Discrete Time Modeling of a Single Node System*, Springer, New York.

[3] Artalejo, J. R., Atencia, I., & Moreno, P. (2005). A discrete-time $Geo^{[X]}/G/1$ retrial queue with control of admission. *Applied Mathematical Modelling*, 29(11), 1100–1120.

[4] Banerjee, A., Gupta, U., & Goswami, V. (2014). Analysis of finite-buffer discrete-time batch-service queue with batch-size-dependent service. *Computers & Industrial Engineering*, 75, 121–128.

[5] Bruneel, H., & Kim, B. G. (1993). Discrete-time models for communication systems including ATM. Kluwer Academic Publishers, Boston.

[6] Bruneel, H., Mélange, W., Steyaert, B., Claeys, D., & Walraevens, J. (2012). A two-class discrete-time queueing model with two dedicated servers and global FCFS service discipline. *European Journal of Operational Research*, 223(1), 123–132.

[7] Chang, S. H., & Choi, D. W. (2005). Performance analysis of a finite-buffer discrete-time queue with bulk arrival, bulk service and vacations. *Computers & Operations Research*, 32(9), 2213–2234.

[8] Chaudhry, M., & Chang, S. H. (2004). Analysis of the discrete-time bulk-service queue $Geo/G^Y/1/N+B$. *Operations Research Letters*, 32(4), 355–363.

[9] Claeys, D., Laevens, K., Walraevens, J., & Bruneel, H. (2010). Complete characterisation of the customer delay in a queueing system with batch arrivals and batch service. *Mathematical Methods of Operations Research*, 72(1), 1–23.

[10] Claeys, D., Walraevens, J., Laevens, K., & Bruneel, H. (2010). Delay analysis of two batch-service queueing models with batch arrivals: $Geo^X/Geo^c/1$. *4OR*, 8(3), 255–269.

[11] Claeys, D., Walraevens, J., Laevens, K., & Bruneel, H. (2011). Analysis of threshold-based batch-service queueing systems with batch arrivals and general service times. *Performance Evaluation*, 68(6), 528–549.

[12] De Clercq, S., Laevens, K., Steyaert, B., & Bruneel, H. (2013). A multi-class discrete-time queueing system under the FCFS service discipline. *Annals of Operations Research*, 202(1), 59–73.

[13] Feyaerts, B., De Vuyst, S., Bruneel, H., & Wittevrongel, S. (2012). Analysis of discrete-time buffers with heterogeneous session-based arrivals and general session lengths. *Computers & Operations Research*, 39(12), 2905–2914.

[14] Fiems, D., & Bruneel, H. (2013). Discrete-time queueing systems with Markovian preemptive vacations. *Mathematical and Computer Modelling*, 57(3), 782–792.

[15] Goswami, V., & Mund, G. (2011). Analysis of discrete-time batch service renewal input queue with multiple working vacations. *Computers & Industrial Engineering*, 61(3), 629–636.

[16] Goswami, V., & Samanta, S. (2009). Discrete-time bulk-service queue with two heterogeneous servers. *Computers & Industrial Engineering*, 56(4), 1348–1356.

[17] Gupta, U., & Goswami, V. (2002). Performance analysis of finite buffer discrete-time queue with bulk service. *Computers & Operations Research*, 29(10), 1331–1341.

[18] Hwang, E., Kim, K. J., Lyakhov, A., & Choi, B. D. (2012). Performance analysis of bandwidth requests under unicast, multicast and broadcast pollings in IEEE 802.16d/e. *Telecommunication Systems*, 50(1), 15–30.

[19] Li, J.-h. (2013). Analysis of the discrete-time $Geo/G/1$ working vacation queue and its application to network scheduling. *Computers & Industrial Engineering*, 65(4), 594–604.

[20] Samanta, S., Chaudhry, M., & Gupta, U. (2007). Discrete-time queues with single and multiple vacations. *Mathematical and Computer Modelling*, 45(1), 93–108.

[21] Samanta, S., Gupta, U., & Chaudhry, M. (2009). Analysis of stationary discrete-time $GI/D-MSP/1$ queue with finite and infinite buffers. *4OR*, 7(4), 337–361.

[22] Takagi, H. (1993). Queueing analysis: A Foundation of Performance Evaluation: Discrete Time Systems, Vol. 3, North-Holland, Amsterdam: Elsevier Science.

