Contemporary Mathematics
http://ojs.wiserpub.com/index.php/CM/

Research Article

Analysis of Batch Arrival General Service Queue with Balking, Feedback and Second Optional Service

P Vijaya Laxmi[1*], Hasan A Qrewi[1], Andwilile A George[2]

[1]Department of Applied Mathematics, Andhra University, Visakhapatnam, Andhra Pradesh, India
[2]College of Business Education, Tanzania
Email: vijayalaxmiau@gmail.com

Received: 15 March 2023; **Revised:** 16 June 2023; **Accepted:** 1 August 2023

Abstract: This paper aims to analyze the steady-state behavior of bulk input general service queue with a second optional service (SOS), balking, and feedback facility. In this study, the server provides two kinds of services such as first essential service (FES) and SOS. The FES is provided to all arriving customers to the system while SOS is only to those customers who demand additional service. When the customer completes FES and is not satisfied with the service, he may choose to rejoin the queue (feedback) or opt for SOS or depart from the system with a certain probability. We have computed the probability-generating function of the queue-length distribution after converting the non-Markovian to Markovian process by using a supplementary variable technique. This technique is used to solve the non-Markov queue model by taking the elapsed service times as the supplementary variable so that the process becomes Markovian. This study contributes to filling the gap in the analysis of batch arrival general service queues with balking, feedback, and SOS. Furthermore, we have presented the numerical results and cost optimization. The results reveal that the higher service rate in both FES and SOS helps the system manager to run the system effectively. Similarly, in cost optimization, the system manager should make emphasize choosing optimal service rates to have a cost-benefit and less congestion in the queueing system.

***Keywords*:** batch arrival, Non-Markovian, feedback, balking, second optional service

MSC: 60K25, 90B22

1. Introduction

Customer impatience must be considered when analyzing queueing systems to reflect real-world scenarios. Often customers become frustrated by long queues at the service centers. Because of this many authors have studied customer behavior in the queueing system whereby some customers upon arrival decide to join the queue or refuse to join the queue because of various reasons like long waiting line or slow rate of service, etc. This customer impatience situation is referred to as balking [1] was the first who studied customer impatience in queueing theory [2] has studied a single server queue model with impatience where the customers lose patience if the wait is more than the pre-fixed threshold value [3] has discussed the $M/G/1$ queue with a retrial and customer feedback with Bernoulli vacation and SOS. The authors have assumed that if the service is not starting immediately upon arrival, the customer may balk and join the

DOI: https://doi.org/10.37256/cm.4420232688

orbit and re-attempts for service. Many related studies on balking are foundin [4-6], and there ferences therein.

Several researchers have studied queueing systems with feedback, such as [7] wherein they have investigated an $M/M^b/1$ with SOS and feedback. The customers are served in batches without exceeding the maximum capacity b. After a batch completes FES, they will rejoin the line and retake the service if the batch is unsatisfied with the service; otherwise, opt for SOS or leave the system with specific probabilities [8] studied an $M/G/1$ queue system with feedback and vacation. They consider the service times as independent and identically distributed with different rates when the customer is served with feedback or without feedback. Other studies on feedback are found in [9-13], etc.

In queueing theory, items may arrive one by one or in batch. When more than one arrival enters simultaneously in the queueing system, the input is referred to as batch arrival [14] have studied a batch arrival queue with general service and server breakdown and repair. The author reported that the server provides two different phases one after the other, defined as the first phase and second phase services [15] have analyzed the steady-state of an $M^X/G/1$ queue with a retrial and two stages of heterogeneous services with admission, general retrial time, and feedback. Because of the server state, the arrivals join with the dependent admission policy. The supplementary variable approach has been used to derive the stationary equations. The probability-generating functions of the system size and state of the server/system are obtained. Other exciting works on this topic are found in [16-21], etc.

In the existing literature, for example [22] and [24], the customer has only two options such as opting for the feedback if he/she is not happy with the quality of service or leaving the system after completing the service. Further, the arriving customers decide to enter the queue with a certain probability or balk with a certain probability when the server is on vacation period or during a busy period. In [23], the author analyzed the retrial queue with customer balking and feedback subjected to server breakdowns. They reported that the balking occurs when there is no available server upon a customer's arrival. Also, the feedback occurs once the customer service is completed, he may join the orbit again to get additional service after being unsatisfied with the first service. However, it is noted that none of the aforementioned literature reviews has tried to obtain a non-Markovian queue system with a combination of SOS, balking, and feedback with batch arrival and general service. In this study, when customers completion of FES, they may opt to join the SOS or depart from the system or rejoin the system (feedback) if not satisfied with FES. Therefore, adding a second optional service to the queue system, with balking and feedback will make the model more adaptable. This motivates us to explore a non-Markovian batch arrival queue with balking, feedback, and SOS under a steady-stateenvironment.

This paper is structured as follows: Model description and governing equations are presented in Section 2. In Section 3, we derive the steady-state solution. Performance measures are obtained in Section 4 followed by particular cases in Section 5. The cost analysis is done by constructing a suitable cost function in Section 6 and Section 7, numerical illustrations are presented. Finally, Section 8 concludes ourpaper.

2. Description and mathematical formulation of the model

Let us study an $M^X/G/1$ queue with SOS, balking, and feedback. A brief description of the model is presented in the following lines:

• Customers (units) arrive in batches of random size, say X, in a compound Poisson process with probability $P(X = j) = k_j$, so that $\lambda k_j dt$ is the probability of first order that $j(j = 1, 2, ...)$ customers (units) arrives at the system during a short interval of time $(t, t + dt]$. Further, $\sum_{j=1}^{\infty} k_j = 1$, $0 \le k_j \le 1$ for all j, where $\lambda > 0$ is the mean arrival rate of batches.

• The service times for FES and SOS are assumed to follow general arbitrary distributions with distribution functions $F(w)$ and $H(w)$ and the density functions are $f(w)$ and $h(w)$, respectively. Let $\mu(w)dw$, $\beta(w)dw$ be the conditional probabilities of the completion of FES and SOS, respectively during the interval $(w, w + dw]$ with elapsed service time w, so that

$$\mu(w) = \frac{f(w)}{1 - F(w)} \text{ and } f(w) = \mu(w)e^{-\int_0^w \mu(t)dt},$$

$$\beta(w) = \frac{h(w)}{1 - H(w)} \text{ and } h(w) = \beta(w)e^{-\int_0^w \beta(t)dt}.$$

• When a customer arrives, he/she joins the line with probability b or refuses to join the line (balking) with probability $1 - b$.

