

On a Retrial Queueing Model with Customer Induced Interruption

Varghese Jacob¹ *∗*

¹*Department of Mathematics, Government College, Kottayam-686013, India.*

Author's contribution

The sole author designed, analyzed, interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/JAMCS/2020/v35i230253 *Editor(s):* (1) Dr. Metin Basarir, Sakarya University, Turkey. *Reviewers:* (1) Jackson Akpojaro, University of Africa Toru-Orua, Nigeria. (2) Myron Hlynka, University of Windsor, Canada. Complete Peer review History: http://www.sdiarticle4.com/review-history/56995

> *[Received: 01 March 2020](http://www.sdiarticle4.com/review-history/56995) Accepted: 06 May 2020*

Original Research Article Published: 09 May 2020

Abstract

This paper presents a retrial queueing system with customer induced interruption while in service. We consider a single server queueing system of infinite capacity to which customers arrive according to a Poisson process and the service time follows an exponential distribution. An arriving customer to an idle server obtains service immediately and customers who find server busy go directly to the orbit from where he retry for service. The inter-retrial time follows exponential distribution. The customer interruption while in service occurs according to a Poisson process and the interruption duration follows an exponential distribution. The customer whose service is got interrupted will enter into a finite buffer. Any interrupted customer, finding the buffer full, is considered lost. Those interrupted customers who complete their interruptions will be placed into another buffer of same size. The interrupted customers waiting for service are given non-preemptive priority over new customers. We analyse the steady-state behavior of this queuing system. Several performance measures are obtained. Numerical illustrations of the system behaviour are also provided with example.

Keywords: Quasi-birth-and-death process; Matrix analytic methods; Customer induced interruption; Retrial queues; Steady-state analysis.

2010 Mathematics Subject Classification: 68K25.

^{}Corresponding author: E-mail: drvarghesejacob@gmail.com;*

1 Introduction

Queueing models with repeated attempts, known as 'retrial queues', have been widely used to model many problems in telecommunication and computer systems. The important feature of a retrial queue is that arriving customers who find all the server busy have to leave the service area and join a retrial group, called orbit, in order to try their luck again after some random time. For a detailed review of the literature on this topic the reader is referred to the a survey paper by Artalejo et al. [1].

In the classical queueing systems, servers are always available to serve customers. But, there are situations where the servers may unavailable for a random period of time due to many reasons such as server interruptions, server vacations, removal of servers due to catastrophic or negative arrival [of](#page-8-0) events, getting preempted due to the arrival of priority customers etc. A review of recent work on different type of server interruption, readers are referred to Krishnamoorthy *et al.* [2]. So far in literature only a few papers studied customer induced interruptions such as customers leaving in the middle of a service.

In many daily life situation where the customers are often leave the service area in the middle of a service due to not having enough information for completing a service. However, the[se](#page-8-1) customers are request their service after some random period of time. These type of interruption is known as *customer induced interruption*. A more commonly occurring example is the following :- In a doctors clinic, while patient is being examined, the physician may find that one or more tests needed for prescription of medicine. Hence he/she is asked to undergo these and return to the clinic. Such patients can be regarded as interruption induced by the customer.

As far our knowledge goes, the first work dealing queues with customer induced interruptions is [3] reported at 8*th* International Workshop on retrial queues in 2010 and the paper of Jacob *et al.* [4]. Subsequently, Krishnamoorthy and Jacob [5] extended the work to a multi-server *M/M/c* model. Jacob and Krishnamoorthy [6] discussed $M/PH/1$ queueing system with customer induced interruption in the retrial set up with a finite orbit. Recentlty, Punalal and Babu [7] studied a [re](#page-8-2)trial queueing model with self-generation of priorities and customer induced interruption.

The [p](#page-8-3)urpose of this work is to introduce custom[er](#page-8-4) induced in a retrial queueing systems with classical retrial policy. In classical r[etr](#page-8-5)ial policy, the rate of retrial of customers for service depends on the number of customers in the orbit. Most of the application problems, retrial of [a](#page-8-6) particular customer need not depends on the retrial the of other. So we consider the constant retrial policy in this model.

The paper is organized as follows. In Section 2, the model under study is described. Section 3 provides the steady-state analysis of the model. Section 4 discuss the main performance measures of the system. Some illustrative examples are discussed in section 5.

