Queue with Batch Service, Batch Size Determined by Emergency Customers

Sinu Lal T. S.[1], A. Krishnamoorthy[1], and V. C. Joshua[1]

Department of Mathematics, CMS College Kottayam 686001, Kerala, India
{sinulal, krishnamoorthy, vcjoshua}@cmscollege.ac.in
http://www.cmscollege.ac.in

Abstract. We consider a queueing system offering batch service for heterogeneous customers. Two types of customers called ordinary customers and emergency customers arrive to the system according to a marked Markovian arrival process of order 2. There is an infinite queue and finite buffer in front of the station. Ordinary customers arrive to the infinite queue and emergency customers to the finite buffer. Service is provided in batches with maximum batch size K. If the number of emergency customers is greater than or equal to K, the first K of them will be taken together into service and if it is less than K, ordinary customers are taken along with emergency customers so as to maximize the batch size. In the absence of emergency customers, if there are at least K ordinary customers, then first K of them will be served next. For a batch of size i, $1 \leq i \leq K$, the service time follows a phase type distribution with representation $PH(\alpha_i, S_i)$. The model possesses the characteristics of a vacation queueing model. The system is analyzed using Matrix analytic Method. Performance characteristics are derived. The model is numerically illustrated with suitable example.

Keywords: Batch Service, Marked Markovian Arrival Process, Emergency Customer, Matrix Analytic Methods, Phase Type Distributions

1 Introduction

In this paper we study the mathematical model of a courier delivery system with emergency arrivals and bulk service. The model has found applications in priority based signal transmission systems as well. Numerous real life problems can be modelled as queues with bulk service and a rich literature on bulk queues is available. Delivery systems of online marketing, courier services, transportation vehicles and airline services etc. are typical examples for bulk service systems. Each of these systems possesses its own special features. In delivering items, the emergency orders are usually served with priority. Sometimes instantaneous service is called for the emergency customers due to urgency (e.g. delivery of highly perishable items like samples of body fluids for diagnostic service, animal sperm, medical essentials upon order placing etc.). It is effective to incorporate threshold policies in bulk queues to maximize revenue generation and a great exploration for mathematical results is possible in this direction.

One of the early works in bulk queues is due to M. F. Neuts [13]. In this work, the customers are taken together to the service station until queue length reaches a specified level (L). If the queue grow beyond L and reaches K all the customers up to K are taken together, the customers beyond K must wait in queue. H. Gold and P. Tran-Gia [6] studied an $M/G[a,b]/1-S$ queueing system. The main motivation for this model is the manufacturing systems with batch serving stations (machines for computer components and chip production).

M. L. Chaudhry and U. C. Gupta [5] describe an $M/G^{a,b}/1/N$ queue and queue is studied using supplementary variables and embedded Markov chain techniques. A finite buffer $M/G/1$ queue with general bulk service rule and single vacation is studied by U. C. Gupta and K. Sidkar [8] and in this model batch size is restricted to a range of values, and when the queue length falls below the infimum, the server goes for a vacation.

W. B. Powell [16] studies a general class of vehicle dispatching strategies for bulk arrivals, bulk service queue. D. J. Van De Rzee, A. Van Haeten and P. C. Schuur [22] describe control strategies for reading the cost of multi server batch operation systems. S. Kuppa and G. Dattatreya [10] describe frame aggregation in unsaturated WLAN's with finite buffers and in this model method for aggregation of more than one data form for transmission is described as an application for bulk queues. A. Banerjee and U. C. Gupta [1] analyze a single server finite buffer queue with Poisson arrival and bulk service and batches are arbitrarily distributed and depends the size of the current batch in service and the aim is to reduce congestion in the system.

J. Baetens, B. Steyaert, D. Claeys, and H. Bruneel [2] analyze a discrete time $BMAP/G^{l,c}/1$ queue in which service time of a batch depends the current batch size and a timing mechanism to reduce waiting time is also proposed. M. Yu and A. S. Alfa [23] describe an algorithm for computing the queue length distribution at various time epochs in a $DMAP/G^{a,b}/1/N$ queue with batch size dependent service times and the impact of correlation on system characteristic is also discussed. [11,7,4] and [2] give live descriptions of batch service queues with dependent batch size.