• After completion of FES, a customer may join the SOS with probability r_0 or depart from the system with probability r_1 or rejoin the system (feedback) if not satisfied with FES with probability r_2 where $r_0 + r_1 + r_2 = 1$.

• All random processes in this model are assumed to be mutually independent.

2.1 *Practical justification of the model*

This model focuses on a practical justification in the field of hospitals. We are aware that COVID-19 is a worldwide epidemic that is affecting the world. Some countries have implemented lockdowns, tracing infected people, and testing to stop the virus fromspreading.

All suspected COVID-19 individuals arrive in batches at the hospital for a primary test. The sample was taken from all incoming suspected and sentfor test. If the test is negative for the coronavirus, the suspected is discharged and excluded from further testing, in contrast, if it tests positive for coronavirus during the primary test, and the patient who is not satisfied *w* ith *t* he result repeats the test or undergoes quarantine or leave the hospital.

Moreover, some patients refuse to join the waiting line if the expected waiting line is too long. Assumption: Testing is done only in densely crowded regions and for people with minimal risk of contracting the infection. This is because there is a significant probability of frequent positive testing. Hospital, primary test, undergo quarantine, repeat the test, and refusing to join the waiting line correspond to the system, FES, SOS, feedback, and balking respectively, in queueing terminology.

2.2 *Formulation of mathematical model*

The state of the system at time t is defined by the Markov process as

$$\{(L_q(t), M(t), \varepsilon_i(t)); i = 1, 2, t \geq 0\},$$

where $L_q(t)$ is the queue length at time t, $M(t)$ be the state of the server at time t which is given by

$$M(t) = \begin{cases} 0, & \text{the server is idle and the queue is empty at time } t, \\ 1, & \text{the server is operating FES at time } t, \\ 2, & \text{the server is operating SOS at time } t. \end{cases}$$

and $\varepsilon_i(t)$ is the elapsed service time of a batch in service ($i = 1$ for FES and $i = 2$ for SOS) at time t.

The state space of the Markov process is given as follows:

$$\Omega = \{\{0, 0\} U\{n, i, \varepsilon_1\} U\{n, i, \varepsilon_2\}; n \geq 0, i = 1, 2\}.$$

The probabilities involved in this model are defined as

$$Q(t) = P\{L_q(t) = 0, M(t) = 0\}, \text{ for } t \geq 0$$

$$P_{n,i}(w, t)dw = Pr\{L_q(t) = n, M(t) = i; w \leq \varepsilon_i(t) \leq w + dw\},$$

for $\varepsilon_i(t)$, $t \geq 0$, $n \geq 0$, $i = 1, 2$.

• $Q(t)$ probability that at time t, the system is empty and the server is idle.

• $P_{n,i}(w, t)$ probability that at time t, there are n (≥ 0) units in the queue, with one unit in the service, elapses service time is w and the server is providing FES for $i = 1$ and SOS for $i = 2$.

According to the description that is given in the previous section, the differential-difference equations are formulated as follows:

$$\frac{d}{dt}Q(t)+\lambda Q(t)=r_1\int_0^\infty P_{0,1}(w,t)\mu(w)dw+\int_0^\infty P_{0,2}(w,t)\beta(w)dw, \tag{1}$$

$$\frac{\partial}{\partial w}P_{0,1}(w,t)+\frac{\partial}{\partial t}P_{0,1}(w,t)=-(\lambda b+\mu(w))P_{0,1}(w,t), \tag{2}$$

$$\frac{\partial}{\partial w}P_{n,1}(w,t)+\frac{\partial}{\partial t}P_{n,1}(w,t)=-(\lambda b+\mu(w))P_{n,1}(w,t)+\lambda b\sum_{i=1}^{n}k_iP_{n-i,1}(w,t)\ \ n\geq 1, \tag{3}$$

$$\frac{\partial}{\partial w}P_{0,2}(w,t)+\frac{\partial}{\partial t}P_{0,2}(w,t)=-(\lambda b+\beta(w))P_{0,2}(w,t), \tag{4}$$

$$\frac{\partial}{\partial w}P_{n,2}(w,t)+\frac{\partial}{\partial t}P_{n,2}(w,t)=-(\lambda b+\beta(w))P_{n,2}(w,t)+\lambda b\sum_{i=1}^{n}k_iP_{n-i,2}(w,t),\ n\geq 1. \tag{5}$$

It is required to solve equations (1)-(5) at $x = 0$ with the following boundary conditions:

$$P_{n,1}(0,t)=\lambda k_{n+1}Q(t)+r_1\int_0^\infty P_{n+1,1}(w,t)\mu(w)dw+r_2\int_0^\infty P_{n,1}(w,t)\mu(w)dw+\int_0^\infty P_{n+1,2}(w,t)\beta(w)dw,\ n\geq 0, \tag{6}$$

$$P_{n,2}(0,t)=r_0\int_0^\infty P_{n,1}(w,t)\mu(w)dw,\ n\geq 0. \tag{7}$$

3. Steady-state solution of the model

At steady state, i.e. as $t\to\infty$, the above probabilities are denoted by Q, $P_{n,i}(w)$, and their derivatives concerning time t vanish. Considering the model in steady-state, the state equations are given as follows:

$$\lambda Q=r_1\int_0^\infty P_{0,1}(w)\mu(w)dw+\int_0^\infty P_{0,2}(w)\beta(w)dw, \tag{8}$$

$$\frac{\partial}{\partial w}P_{0,1}(w)+(\lambda b+\mu(w))P_{0,1}(w)=0, \tag{9}$$

$$\frac{\partial}{\partial w}P_{n,1}(w)+(\lambda b+\mu(w))P_{n,1}(w)=\lambda b\sum_{i=1}^{n}k_iP_{n-i,1}(w),\ n\geq 1, \tag{10}$$

$$\frac{\partial}{\partial w}P_{0,2}(w)+(\lambda b+\beta(w))P_{0,2}(w)=0, \tag{11}$$

$$\frac{\partial}{\partial w}P_{n,2}(w)+(\lambda b+\beta(w))P_{n,2}(w)=\lambda b\sum_{i=1}^{n}k_iP_{n-i,2}(w),\ n\geq 1. \tag{12}$$