2 Model Description

We consider an infinite capacity queueing system with a single sever to which customers arrive according to a Poisson process with rate λ . The service times are assumed to follow an exponential distribution with parameter μ . An arriving customer to the idle server obtains service immediately. Customers who find server busy go directly to the orbit from where he retry for service. In classical retrial policy, the rate of retrial of customers for service depends on the number of customers in the orbit. Most of the real situation, retrial of a particular customer need not depends on the retrial the of other. So we consider the constant retrial policy in this model and the inter-retrial time is exponentially distributed with rate σ . Here a customer induced interruption occurs while in service according to a Poisson process of rate θ . When an interruption occurs, the currently in service will be forced to leave the service facility. The freed server is ready to offer services to other customers. The interrupted customer will enter into a buffer (referred to as *BIP*) of finite capacity, *K*, should there be a space available. Otherwise, the customer will be lost from the system. The interrupted customers will spend a random period of time that is independent of other customers and the interruption time follows an exponential distribution with parameter *η*. Also it is assumed that the maximum number of interruptions allowed for a customer is one. That is, an an interrupted customer cannot be interrupted again and hence will leave the system after getting a service. All interrupted customers upon completing their interruption enter into a finite buffer (referred to as *BIC*) whose size is *K*. Customers who are in *BIC* are given non-preemptive priority over new customers but are served in the order in which they enter into this buffer. Thus, a free server will offer services to those customers waiting in *BIC* before serving new customers by maintaining the first-in-first-served order. Because of this restriction coupled with the fact that at most one interruption is allowed for any customer, the total number of customers in *BIC* will never exceed the size of *BIP* and hence we assume the buffer sizes to be the same.

In the sequel we use the following notations.

- n_t = Number of customers in the orbit at time *t*.
- j_t = Number of customers in *BIC* at time *t*.
- m_t = Number of customers in *BIP* at time *t*.
	- $\sqrt{ }$ 0 if server is idle
- \bullet $i_t =$ $\frac{1}{2}$ 1 if server is busy with primary/orbital customer
	- \mathbf{I} 2 if server is busy with a customer from BIC
- *•* $a_i = (1, 2, \ldots, i); 1 \leq i \leq K$
- *•* **e** denote column vector of 1's with appropriate dimension,
- $e_j(r)$ denote column vector of dimension *r* with 1 in the j^{th} position and 0 elsewhere
- *• I^r* denote identity matrix of dimension *r*.
- *•* ∆(*aⁱ*) is a diagonal matrix whose diagonal entries are the components of the vector *ai*.
- *• A ⊗ B* denotes the Kronecker product of matrices *A* and *B*.
- $M = (K+1)(K+2)/2$.
- $M1 = (K+1)(K+3)$.

The process $\mathbf{X} = \{(n_t, i_t, j_t, m_t) : t \geq 0\}$ is a continuous-time Markov chain (*CTMC*) state space

$$
\Omega = \bigcup_{n=0}^{\infty} (l^*(n) \cup l(n)),
$$

where $l^*(n) = \{ (n, 0, 0, m) : m = 0, 1, ..., K \}, l(n) = \{ (n, i, j, m) : i = 1, 2; j, m = 0, ..., K; 0 \leq$ $j + m \leq K$, $n \geq 0$.

The constant retrial rate makes the *CTMC* under consideration is a level independent QBD (*LIQBD*) with infinitesimal generator matrix *Q*.

$$
Q = \begin{bmatrix} Q_1 & Q_0 & & & \\ Q_2 & Q_1 & Q_0 & & \\ & Q_2 & Q_1 & Q_0 & \\ & & Q_2 & Q_1 & Q_0 & \\ & & & \ddots & \ddots & \ddots \end{bmatrix},
$$
(2.1)

114

where Q_0, Q_1, Q_2 are square matrices of dimension $(K+1)(K+3)$.

To write down the generator matrix *Q*, we consider the transitions from one level to another :

• The entries of *Q*²

Transitions due to successful retrial when server is idle.

 $(n, 0, 0, m) \rightarrow (n-1, 0, 0, m):$ σ for $m = 0, 1, ..., K; n \ge 0$

• The entries of *Q*⁰

Transitions due to arrival of primary customers when server is in busy state.