A. Sikadar and S. K. Samantha [19] studied a bulk service queue with service vacation and vacation starts when all the customers are exhaustively served. In [17] G. V. Krishna Reddy, R. Nadarajan and P. R. Kandasamy considered a bulk service system with heterogeneous arrivals and bulk service is provided to the class of customers with low priority. In [2] J. Baetens, B. Steyaert, D. Claeys, and H. Bruneel analyze a two class batch service queueing model with variable server capacity and batch size is determined by number of consecutive same-class customers. In [3] J. Baetens, B. Steyaert, D. Claeys, and H. Bruneel analyze delay of a random customer in a two class batch service queueing model with variable service capacity and batch size is determined by the length of the sequence of same class customers. An excellent survey by S. Sasikala, K. Indhira on bulk service queueing models is demonstrated in [18].

A phase type distribution may be defined as the distribution of the time until absorption takes place in a Markov process with a finite state space and a single

absorption state defined over nonnegative real line. A phase type distribution with transient states $\{1, 2, \ldots, n\}$ and an absorbing state $n+1$ is represented by a two tuple of the form $(\alpha,\ T)$, where α is the probability vector of length n according to which the process selects the initial state from $\{1, 2, \ldots, n\}$ and T is an $n \times n$ matrix such that $\begin{pmatrix} T & T^0 \\ \mathbf{0} & 0 \end{pmatrix}$ generates the process, given the column vector T^0 satisfies the condition $Te+T^0 = \mathbf{0}$, where e is the vector of ones. (α, T) is called the representation of the phase type distribution. The distribution F of time until the chain gets absorbed into the state $n+1$ is given by $F(x) = 1 - \alpha exp(Tx)e,\ x \geq 0$. The set of all phase type distributions is a dense subset of the set of all distributions on the non-negative real line and hence it is a best tool to approximate any arbitrary distribution in this set. For more descriptions on phase type distributions see [14].

A Marked Markovian arrival process (MMAP) is a stochastic point process with heterogeneous arrivals in discrete or continuous time. MMAP may be described as follows: Let C be set of indices which describes different types of customers. Let $N_h(t)$ be the number of arrivals of type h in $[0, t]$ such that $N_h(0) = 0, \forall h \in C$. Consider the set of nonnegative matrices $\{D_h : h \in C\}$. Let D_0 is a matrix with nonnegative off-diagonal elements and negative diagonal elements, and $D = D_0 + \sum_{h \in C} D_h$ be an infinitesimal generator of order m, and $\{I(t) : t \geq 0\}$ be a continuous time Markov chain defined by D. Then $(D_0, D_h, h \in C)$ defines an MMAP $\{N_h(t), h \in C, I(t), t \geq 0\}$. The Markov chain $\{I(t) : t \geq 0\}$ is called the underlying Markov chain of $\{N_h(t), h \in C, I(t), t \geq 0\}$. For description of MMAP model see [9]. Analysis of queues using matrix analytic method can be seen in [20,21]

Upcoming sections are arranged as follows: Section 2 describes the mathematical model stability condition and stationary distribution is also obtained in this section. Performance characteristics are included in Section 3. Service time analysis is done in Section 4. Waiting time analysis is given in Section 5. The model is numerically illustrated in Section 6. Section 7 concludes the work.

2 Mathematical Model

The customers arriving to the system are of two types, namely ordinary customers and emergency customers. The arrival is according to a MMAP determined by the matrices D_0, D_1 and D_2. D_0 gives the rate of transitions in the underlying process without an arrival, D_1 and D_2 gives the rate of transitions for ordinary and emergency arrivals respectively. The matrix $D = D_0 + D_1 + D_2$ is an infinitesimal generator of the underlying process. If θ is the steady state distribution of D then $\lambda_i = \theta D_i e, i = 1, 2$ is the fundamental rates of ordinary and emergency arrivals.

If the covariance of number of ordinary customers $n_1(t)$ and that of emergency customers $n_2(t)$ is given by

$$cov(n_1(t), n_2(t)) = -(2\lambda_1\lambda_2 + \theta(\sum_{k=1}^{2} D_k(D - e\theta)^{-1}D_{3-k})e)t + \theta(\sum_{k=1}^{2} D_k(D - e\theta)^{-1}exp(Dt_I(D - e\theta)^{-1})D_{3-k}))e. \quad (1)$$

The service discipline is as follows: service is provided in batches with maximum batch size K. If the number of emergency customers is greater than or equal to K, the first K of them will be taken together into service and if it is less than K, ordinary customers are taken along with emergency customers so as to maximize the batch size. In the absence of emergency customers, if there are at least K ordinary customers, then first K of them will be served next. Service time for a batch of size j is distributed with a phase type representation $PH(\alpha_j, S_j)$.