The boundary conditions are given by

$$P_{n,1}(0)=\lambda k_{n+1}Q+r_1\int_0^\infty P_{n+1,1}(w)\mu(w)dw+r_2\int_0^\infty P_{n,1}(w)\mu(w)dw+\int_0^\infty P_{n+1,2}(w)\beta(w)dw,\ n\geq 0, \tag{13}$$

$$P_{n,2}(0)=r_0\int_0^\infty P_{n,1}(w)\mu(w)dw,\ n\geq 0. \tag{14}$$

3.1 *Generating functions of the queue length*

The main purpose of this subsection is to solve the equations (8)-(14) using probability generating functions (PGFs). The PGFs are defined as follows:

$$P_i(w,z)=\sum_{n=0}^{\infty}P_{n,i}(w)z^n,\ |z|\le 1,\ w>0,\ i=1,2. \tag{15}$$

$$P_i(0,z)=\sum_{n=0}^{\infty}P_{n,i}(0)z^n,\ |z|\le 1,\ i=1,2. \tag{16}$$

$$K(z)=\sum_{j=1}^{\infty}k_j z^j,\ |z|\le 1. \tag{17}$$

Lemma 1 For $w>0$, we have

$$\text{(I)}\ \frac{\partial}{\partial w}P_1(w,z)+\big(\lambda b(1-K(z))+\mu(w)\big)P_1(w,z)=0, \tag{18}$$

$$\text{(II)}\ \frac{\partial}{\partial w}P_2(w,z)+\big(\lambda b(1-K(z))+\beta(w)\big)P_2(w,z)=0. \tag{19}$$

Proof. (I) Multiplying equations (9) and (10) by appropriate power z^n, summing them from $n=0$ to ∞, and using the definition of PGFs, we get the result. (II) Similarly, from equations (11) and (12), we get the desired result. □

Lemma 2 For $w>0$, we have

$$\text{(I)}\ P_1(w,z)=P_1(0,z)e^{-\eta(z)w-\int_0^w\mu(t)dt}, \tag{20}$$

$$\text{(II)}\ P_2(w,z)=P_2(0,z)e^{-\eta(z)w-\int_0^w\beta(t)dt}, \tag{21}$$

where $\eta(z)=\lambda b(1-K(z))$.

Proof. Integrating equations (18) and (19) in the interval $[0,w]$, we get the desired result. □

Lemma 3 For $w>0$, we have

$$\text{(I)}\ \int_0^{\infty}P_1(w,z)\mu(w)dw=P_1(0,z)F^*(\eta(z)). \tag{22}$$

$$\text{(II)}\ \int_0^{\infty}P_2(w,z)\beta(w)dw=P_2(0,z)H^*(\eta(z)). \tag{23}$$

where $F^*[\eta(z)]$, $H^*[\eta(z)]$ are the Laplace-Steiltjes transform (LST) of the service times $F(w)$ and $H(w)$, respectively.

$$F^*[\eta(z)]=\int_0^{\infty}e^{-\eta(z)w}dF(w),$$

$$H^*[\eta(z)]=\int_0^{\infty}e^{-\eta(z)w}dH(w).$$

Proof. Multiplying equations (20) and (21) by $\mu(w)$ and $\beta(w)$, respectively and integrating with respect to w, we get the result. □

Lemma 4

$$\text{(I) } zP_1(0,z)=\lambda(K(z)-1)Q+r_1F^*(\eta(z))P_1(0,z)+r_2zF^*(\eta(z))P_1(0,z)+P_2(0,z)H^*(\eta(z)). \tag{24}$$

$$\text{(II) } P_2(0,z)=r_0F^*(\eta(z))P_1(0,z). \tag{25}$$

Proof. (I) Multiplying equations (13) by appropriate powers of z^n, summing them from n = 0 to n = ∞, and using the PGFs definition, we get

$$zP_1(0,z)=\lambda K(z)Q+r_1\int_0^\infty P_1(w,z)\mu(w)dw+zr_2\int_0^\infty P_1(w,z)\mu(w)dw$$

$$-\left[r_1\int_0^\infty P_{0,1}(w)\mu(w)dw+\int_0^\infty P_{0,2}(w)\mu_2(w)dw\right]+\int_0^\infty P_2(w,z)\beta(w)dw. \tag{26}$$

Using equation (8) into equation (26), we get

$$zP_1(0,z)=\lambda K(z)Q+r_1\int_0^\infty P_1(w,z)\mu(w)dw+r_2z\int_0^\infty P_1(w,z)\mu(w)dw+\int_0^\infty P_2(w,z)\beta(w)dw-\lambda Q. \tag{27}$$

Substituting equations (22) and (23) in equation (27), we get the result.
(II) Similarly, from equation (14), we get

$$P_2(0,z)=r_0\int_0^\infty P_1(w,z)\mu(w)dw. \tag{28}$$

Substituting equation (22) in equation (28), we obtain the result. □
Lemma 5 Based on the previous results, we have

$$\text{(I) } P_1(0,z)=\frac{\lambda(K(z)-1)Q}{z-r_1F^*(\eta(z))-r_2zF^*(\eta(z))-r_0F^*(\eta(z))H^*(\eta(z))}. \tag{29}$$

$$\text{(II) } P_2(0,z)=\frac{r_0\lambda(K(z)-1)F^*(\eta(z))Q}{z-r_1F^*(\eta(z))-r_2zF^*(\eta(z))-r_0F^*(\eta(z))H^*(\eta(z))}. \tag{30}$$

Proof. (I) Substituting equation (25) in equation (24), we get

$$zP_1(0,z)=\lambda(K(z)-1)Q+r_1F^*(\eta(z))P_1(0,z)+r_2zF^*(\eta(z))P_1(0,z)+r_0F^*(\eta(z))H^*(\eta(z))P_1(0,z).$$

After algebraic calculations, the equation (29) is obtained.
(II) Substituting equation (29) in equation (25), we get the desired result. □
Lemma 6 The PGFs $P_i(z)$, i = 1, 2 are given by

$$\text{(I) } P_1(z)=\frac{-\left[1-F^*(\eta(z))\right]Q}{b\left[z-r_1F^*(\eta(z))-r_2zF^*(\eta(z))-r_0F^*(\eta(z))H^*(\eta(z))\right]}, \tag{31}$$

$$\text{(II) } P_2(z)=\frac{-r_0F(\eta(z))\left[1-H^*(\eta(z))\right]Q}{b\left[z-r_1F^*(\eta(z))-r_2zF^*(\eta(z))-r_0F^*(\eta(z))H^*(\eta(z))\right]}, \tag{32}$$

where $P_i(z) = \int_0^{\infty} P_i(w, z)dw$, $i = 1, 2$.

Proof. Integrating equations (20) and (21) by parts, we get

$$P_1(z) = P_1(0, z)\left(\frac{1 - F^*(\eta(z))}{\eta(z)}\right) \tag{33}$$

$$P_2(z) = P_2(0, z)\left(\frac{1 - H^*(\eta(z))}{\eta(z)}\right). \tag{34}$$

After substituting equations (29) and (30) in equations (33) and (34) respectively, and some algebraic calculations, we get the result. □

Lemma 7 Based on the previous results, we have

$$Q = \frac{b\left[1 - r_2 - \lambda bE(X)E(s) - r_0\lambda bE(X)E(v)\right]}{(1-b)(\lambda bE(X))\left[E(s) + r_0E(v)\right] + b(1 - r_2)}, \tag{35}$$

where $r_2 + \lambda bE(X)[E(s) + r_0E(v)] < 1$.