[23] Xeung, W., Kim, N., Yoon, B., & Chae, K. (2007). Analysis of the queue-length distribution for the discrete-time batch-service $Geo/G^{a,Y}/1/K$ queue. *European Journal of Operational Research*, 181(2), 787–792.

# Appendix

**Theorem 3.** *Show that $K'(1) < B \Leftrightarrow \rho < 1$.*

*Proof.* We have $K(x) = \sum_{j=a}^{B} y_j s_j^*(\overline{\lambda} + \lambda x)x^{B-j}$. Now differentiating $K(x)$ wrt $x$, we get

$$K'(x) = \sum_{j=a}^{B} y_j \left(\lambda s_j^{*(1)}(\overline{\lambda} + \lambda x)x^{B-j} + s_j^*(\overline{\lambda} + \lambda x)(B-j)x^{B-j-1}\right). \quad (47)$$

Now setting $x = 1$ in (47), we get

$$\begin{aligned} K'(1) &= \sum_{j=a}^{B} y_j \left(\lambda s_j^{*(1)}(1) + s_j^*(1)(B-j)\right), \\ K'(1) &= \sum_{j=a}^{B} y_j \left(\lambda s_j + B - j\right), \quad \text{using } s_j = s_j^{*(1)}(1) \text{ and } s_j^*(1) = 1, \\ K'(1) &= \lambda \sum_{j=a}^{B} y_j s_j + B - \overline{y}. \end{aligned} \quad (48)$$

Now from Abolnikov and Dukhovny [1], we have

$$\begin{aligned} K'(1) < B \quad &\Leftrightarrow \quad \lambda \sum_{j=a}^{B} y_j s_j + B - \overline{y} < B, \ \text{using (48)} \\ &\Leftrightarrow \quad \lambda \sum_{j=a}^{B} y_j s_j < \overline{y}, \\ &\Leftrightarrow \quad \lambda \frac{\sum_{j=a}^{B} y_j s_j}{\overline{y}} < 1. \\ &\Leftrightarrow \quad \rho < 1. \end{aligned}$$

**Theorem 4.** *For a small $\epsilon > 0$, $\chi(z)$ admits exactly $B$ number of zeroes inside the open disc $|z| < 1 + \epsilon$.*

*Proof.* $|z| = 1+\epsilon$ is a simple and closed contour. Let $f(z) = z^B$ and $g(z) = -\sum_{i=a}^{B} y_j \widehat{\Gamma}^{(j)}(z) z^{B-j}$. Clearly $f(z)$ and $g(z)$ are both holomorphic inside and on $|z| = 1 + \epsilon$. Now we can write,

$$
\begin{aligned}
|g(z)| &= |\sum_{i=a}^{B} y_i z^{B-i} \widehat{\Gamma}^{(i)}(z)| \\
&= \sum_{i=a}^{B} y_i |z|^{B-i} \widehat{\Gamma}^{(i)}(|z|) \\
&= \sum_{i=a}^{B} y_i (1+\epsilon)^{B-i} \widehat{\Gamma}^{(i)}(1+\epsilon) \\
&= \sum_{i=a}^{B} y_i \{1 + (B-i)\epsilon + o(\epsilon)\}\{1 + \Gamma^{(i)(1)}(1) + o(\epsilon)\} \\
&= \sum_{i=a}^{B} y_i \{1 + (B-i)\epsilon + \epsilon \frac{\lambda}{\mu_i} + o(\epsilon)\} \\
&= 1 + \epsilon B + \epsilon\lambda \sum_{i=a}^{B} \frac{y_i}{\mu_i} - \epsilon \overline{y} + o(\epsilon). \\
&= 1 + \epsilon B + \epsilon \overline{y}(\rho - 1) + o(\epsilon) \\
&\leq 1 + \epsilon B \qquad \text{(since under stability condition } \rho < 1) \\
&\leq (1+\epsilon)^B \\
&= |z|^B = |f(z)|.
\end{aligned}
$$

Thus by Rouché's theorem, the functions $f(z)$ and $f(z) + g(z)$ admits same number of zeros inside the circle $|z| < 1 + \epsilon$, where $\epsilon$ is arbitrarily small. Since $f(z) = |z|^B$ has $B$ number of zeros inside the the unit disk $|z| = 1$ i.e. in $|z| \leq 1$ (assuming $\epsilon \to 0$), hence$f(z) + g(z) = \chi(z)$ has $B$ number of zeros in $|z| \leq 1$.