The following are system descriptors at the time t.

- $N_1(t)$ - Length of the infinite queue.
- $N_2(t)$ - Number of customers in finite buffer.
- $I(t)$ - Status of the server.
- $S(t)$ - Phase of the service.
- $a(t)$ - Phase of MMAP.

The server status is defined as follows.

$$I(t) = \begin{cases} 0, \text{ if the server is idle} \\ i, \text{ if the sever is busy with } i \text{ customers}, 1 \leq i \leq K. \end{cases}$$

We define

$$\mathcal{Y}(t) = \{N_1(t), N_2(t), I(t), S(t), a(t)\}.$$

Then $\{\mathcal{Y}(t), t \geq 0\}$ is a continuous time Markov process on the state space

$$\mathcal{S} = \cup_{i \geq 0}\mathcal{L}(i).$$

For each $i < K$,

$$\mathcal{L}(i) = \mathcal{L}_1(i) \cup \mathcal{L}_2(i),$$

where

$$\mathcal{L}_1(i) = \{(i, 0, l), 1 \leq l \leq a\}$$

and

$$\mathcal{L}_2(i) = \{(i, i_1, 1, j, k), 0 \leq i_1 \leq M, 1 \leq k \leq a, 1 \leq j \leq r\}.$$

$\mathcal{L}_1$ contains the states in which there are no priority customers and the server is idle and $\mathcal{L}_2(i)$ corresponds to states in which server is active and in this case buffer can be empty or non empty.

For $i \geq K$,

$$\mathcal{L}(i) = \{(i, i', 1, j, k), 0 \leq i' \leq M, 1 \leq k \leq a, 1 \leq j \leq r\}.$$

The states in all the levels greater than or equal to K will only contain states in which server is busy.

The infinitesimal generator Q of $\mathcal{Y}$ takes the form

$$Q = \begin{pmatrix}
A_{00} & A_{01} & & & & & & & & & \\
A_{10} & A_{11} & A_{12} & & & & & & & & \\
A_{20} & & A_{21} & A_{22} & & & & & & & \\
\vdots & & & \ddots & \ddots & & & & & & \\
A_{K-10} & & & & A_{K-11} & A_{K-12} & & & & & \\
 & A_1^{\#} & & & & A_{K1} & A_0 & & & & \\
 & & A_2^{\#} & & & & A_{K+11} & A_0 & & & \\
 & & & \ddots & & & & \ddots & \ddots & & \\
 & & & & A_K^{\#} & & & & A_{2K-11} & A_0 & \\
 & & \vdots & & & & & & & \ddots & \ddots
\end{pmatrix},$$

$$A_{00} = D_0, A_{01} = \begin{pmatrix} D_1 & \alpha_1 \otimes D_2 \end{pmatrix},$$

$$A_{j0} = e_r \otimes \big(\sum_{i=0}^{j} S_j^0 \otimes I_a \big), 1 \leq j \leq K,$$

$$A_{i1} = diag\big(D_0, (\sum_{j=1}^{i} S_i) \oplus D_0\big), 1 \leq i \leq K-1,$$

$$A_{i1} = diag\big(D_0, S_K \oplus D_0\big), i > K,$$

$$A_{i2} = \begin{pmatrix} D_1 & \alpha_i \otimes D_2 \\ & I_r \otimes (D_1 + D_2) \end{pmatrix}, 1 \leq i \leq K-1,$$

$$A_{K2} = \begin{pmatrix} \alpha_K \otimes (D_1 + D_2) \\ I_r \otimes (D_1 + D_2) \end{pmatrix}, A_{i2} = \begin{pmatrix} I_r \otimes (D_1 + D_2) \end{pmatrix}, i > K,$$

$$A_l^{\#} = \begin{pmatrix} e_r \otimes (S_K^0 \otimes I_a) & e_r \otimes (\alpha_l \otimes (S_K^0 \otimes I_a)) \end{pmatrix}, K+1 \leq l \leq 2K,$$

$$A_l^{\#} = \begin{pmatrix} e_r \otimes (S_K^0 \otimes I_a) & e_r \otimes (\alpha_l \otimes (S_K^0 \otimes I_a)) \end{pmatrix}, l > 2K.$$

Note that the Kronecker sum $\oplus$ and product $\otimes$ are used. The non zero blocks in the above matrix arise due to the transitions described below.