Proof. To get Q, we have to use the normalizing condition

$$P_1(1) + P_2(1) + Q = 1. \tag{36}$$

Now, clearly $z = 1$ brings equations (31) and (32) to indeterminate $\left(\frac{0}{0}\right)$ form. Therefore using L'Hospital's rule, we obtain

$$P_1(1) = \lim_{z\to 1} P_1(z) = \frac{-\lambda bE(X)F^{*\prime}(0)Q}{b\left[1 - r_2 - F^{*\prime}(0) - r_0H^{*\prime}(0)\right]} \text{ and} \tag{37}$$

$$P_2(1) = \lim_{z\to 1} P_2(z) = \frac{-r_0\lambda bE(X)H^{*\prime}(0)Q}{b\left[1 - r_2 + \lambda bE(X)F^{*\prime}(0) + r_0\lambda bE(X)H^{*\prime}(0)\right]}, \tag{38}$$

where $'$ indicate the derivative with respect to z of the respective functions. Substituting equations (37) and (38) in equation (36), we get

$$Q = \frac{b\left[1 - r_2 + \lambda bE(X)F^{*\prime}(0) + r_0\lambda bE(X)H^{*\prime}(0)\right]}{(1-b)\lambda bE(X)\left[-F^{*\prime}(0) - r_0H^{*\prime}(0)\right] + b(1 - r_2)}. \tag{39}$$

Substituting $K(1) = 1$, $K'(1) = E(X)$, $F^*(0) = 1$, $F^{*\prime}(0) = -E(s)$, $H^*(0) = 1$, $H^{*\prime}(0) = -E(v)$ in (39), we get the result, where $E(v)$ and $E(s)$ are the mean service times for FES and SOS, respectively. $E(X)$ is the mean batch size of the arriving units. □

Remark The system utilization factor ρ is given by

$$\rho = \frac{\lambda bE(X)\left[E(s) + r_0E(v)\right]}{(1-b)(\lambda bE(X))\left(E(s) + r_0E(v)\right) + b(1 - r_2)},$$

which is obtained by letting $\rho = 1 - Q$, where Q is given by (35). One observes that $\rho < 1$ is the stability condition under which the steady-state system exists.

4. Performance measures

In this section, using the PGF of the queue size distribution that we obtained in the previous section, we get the mean queue size and the waiting time of a customer in the queue.

Lemma 8 The PGF of the queue size is given by

$$P_q(z)=\frac{(-Q)\left[1-F^*(\eta(z))+r_0F^*(\eta(z))-r_0F^*(\eta(z))H^*(\eta(z))\right]}{b\left[z-r_1F^*(\eta(z))-r_2zF^*(\eta(z))-r_0F^*(\eta(z))H^*(\eta(z))\right]}. \tag{40}$$

Proof. By substituting equations (31) and (32) in the below relation and simplifying it, the equation (40) follows.

$$P_q(z)=P_1(z)+P_2(z)$$ □

Let L_q be the mean queue size which is define as following:

$$L_q=\lim_{z\to 1}\frac{d}{dz}P_q(z), \tag{41}$$

where $P_q(z)$ denote the PGF of the queue size.

Lemma 9 The mean queue size (L_q) is given by

$$L_q=\frac{L_1}{L_2}, \tag{42}$$

where $L_1 = Q[(1-r_2)(\lambda bE(X))^2E(s^2)+\lambda bE(X(X-1))E(s)+(\lambda bE(X))^2E(v^2)+\lambda bE(X(X-1))E(v)]+2r_2(\lambda bE(X)E(s))^2+2r_0(\lambda bE(X)^2E(s)E(v)]$ and $L_2=2b[1-r_2-\lambda bE(X)E(s)-r_0\lambda bE(X)E(v)]^2$.

Proof. Taking the limit of derivative of $P_q(z)$ at $z = 1$ brings equation (41) to indeterminate $\left(\frac{0}{0}\right)$ form. Then using L'Hospital's rule twice and carrying out the derivatives at $z = 1$, we obtain

$$L_q=\frac{L_1}{L_2}, \tag{43}$$

where $L_1 = [(1-r_2)[(\lambda bK'(1))^2F^{*}{''}(0)-\lambda bK''(1)F^{*}{'}(0)+r_0(\lambda bK'(1))^2H^{*}{''}(0)-r_0\lambda bK''(1)H^{*}{'}(0)]+2r_2[\lambda bK'(1)F^{*}{'}(0)]^2+2r_0(\lambda bK'(1))^2F^{*}{'}(0)H^{*}{'}(0)]Q$, $L_2=2b[1-r_2+\lambda bK'(1)F^{*}{'}(0)+r_0\lambda bK'(1)H^{*}{'}(0)]^2$. now, setting $K(1)=1$, $K'(1)=E(X)$, $K''(1)=E(X(X-1))$, $F^*(0)=1$, $F^{*}{'}(0)=-E(s)$, $F^{*}{''}(0)=E(s^2)$, $H^*(0)=1$, $H^{*}{'}(0)=-E(v)$, $H^{*}{''}(0)=E(v^2)$ in equation (43), the equation (42) is obtained. □

Let W_q be the mean of waiting time of a customer in the queue. Using Little's formula, we have

$$W_q=\frac{L_q}{\lambda bE(X)}, \tag{44}$$

where L_q is found in equation (44), $E(s^2)$ and $E(v^2)$ are the second moments of the service time FES and SOS, respectively. $E(X(X-1))$ is the second factorial moment of the batch size of the arriving units.

5. Particular cases

In this section, we derived some interesting particular cases of our results obtained in the previous sections.

Case 1 Consider $r_2 = 0$ (no feedback), an $M^X/G/1$ queueing system with SOS and balking is obtained.

$$Q = b[1 - \lambda bE(X)E(s) - r_0\lambda bE(X)E(v)], \tag{45}$$

where $\lambda bE(X)[E(s) + r_0E(v)] < 1$.

$$L_q = \frac{L_1}{L_2}, \tag{46}$$

where $L_1 = Q[(\lambda bE(X))^2E(s^2) + \lambda bE(X(X-1))E(s) + r_0(\lambda bE(X))^2E(v^2) + r_0\lambda bE(X(X-1))E(v) + 2r_0(\lambda bE(X))^2E(s)E(v)]$ and $L_2 = 2b[1 - \lambda E(X)E(s) - r_0\lambda E(X)E(v)]^2$.