From	To	Transition rate
(i,0,0,l), $0 \leq i < K-1$	(i,i',0,k)	$D_0(l,k)$
(i,0,0,l), $i \geq K-1$	(i,1,j,k)	$D_1(l,k)$
(h,0,0,l), $h \geq 0$	(h,1,j,k)	$\alpha_h^j D_2(l,k)$
(i,0,g,j,k), $1 \leq i \leq K-1$	(i,0,g,0,k)	$S_g^0(j)$
(i,i',g,j,k), $1 \leq k,k^* \leq a$	(i,i',g,j,k^*)	$U(k,k^*), U = (S_g) \oplus D_0$
(i,i',g,j,k), $\leq k,k^* \leq a$	$(i,1,j,k^*)$	$U(k,k^*), U = S_i \oplus D_0, 1$
(i,1,j,k), $K+1 \leq i \leq 2K-1$	$(0,0,k)$	S_K^{j0}
(i,1,j,k), $K+1 \leq i \leq 2K-1$	(0,0,k)	S_m^{j0}
(i,1,j,k),$K+1 \leq i \leq 2K-1$, $1 \leq k,k^* \leq a$	$(i,1,j,k^*)$	$V(k,k^*), V = \alpha_{i-K} \otimes (S_K^0 \otimes I_a)$
(i,1,j,k), $0 \leq i \leq K-2$	$(i,1,j,k^*)$	$\Omega(k,k^*), \Omega = I_r \otimes (D_1 + D_2)$
(i,1,j,k),$0 \leq i \leq K-2$	$(i,0,k^*)$	$\Omega'(k,k^*), \Omega' = I_r \otimes (D_0)$
(i,1,j,k),$i > K-2$	$(i,1,j,k^*)$	$\Omega(k,k^*), \Omega = I_r \otimes (D_1 + D_2)$

Clearly in the infinitesimal generator Q there are non-zero blocks $A_j^*, 1 \leq j \leq K$, due to which Q loses its QBD structure. In order to gain a QBD structure for the matrix Q, the levels are redefined by merging them appropriately. The levels $lK+1$ to $(l+1)K$, $l = 0, 1, 2, \ldots$, are merged together, and modified generator Q' is described below.

After merging of cells the modified form of Q take the structure

$$Q' = \begin{pmatrix} A_{00} & A_{01} & & & & \\ B_{10} & B_{11} & B_{12} & & & \\ & B_{20} & B_{21} & B_0 & & \\ & & B_2 & B_1 & B_0 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix}.$$

$$A_{00} = D_0, A_{01} = (D_1\alpha_1 \otimes D_2), B_{10} = [A_1^* A_2^* \ldots A_K^*]^T,$$

$$B_{11} = diag(A_{11}, A_{21}, \ldots, A_{K1}) + (\Omega_0 + diag(A_{12}, A_{22}, \ldots, A_{k-12})),$$

$$B_{21} = diag(A_{K+11}, A_{K+21}, \ldots, A_{2K1}) + (\Omega_0 + diag(A_{12}, A_{22}, \ldots, A_{k-12})),$$

$$B_1 = diag(A_{2K+11}, A_{2K+21}, \ldots, A_{3K1}) + (\Omega_0 + diag(A_{12}, A_{22}, \ldots, A_{k-12})),$$

$$B_{12} = diag(diag(D_0, I_r \otimes (D_1 + D_2), \ldots, diag(D_0, I_r \otimes (D_1 + D_2)),$$

$$B_2 = diag(A_1^\#, A_2^\#, \ldots, A_K^\#),$$

$$B_0 = diag(I_r \otimes (D_1 + D_2), \ldots, I_r \otimes (D_1 + D_2)).$$

Clearly Q' posses a QBD structure. The matrix $B = B_0 + B_1 + B_2$ is an infinitesimal generator and the steady state distribution of B is obtained below.