Case 2 Consider $r_2 = 0$ (no feedback), $b = 1$ (no balking), the customers arrive at the system one by one, then $c_1 = 1$ and $c_i = 0$ for all $i > 1$. Consequently $K(z) = z$, $E(X) = 1$, $E(X(X-1)) = 0$, the model reduces to an $M/G/1$ queueing system with SOS.

$$Q = 1 - \lambda E(s) - r_0\lambda E(v), \tag{47}$$

where $\lambda[E(s) + r_0E(v)] < 1$.

$$L_q = \frac{L_1}{L_2}, \tag{48}$$

where $L_1 = Q[((\lambda)^2E(s^2) + (\lambda)^2E(v^2) + 2r_0(\lambda)^2E(s)E(v)]$ and $L_2 = 2[1 - \lambda[E(s) + r_0E(v)]]^2$.

We notice this result agrees with the result of An $M/G/1$ queue with a second optional service with general service time distribution (see [26]).

Case 3 Consider $r_2 = 0$ (no feedback), $r_0 = 0$ (no SOS), $b = 1$ (no balking), a simple $M^X/G/1$ queue follows.

$$Q = 1 - \lambda E(X)E(s), \tag{49}$$

where $\lambda E(X)E(s) < 1$.

$$L_q = \frac{L_1}{L_2}, \tag{50}$$

where $L_1 = [(\lambda E(X))^2E(s^2) + \lambda E(u/u-1)E(s)]Q$ and $L_2 = 2[1 - \lambda E(X)E(s)]^2$.

We note that this result agrees with the known result of the $M^X/G/1$ queue.

Case 4 Consider $r_2 = 0$ (no feedback), $r_0 = 0$ (no SOS), one gets an $M^X/G/1$ queueing system with balking.

$$Q = 1 - \lambda bE(X)E(s), \tag{51}$$

where $\lambda bE(X)E(s) < 1$.

$$L_q = \frac{L_1}{L_2}, \tag{52}$$

where $L_1 = [(\lambda E(X))^2E(s^2) + \lambda E(X(X-1))E(s)]Q$ and $L_2 = 2b[1 - \lambda bE(X)E(s)]^2$.

Case 5 Consider $r_0 = 0$ (no SOS), the model reduces to an $M^X/G/1$ queueing system with feedback and balking.

$$Q = \frac{b[1 - r_2 - \lambda bE(X)E(s)]}{[1-b][\lambda bE(X)E(s)] + b(1 - r_2)}, \tag{53}$$

where $r_2 + \lambda bE(X)E(s) < 1$.

$$L_q = \frac{L_1}{L_2}, \tag{54}$$

where $L_1 = (1-r_2)[(\lambda bE(X))^2E(s^2) + \lambda bE(X(X-1))E(s)]Q$ and $L_2 = 2b[1-r_2-\lambda bE(X)E(s)]^2$.

Case 6 Consider $r_0 = 0$ (no SOS), $b = 1$ (no balking), a feedback model in $M^X/G/1$ queue is obtained.

$$Q = \frac{[1-r_2-\lambda E(X)E(s)]}{1-r_2}, \tag{55}$$

where $r_2 + \lambda E(X)E(s) < 1$.

$$L_q = \frac{L_1}{L_2}, \tag{56}$$

where $L_1 = [(1-r_2)(\lambda E(X))^2E(s^2) + \lambda E(X(X-1))E(s) + 2r_2(\lambda bE(X(X-1))E(s))^2]Q$ and $L_2 = 2[1-r_2-\lambda E(X)E(s)]^2$.

We note that this result agrees with the result of $M^X/G/1$ queue with feedback and optional server vacations (see [13]).

6. The cost model and numerical results

To achieve the optimal service rate in FES and SOS with a minimum expected cost function, we develop the expected cost function per unit time as

$$f(\mu, \beta) = CL + C_1\mu + C_2\beta, \tag{57}$$

where :

- C = cost per unit time per customer present in the queue.
- C_1 = cost per unit time during FES.
- C_2 = cost per unit time during SOS.

The cost minimization problem $f(\mu, \beta)$ can be presented mathematically as

$$f(\mu^*, \beta^*) = \underset{s.t\ \mu,\ \beta>0}{\text{Minimize}}\, f(\mu, \beta). \tag{58}$$

We use the Quasi-Newton method to search for the optimum values of (μ, β). For more details of Quasi-Newton method, one may refer Lewis and Overton [25].

6.1 *Numerical results and discussion*

Some numerical illustrations with discussion based on Q, L_q, and W_q are provided with the purpose to illustrate the effect of the parameters (λ, μ, β, b, r_0, r_1, r_2) on Q, L_q and W_q. For this purpose, we consider exponential distribution for the service time FES and SOS as

$$E(s) = \frac{1}{\mu},\ E(s^2) = \frac{2}{\mu^2},\ E(v) = \frac{1}{\beta},\ E(v^2) = \frac{2}{\beta^2}.$$

From Table 1, we obtain that the minimum cost per unit time is $f(\mu^*, \beta^*) = 55.8463$ at $(\mu^*, \beta^*) = (1.63943, 1.36848)$ achieved at tenth iteration.

Table 1. Quasi-Newton method

Iteration	0	1	2	3
μ	2.0000	1.62709	1.67468	1.64685
β	1.0000	1.17857	1.32044	1.35784
$f(\mu, \beta)$	58.7500	56.3289	55.8812	55.8480
$\frac{\partial f}{\partial \mu}$	5.56412	-1.64937	0.780597	0.169144
$\frac{\partial f}{\partial \beta}$	-11.9156	-5.5587	-0.889916	-0.187841
Hessian	$\begin{bmatrix} 14.9207 & 6.67271 \\ 6.67271 & 66.7271 \end{bmatrix}$	$\begin{bmatrix} 34.6534 & 8.42502 \\ 8.42502 & 39.1811 \end{bmatrix}$	$\begin{bmatrix} 28.0483 & 5.24947 \\ 5.24947 & 23.7945 \end{bmatrix}$	$\begin{bmatrix} 29.8940 & 5.14955 \\ 5.14955 & 21.5380 \end{bmatrix}$
Iteration	4	5	6	7
μ	1.64119	1.63975	1.6395	1.63944
β	1.36657	1.36805	1.3684	1.36846
$f(\mu, \beta)$	55.8464	55.8463	55.8463	55.8463
$\frac{\partial f}{\partial \mu}$	0.043693	0.00751339	0.00181581	0.000310087
$\frac{\partial f}{\partial \beta}$	-0.0312059	-0.00744531	-0.00126942	-0.000305849
Hessian	$\begin{bmatrix} 30.28 & 5.11987 \\ 5.11987 & 21.0468 \end{bmatrix}$	$\begin{bmatrix} 30.389 & 5.12093 \\ 5.12093 & 20.9703 \end{bmatrix}$	$\begin{bmatrix} 30.4065 & 5.11997 \\ 5.11997 & 20.9509 \end{bmatrix}$	$\begin{bmatrix} 30.411 & 5.12003 \\ 5.12003 & 20.9478 \end{bmatrix}$
Iteration	8	9	10	
μ	1.63943		1.63943	
β	1.36848	1.36848	1.36848	
$f(\mu, \beta)$	55.8463	55.8463	55.8463	
$\frac{\partial f}{\partial \mu}$	0.0000747509	0.0000127618	0.00000307601	
$\frac{\partial f}{\partial \beta}$	-0.0000522121	-0.0000125848	-0.00000214857	
Hessian	$\begin{bmatrix} 30.4117 & 5.11999 \\ 5.11999 & 20.947 \end{bmatrix}$	$\begin{bmatrix} 30.4119 & 5.11999 \\ 5.11999 & 20.9469 \end{bmatrix}$	$\begin{bmatrix} 30.4119 & 5.11999 \\ 5.11999 & 20.9469 \end{bmatrix}$	