$$B = I_K \otimes diag(V_1, V_2, \ldots, V_{ar}),$$

$$V_i = e \otimes \alpha_i \otimes (S_3^0 \otimes I_a + S_K \otimes I_a + I_r \otimes D), 1 \leq i \leq ar.$$

Let $\pi = (\pi_1, \pi_2, \dots \pi_K)$ be the corresponding steady state probability vector,

$$\pi_i = (\pi_{i1}, \pi_{i2}, \dots, \pi_{iK}), i \geq 1.$$

π may be obtained as the solution to the system $\pi B = 0$ and $\pi e_{Kar} = 1$ where e_{Kar} is a column vector of ones having length Kar. B is a singular matrix as it can have at most $Kar - 1$ linearly independent rows. Hence $\pi B = 0$ has a non trivial solution π, this solution can be normalized to satisfy the condition $\pi e_{Kar} = 0$.

Theorem 1. *The system is stable if and only if*

$$\sum_{i=1}^{K} \pi_i I_r \otimes (D_1 + D_2) < \sum_{i=1}^{K} \pi_i (e \otimes \alpha_i \otimes (S_3^0 \otimes I_a)).$$

Proof. The system is stable if and only if $\pi B_2 e < \pi B_0 e$, see [15]. From matrices defined in Lemma 2

$$\pi B_2 e = \sum_{i=1}^{K} \pi_i I_r \otimes (D_1 + D_2),$$

$$\pi B_0 e = \sum_{i=1}^{K} \pi_i (e \otimes \alpha_i \otimes (S_3^0 \otimes I_a)).$$

□

The stationary distribution of the system process is obtained as follows. Under the assumption of the stability condition, the steady state probability distribution exists. Let $y = (y_0, y_1, y_2, \dots)$ be the steady state probability vector of the Markov Chain $\mathcal{Y}$, where $y_i = (y_i^0, y_i^1, \dots, y_i^M)$. Then y is the unique solution to the system of equations $yQ = 0$ and $ye = 1$. Then each component y_i ia a vector of length Kar.

From $yQ' = 0$ and $ye = 1$, we have the system of equations

$$\begin{gathered} y_0 A_{00} + y_1 B_{10} = 0 \\ y_0 A_{01} + y_1 B_{11} + y_2 B_{20} = 0 \\ y_1 B_{12} + y_2 B_{21} + y_3 B_2 = 0 \\ y_1 B_0 + y_3 B_1 + y_4 B_2 = 0 \\ \vdots \end{gathered}$$

Now from Matrix analytic method, $y_{c+i} = y_c R^i$, $i = 0, 1, 2 \dots$,where R is the minimal nonnegative solution of matrix quadratic equation $R^2 A_2 + R A_1 + A_0 = 0$. R is computed algorithmically, using the logarithmic reduction algorithm [12].

3 Performance Characteristics of the System

- Expected number of ordinary customers in the system.

$$E_{N_1} = \sum_{i=0}^{\infty} i y_i e.$$

- Expected number of emergency customers.

$$E_{N_2} = \sum_{i=0}^{\infty} \sum_{j=0}^{M} j y_i e.$$

- Expected rate of departure from the system.

$$E_r = \sum_{i=0}^{K-1} y_i A_i e + \sum_{l=0}^{\infty} \sum_{i=l+K}^{l+2K-1} y_i A i^{\#} e.$$

- Probability that server is idle with i customers $i < K$ customers in the queue.

$$P_{idle}^{i} = \sum_{i=0}^{K-1} y_{i0}^{0} e.$$

4 Expected Service Time of a Customer

The expected service time of a customer is described as follows. The service process can be considered as a continuous time Markov process on the finite state space

$$\{\#\} \cup \{(j,l,m), 1 \leq j \leq K, 1 \leq l \leq r, 1 \leq m \leq a\}$$

and $\{\#\}$ is the absorbing state.

$$\mathcal{Q}_s = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \Sigma & \Sigma^0 \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} A_{11} & A_{12} & & & \\ & A_{21} & A_{22} & & \\ & & \ddots & \ddots & \\ & & & A_{K-11} & A_{K-12} \\ & & & & A_{K1} \end{pmatrix}, \Sigma^0 = \begin{pmatrix} A_{10} & A_{20} & \dots & A_{K0} \end{pmatrix}^T.$$

The service time of an arbitrary customer follows phase type distribution with representation (β, Σ), where $\beta = (\beta_1, \beta_2, \dots, \beta_K, \beta_{K+1})$, $\beta_i = \frac{1}{(K+1)ar} \alpha_i \otimes e^T$ and $\beta_K = \frac{1}{ar} \sum_{i=1}^{K} \alpha_{ir+1}$.