Table 2 shows the impact of feedback probability r_2 on the minimum expected cost function $f(\mu^*, \beta^*)$ for different values of joining probability b. We observe that the optimal service rates μ^*, β^* and expected cost $f(\mu^*, \beta^*)$ increase as both r_2 and b increase. Particularly, as r_2 increases, customers rejoin the queue as feedback, which leads to an increase in the service rates μ^*, β^*, and cost to balance the system profitability. Here we take; the service time (FES and SOS) follow Exponential distribution and $\lambda = 2$, $r_0 = 0.4$, $\mu = 2$, $\beta = 1$, $E(X) = 1$, $E(X(X-1)) = 0$.

Table 3 shows the impact of the probability of feedback (r_2) and the probability of joining SOS (r_0) on the mean queue size (L_q), considering that the probability of departure remains constant. We observe that as r_2 increases, r_0 decreases. This situation leads to an increase in L_q. This is because the patients are not satisfied with the primary test results, and for all patients who tested positive or negative, they want to confirm whether they are positive or negative by repeating the primary test, which in turn, increases the mean queue size at the hospital.

Here we take; the service time (FES and SOS) follow exponential distribution and $\lambda = 2$; $\mu = 5$; $\beta = 4$; $b = 0.5$; $E(X) = 1$, $E(X(X-1)) = 0$.

Table 4 depicts the effect of the probability of joining the queue and the arrival rate on the L_q. In this table, we observe that as both probability of joining a queue and the rate of flow of patients to the queue increase, leads to an increase the mean queue size. Particularly as the flow of patients into the queue increases, it indicates a high number of suspected people in society, which in turn increases the congestion at the hospital.

Here we take; the service time (FES and SOS) follow exponential distribution and $\lambda = 2, 3, 4$; $r_0 = 0.6$; $r_2 = 0.2$; $\mu_1 = 5$; $\beta = 4$; $b = 0.3, 0.4, 0.5$; $E(X) = 1$, $E(X(X-1)) = 0$.

Table 2. Impact of r_2 and b on the expected cost

r_2	b	μ^*	β^*	$f(\mu^*, \beta^*)$
$r_2 = 0.1$	$b = 0.20$	1.49909	1.29135	51.9987
	$b = 0.25$	1.76145	1.50907	59.7216
	$b = 0.30$	2.01407	1.71764	67.0097
$r_2 = 0.2$	b = 0.20	1.63943	1.36848	55.8463
	$b = 0.25$	1.92951	1.60415	64.2728
	$b = 0.30$	2.20919	1.83057	72.2418
$r_2 = 0.3$	$b = 0.20$	1.81612	1.46562	60.6373
	$b = 0.25$	2.14139	1.72413	69.9516
	$b = 0.30$	2.45551	1.97325	78.7812

Table 3. The impact of r_0 and r_2 on Q, L_q and W_q

r_2	r_0	Q	ρ	L_q	W_q
0.1	0.5	0.469388	0.530612	0.264419	0.264419
0.2	0.4	0.454545	0.545455	0.290909	0.290909
0.3	0.3	0.435897	0.564103	0.328808	0.328808
0.4	0.2	0.411765	0.588235	0.386555	0.386555
0.5	0.1	0.379310	0.620690	0.482759	0.482759

Table 4. Impact of λ and b on L_q

λ	b	Q	ρ	L_q	W_q
$\lambda = 2$	$b = 0.3$	0.4573640	0.542636	0.157667	0.2627780
	$b = 0.4$	0.4262300	0.573770	0.252207	0.3152597
	$b = 0.5$	0.3913040	0.608696	0.386473	0.3864730
$\lambda = 3$	$b = 0.3$	0.3159610	0.684039	0.362672	0.4029690
	$b = 0.4$	0.2657340	0.734266	0.662495	0.5520790
	$b = 0.5$	0.2075470	0.792453	1.234990	0.8233280
$\lambda = 4$	$b = 0.3$	0.2134830	0.786517	0.709639	0.5913660
	$b = 0.4$	0.1463410	0.853659	1.626020	1.0162600
	$b = 0.5$	0.0666667	0.933333	5.333330	2.6666700

In Figure 1, we plot the Iterations versus the expected cost. We notice that the expected cost decrease gradually from 58.75 value at the beginning up to 55.8463 value and with increasing iterations it reaches a minimum expected cost value of 55.8463, where there is no further decrease as iteration increases.

In Figure 2, we show the effect of joining probability (b) on the expected queue length variation of inequality of departure probability r_1 and feedback probability r_2. We observe that as b increases, the expected queue length

L_q increases, as we expected. Furthermore, we remark that the expected queue length L_q gets reduced as $r_2 < r_1$. The implication of $r_2 < r_1$ is that when the majority of patients test negative at the primary test, they opt to leave the hospital, resulting on reducing congestion at the hospital. Here we take; the service time (FES and SOS) follow Exponential distribution and $\lambda = 2$, $\mu = 5$, $\beta = 4$, $E(X) = 1$, $E(X(X - 1)) = 0$.

Also in Figures 3 and 4, we show the effect of service rate on L_q for different b. We observe that L_q decreases as the service rate increases for both FES and SOS. Further, we notice that as b increases, the L_q increases, which in turn reflects the intuition. Here we take; the service time (FES and SOS) follow Exponential distribution and $\lambda = 2$; $r_0 = 0.6$, $r_2 = 0.2$, $\beta = 4$, $E(X) = 1$, $E(X(X - 1)) = 0$ in Figure 3, and we take $\lambda = 2$, $r_0 = 0.6$, $r_2 = 0.2$, $\mu = 4$, $E(X) = 1$, $E(X(X - 1)) = 0$ in Figure 4.

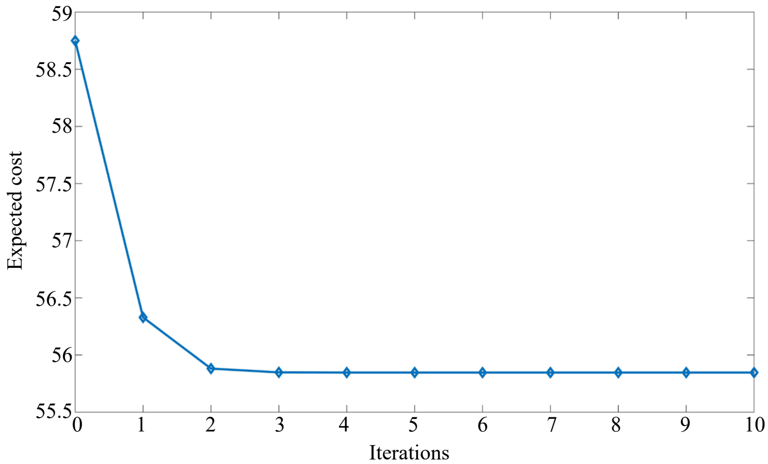

Figure 1. The impact of departure probability r_1 and feedback probability r_2 on the expected queue (L_q)

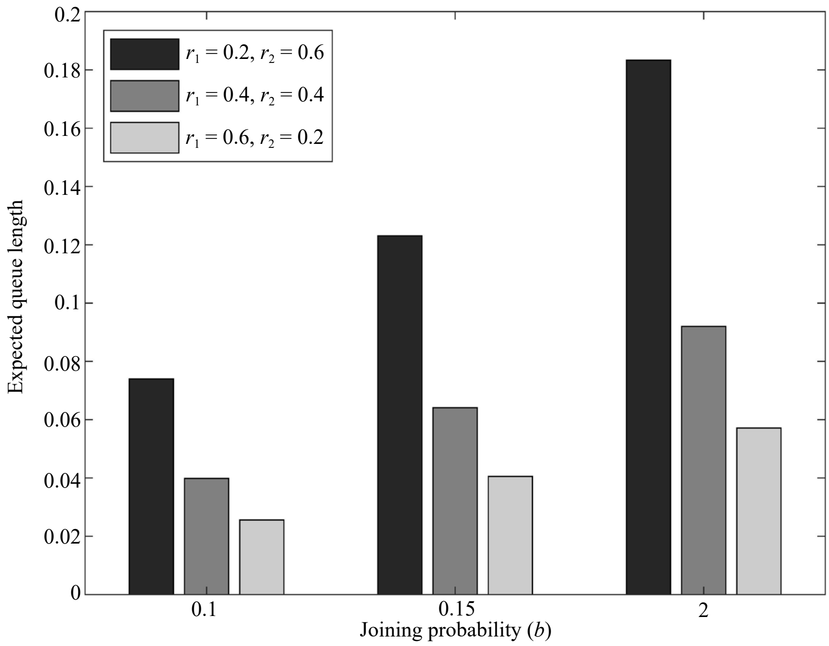

Figure 2. The impact of departure probability r_1 and feedback probability r_2 on the expected queue (L_q)

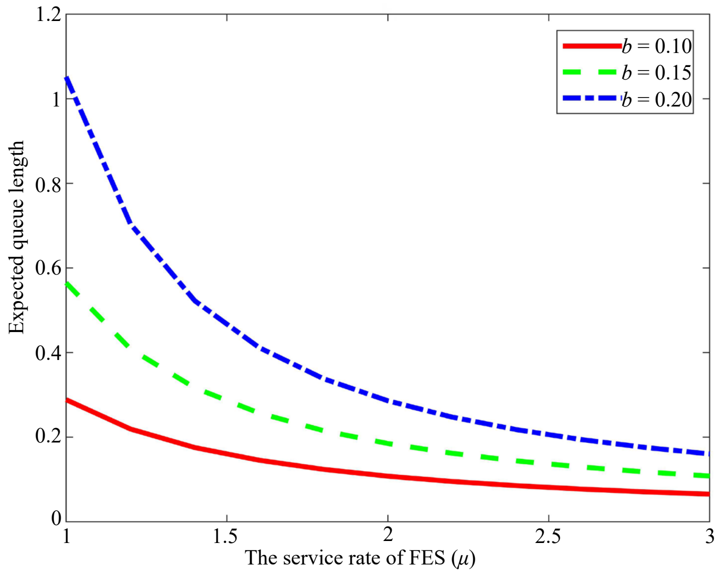

Figure 3. The effect of FES rate (μ) on (L_q) in different joining probability b

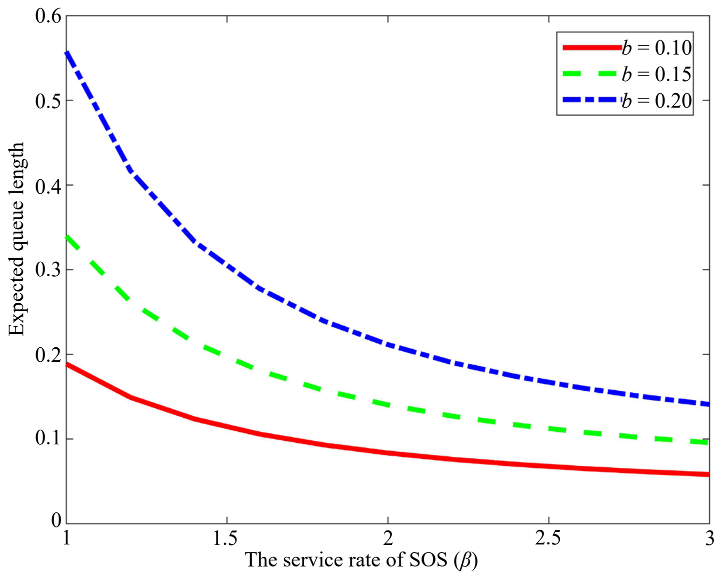

Figure 4. The effect of the SOS rate (β) on (L_q) in different joining probability b

7. Conclusion

In this paper, we analyzed the steady-state behavior of a batch arrival queue single server non-Markovian service queue with a second optional service, balking, and feedback. Using the supplementary variable technique and the probability generating functions, we have obtained the mean of the queue size and waiting time of a customer in the queue. The cost model was presented to determine the optimal service rates to minimize the expected cost. Finally, the numerical results through graphical illustrations and tables were presented. In future work, we will incorporate a batch arrival general service queue with balking, feedback, and SOS, adding the concepts of working vacations and vacation interruption. Also, we will consider the transient state in the current model.