Expected service time of an arbitrary customer is given by

$$E_{st} = -\beta \Sigma^{-1} e.$$

5 Expected Waiting Time of an Emergency Customer When at Least K Ordinary Customers are in Waiting

The expected waiting time of an emergency customer is described as follows. Consider the process $\mathcal{H} = \{(\tau(t), I(t), s(t)), t \geq 0\}$. Then $\mathcal{H}$ is clearly an irreducible continuous time stochastic process defined on a finite state space $\{(i,j,k), 1 \leq i \leq M, 1 \leq j \leq K, 1 \leq k \leq r\} \cup \{*\}$, where $*$ represents the absorption state and all other states are transient. The infinitesimal generator of this process is as shown below.

$$\mathcal{Q}_{\mathcal{H}} = \begin{pmatrix} \mathcal{C} & \mathcal{C}^0 \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

where

$$\mathcal{C} = \begin{pmatrix} C_M & & C^*_{M-K} & & & \\ & C_{M-1} & & C^*_{M-(K+1)} & & \\ & & \ddots & & & \\ & & & C_K & & C^*_0 \\ & & & & \ddots & \\ & & & & & C_1 \end{pmatrix},$$

$$C_h = diag(S_1, S_2, \ldots, S_K), 1 \leq h \leq M,$$

$$C^*_g = \begin{pmatrix} 0 \ldots & 0 & S^0_1 \otimes \alpha_K \\ \vdots \ddots & & S^0_2 \otimes \alpha_K \\ & \ddots & \vdots \\ & & S^0_{K-1} \otimes \alpha_K \\ 0 & 0 & 0 \end{pmatrix}, 0 \leq g \leq M-K,$$

$$\mathcal{C}^0 = (C^0_M, C^0_{M-1}, \ldots, C^0_1)^T,$$

$$C^0_i = (0, 0, \ldots, S^0_K \otimes \alpha_K), 0 \leq i \leq M.$$

The waiting time of an emergency customer follows a phase type distribution with representation $(\gamma, \mathcal{C})$, where $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_K)$ and $\gamma_i = \frac{1}{Kr}\alpha_i \otimes e^T$. Expected waiting time of an emergency customer is

$$E^w_e = -\gamma \mathcal{C}^{-1} e.$$

6 Numerical Example

We illustrate the model by considering a system with $K = 2$, the service times are exponentially distributed with parameters μ_1 and μ_2. We take arrival rate of ordinary customers as λ and arrival rate of emergency customers as γ. The matrices defining the MMAP process takes the form $D_0 = (-\lambda - \gamma), D_1 = (\lambda), D_2 = (\gamma)$.

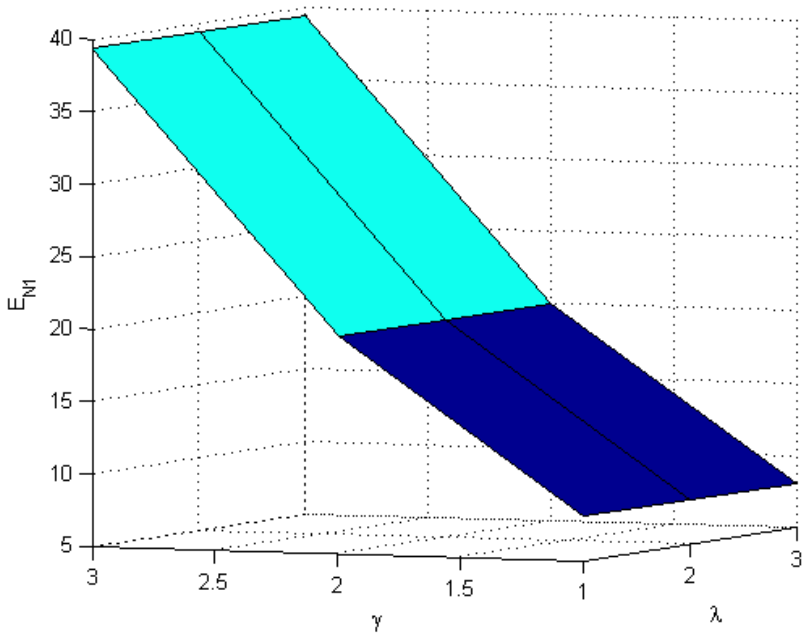

Fig. 1. Variation in queue length with respect arrival rates.

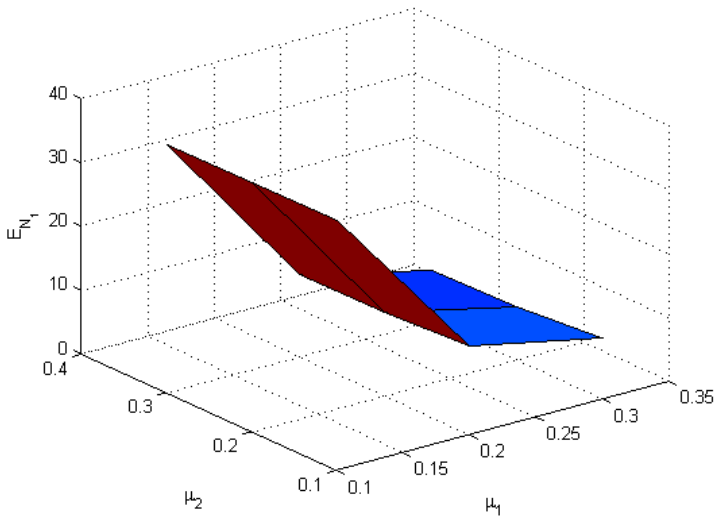

Fig. 2. Variation in number of customers in queue with variation in μ_1 ans μ_2.

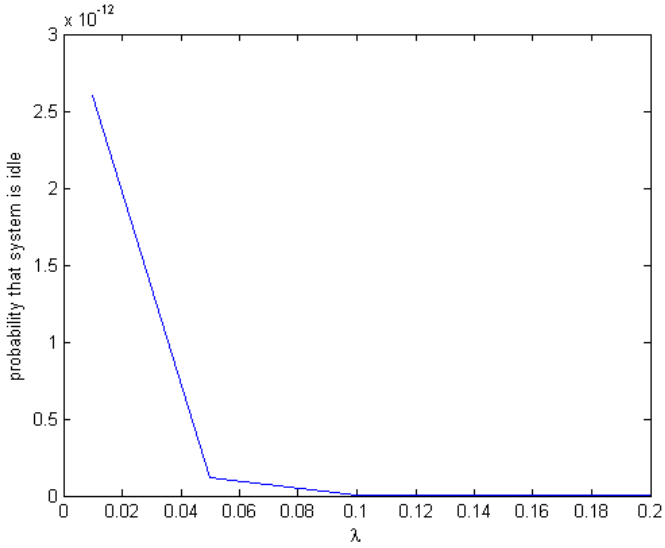

Fig. 3. Idleness with λ.

In Figure 1 variation in expected number of ordinary customers in the system, (E_{N_1}) with respect to the variations in arrival rates of ordinary and emergency customers is depicted. It is observed from Figure 1 that accumulation of ordinary customers increases in the queue as arrival rate of emergency customers increases. Figure 2 describes the variation in (E_{N_1}) with respect to the service rates μ_1 and μ_2. Figure 3 shows that the probability that system is in idle state decreases with the increase of rate arrival of the ordinary customers.

7 Conclusion

In this paper we have studied a queue with batch service and batch size depends the number of customers in the buffer. The model is observed in a courier delivering system with two kinds of arrivals. The strategy for batch size determination is designed for reducing system cost. The model is analyzed using matrix analytic method and is illustrated with a numerical example.

8 Acknowledgment

Sinu Lal T. S. thanks University Grants Commission(UGC) of India for UGC-Junior Research Fellowship(Roll no.-432693,June 2017).