8. Limitations

While applying queueing theory to simulate a queuing system's behavior to improve efficiency and service levels, it has certain drawbacks. This is because the traditional queueing theory could be too limited to accurately represent actual circumstances. In addition, we model the system and derive solutions simply based on the assumptions of queueing theory. However, the queueing theory is a growing area that uses more complex models to solve the challenges by providing enough approximations solution. Based on our models discussed in this paper it is very difficult to extract the probabilities into analytical explicit expressions, instead, we have obtained the probability generating function of the number of customers in the queue. Therefore, queue modeling is frequently seen as a difficult research project due to the intricacy of the mathematics entailed in the theory.

Conflict of interest

The authors declare no competing financial interest.

References

[1] Daley DJ. General customer impatience in queue GI/G/1. *Journal of Applied Probability*. 1965; 2(1): 186-205.

[2] Choudhury A. Impatience in single server queueing model. *American Journal of Mathematical and Management Sciences*. 2008; 28(1-2): 177-211.

[3] Soundararajan PM, Josephine SA. Analysis of an M/G/1 retrial queue with second optional service and customer feedback under Bernoulli vacation schedule. *Malaya Journal of Matematik*. 2019; 7(4): 795-807.

[4] Baruah M, Madan KC, Eldabi T. Balking and re-service in a vacation queue with batch arrival and two types of heterogeneous service. *Journal of Mathematics Research*. 2012; 4(4): 114-124.

[5] Singh CJ, Jain M, Kumar B. Analysis of $M^X/G/1$ queueing model with balking and vacation. *International Journal of Operational Research*. 2014; 19(2): 154-173.

[6] Inoue Y, Boxma O, Perry D, Zacks SH. Analysis of $M^X/G/1$ queues with impatient customers. *Queueing System*. 2017; 89: 303-350.

[7] Laxmi PV, Qrewi HA, George AA. Analysis of markovian batch service queue with feedback and second optional service. *Reliability: Theory & Applications*. 2022; 17(2): 507-518.

[8] Shanmugasundaram S, Sivaram G. $M^X/G/1$ feedback queue when server is off and on vacation. *International Journal of Applied Engineering Research*. 2020; 15(10): 1025-1028.

[9] Arivudainambi D, Godhandaraman P. A batch arrival retrial queue with two phases of service, feedback and K optional vacations. *Applied Mathematical Sciences*. 2012; 6(22): 1071-1087.

[10] Ayyappan G, Devipriya G, Subramanian AMG. Analysis of single server queueing system with batch service under multiple vacations with loss and feedback. *Mathematical Theory and Modeling*. 2014; 14(11): 79-90.

[11] Ayyappan G, Shyamala S. Transient solution of an $M^X/G/1$ queueing model with feedback, random breakdowns, Bernoulli schedule server vacation and random setup time. *International Journal of Operational Research*. 2016; 25(2): 196-211.

[12] Jain M, Singh M. Transient analysis of a markov queueing model with feedback, discouragement and disaster. *International Journal of Applied and Computational Mathematics Volume*. 2020; 6: 31. Available from: https://doi.org/10.1007/s40819-020-0777-x.

[13] Madan KC, Al-Rawwash M. On the $M^X/G/1$ queue with feedback and optional server vacations based on a single vacation policy. *Applied Mathematics and Computation*. 2005; 160(3): 909-919.

[14] Rajan BS, Ganesan V, Rita S. Batch arrival poisson queue with breakdown and repairs. *International Journal of Mathematics in Operational Research*. 2020; 17(3): 424-435.

[15] Abdollahi S, Salehi Rad MR. Analysis of a batch arrival retrial queue with two-phase services, feedback and admission. *Bulletin of the Iranian Mathematical Society*. 2021; 48(3): 791-804.

[16] Khalaf R, Madan KC, Lucas C. An $M^X/G/1$ queue with bernoulli schedule general vacation times, general extended vacations, random breakdowns, general delay times for repairs to start and general repair times. *Journal of Mathematics Research*. 2011; 3(4): 8-20.

[17] Suganya S. $M^X/G/1$ with second optional service, multiple vacation, breakdown and repair. *International Journal of Research in Engineering and Science*. 2014; 2(11): 70-77.
[18] Baruah M, Madan KC, Eldabi T. A batch arrival queue with second optional service and reneging during vacation periods. *Investigacion Operacional*. 2013; 34(3): 244-258.
[19] Zirem D, Boualem M, Aïssani D. $M^X/G/1$ queue with breakdowns and repairs. *International Journal of Advances in Computer Science & Its Applications*. 2016; 6(3): 18-21.
[20] Madan KC, Eldabi T. On an $M^X/G/1$ queue with a random set up time, random breakdowns and delayed deterministic repairs. *Journal of Mathematical and Computational Science*. 2021; 11(5): 6568-6580.
[21] Bouchentouf AA, Houalef M, Guendouzi A. Analysis of a batch arrival multi-server queueing system with waiting servers, synchronous working vacations and impatient customers. *Acta Universitatis Sapientiae, Mathematica*. 2022; 14(1): 37-60.
[22] Bouchentouf AA, Medjahri L, Boualem M, Kumar A. Mathematical analysis of a Markovian multi-server feedback queue with a variant of multiple vacations, balking and reneging. *Discrete & Continuous Models & Applied Computational Science*. 2022; 30(1): 21-38.
[23] Ke JC, Liu TH, Su S, Zhang ZG. On retrial queue with customer balking and feedback subjected to server breakdowns. *Cummunications in Statistics-Theory and Methods*. 2022; 51(17): 6049-6063.
[24] Jain M, Kumar A. Unreliable server $M^{[X]}/G/1$ retrial feedback queue with balking, working vacation and vacation interruption. *Proceedings of the National Academy of Science, India Section A: Physical Sciences*. 2023; 93: 57-73. Available from: https://doi.org/10.1007/s40010-022-00777-w.
[25] Lewis AS, Overton ML. Nonsmooth optimazation via quasi-Newton methods. *Mathematical Programming*. 2013; 141: 135-163. Available from: https://doi.org/10.1007/s10107-012-0514-2.
[26] Al-jararha J, Kailash M. An M/G/1 queue with second optional service with general service time distribution. *International Journal of Information and Management Sciences*. 2003; 14(2): 47-56.