References

1. Banerjee, A., Gupta, U. C., Sikdar, K.: Analysis of finite buffer bulk-arrival bulk-service queue with variable service capacity and batch size-dependent service, International Journal of Mathematics in Operational Research, **5**, (3), 358-386 (2013).
2. Baetens, J., Steyaert, B., Claeys, D., Bruneel, H.: System occupancy of a two class batch service queue with class dependent variable server capacity, International Conference on Analytical and Stochastic Modeling Techniques and Applications, 32-44, Springer (2016).
3. Baetens, J., Steyaert, B., Claeys, D., Bruneel, H.: Delay analysis of a two class batch service queue with class dependent variable server capacity. Mathematical Methods of Operations Research,**88**(1), 37-57 (2018).
4. Baetens, J., Steyaert, B., Claeys, D., Bruneel, H.: System occupancy in a multi-class batch-service queueing system with limited variable service capacity. Annals of Operations Research, 1-24 (2019).
5. Chaudhry, M. L., Gupta, U. C.: Modelling and analysis of $M/G^{a,b}/1/n$ queue a simple alternative approach, Queueing Systems, **31**(1-2), 95-100 (1999).
6. Gold, H., Tran-Gia, P.: Performance analysis of a batch service queue arising out of manufacturing system modelling. Queueing Systems, **14**(3-4), 413-426 (1993).
7. Gupta, G., Banerjee, A.: On $M/G^{(a;b)}/1/n$ queue with batch size and queue length dependent service. International Conference on Mathematics and Computing, 249-262, Springer (2018).
8. Gupta, U. C., Sikdar, K.: The fnite-buffer m/g/1 queue with general bulk-service rule and single vacation,Performance Evaluation, **57**(2), 199-219 (2004).
9. He, Q-M.: Fundamentals of matrix-analytic methods, Vol. 365, Springer, New York (2014).
10. Kuppa, S., Dattatreya, G.: Modeling and analysis of frame aggregation in unsaturated wlans with finite buffer stations, IEEE International Conference on Communications, **3**, 967-972, IEEE (2006).
11. Maity, A., Gupta, U. C.: Analysis and optimal control of a queue with infinite buffer under batch size dependent versatile bulk-service rule. Oper. research, **52**(3), 472-489 (2015).
12. Latouche, G., Ramaswami, V.: Introduction to matrix analytic methods in stochastic modeling, Vol. 5, Siam (1999).
13. Neuts, M. F.: A general class of bulk queues with Poisson input. The Annals of Mathematical Statistics, **38**(3), 759–770, JSTOR (1967).
14. Neuts, M. F.: Probability distributions of phase type. Liber Amicorum Prof. Emeritus H. Florin (1975).
15. Neuts, M. F.: Matrix-geometric solutions in stochastic models: an algorithmic approach. Courier Corporation (1994).
16. Powell, W. B.: Analysis of vehicle holding and cancellation strategies in bulk arrival, bulk service queues. Transportation Science. **19**(4), 352–377, INFORMS (1985).

17. Reddy, G. K., Nadarajan, R., Kandasamy, P.: A nonpreemptive priority multiserver queueing system with general bulk service and heterogeneous arrivals. Computers and operations research, **20**(4), 447-453 (1993).
18. Sasikala, S., Indhira, K.: Bulk service queueing models survey. International Journal of Pure and Applied Mathematics **106**(6), 43-56 (2016).
19. Sikdar, K., Samanta, S.: Analysis of a finite buffer variable batch service queue with batch markovian arrival process and servers vacation. Oper. research, **53**(3), 553-583 (2016).
20. Sinu Lal, T. S., Krishnamoorthy, A., Joshua, V. C.: A Multiserver Tandem Queue with a Specialist Server Operating with a Vacation Strategy. Automation and Remote Control, **81**, 760-773 (2020).
21. Sinu Lal, T. S., Krishnamoorthy, A., Joshua, V. C. September. A Queueing Inventory System with Search and Match-An Organ Transplantation Model. International Conference on Distributed Computer and Communication Networks, 273-287, Springer, Cham (2019).
22. Van De Rzee, D. J., Van Harten, A., Schuur, P. C.: Dynamic job assignment heuristics for multi-server batch operations-a cost based approach. International Journal of Production Research, **35**(11), 3063-3094 (1997).
23. Yu, M., Alfa, A. S.: Algorithm for computing the queue length distribution at various time epochs in $DMAP/G^{(1,a,b)}/1/n$ queue with batch size dependent service time. European Journal of Operational Research, **244**(1), 227-239 (2015).