 $(n, i, j, m) \rightarrow (n + 1, i, j, m)$: λ for $i = 1, 2; i = 0, 1, \dots, K; m = 0, 1, \dots, K - j; n \ge 0$ *•* The entries of *Q*¹

Transitions that the levels are unchanged when the system is in busy state.

- (i). When all the states are unchanged.
	- ρ (*n*, 0*,* 0*, m*) → (*n,* 0*,* 0*, m*) : $-(\lambda + \sigma + m\eta)$
	- \triangleright $(n, i, j, m) \rightarrow (n, i, j, m) :$ $-(\lambda + \mu + \theta + m\eta)$
	- \triangleright $(n, i, j, m) \rightarrow (n, i, j, m) :$ $-(\lambda + \mu + m\eta)$

for
$$
i = 1, 2, j = 0, 1, ..., K
$$
; $m = 1, ..., K - j$; $n \ge 0$.

(ii). Transitions due to customer interruption.

- ρ (*n*, 1*,* 0*, m*) \rightarrow (*n*, 0*,* 0*, m*): θ for $m = 1, ..., K; n > 0$.
- $p \in (n, 1, j, m) \rightarrow (n, 2, j-1, m+1):$ θ for $j = 1, 2, ..., K;$ $m = 0, 1, ..., K j;$ $n \ge 0.$

(iii). Transitions due to customer interruption completion.

$$
\triangleright (n, 0, 0, m) \to (n, 2, 0, m - 1): \quad m\eta \quad \text{for } m = 1, ..., K; n \ge 0.
$$

$$
\triangleright (n, i, j, m) \to (n, i, j + 1, m - 1): \quad m\eta \quad \text{for } i = 1, 2; j = 0, 1, ..., K - 1;
$$

$$
m = 1, ..., K - j; n \ge 0.
$$

- (iv). Transitions due to service completion.
	- φ $(n, i, 0, m) \rightarrow (n, 0, 0, m)$: μ for $i = 1, 2; m = 1, 2, \ldots, K; n \ge 0$. $p \in (n, 1, j, m) \rightarrow (n, 2, j - 1, m):$ μ for $j = 1, 2, \dots, K; m = 0, 1, \dots, K - j; n \ge 0.$ $p \in (n, 2, j, m) \rightarrow (n, 2, j - 1, m):$ μ for $j = 1, 2, \dots, K; m = 0, 1, \dots, K - j; n \ge 0.$

(v). Transitions due to arrival of primary customer to system when idle.

 ρ (*n,* 0*,* 0*, m*) \rightarrow (*n,* 1*,* 0*, m*): *m* for *m* = 0*,* 1*, ..., K; n* \geq 0*.*

3 Steady-state Analysis

In this section we perform the steady-state analysis of the queueing model under study.

3.1 Stability condition

Let π denote the steady-state probability vector of the generator $Q_0 + Q_1 + Q_2$. That is, $\pi(Q_0 + Q_1 + Q_2) = 0$, $\pi e = 1$. The *LIQBD* description of the truncated system is stable (see Neuts [8]) if and only if

$$
\pi Q_0 e < \pi Q_2 e. \tag{2.2}
$$

The vector, π , cannot be obtained explicitly in terms of the parameters of the model, and hence the stability condition is known only implicitly. For future reference, we define the traffic intensity, ρ as

$$
\rho = \frac{\pi Q_0 e}{\pi Q_2 e}.\tag{2.3}
$$

3.2 Steady-state vector

Suppose x denote the steady-state probability vector of the generator Q given in (2.1) . That is,

$$
x Q = 0, \quad x e = 1. \tag{2.4}
$$

When the stability condition holds, we see that there exists a unique steady-state probability vector x. We define the the steady-state distribution of $\{(n_t = n, i_t = i, j_t = j, m_t = m) : t \geq 0\}$ as follows :

$$
\mathbf{x}_{i,j,m}(n) = \lim_{t \to \infty} P(n_t = n, i_t = i, j_t = j, m_t = m); \ \ (n, i, j, m) \in \Omega
$$

For the computation of stationary probabilities $x_{i,j,m}(n)$, we adopt the matrix-geometric method (see [8]).

Now Partitioning *x* as

$$
\mathbf{x} = (\mathbf{x}(0), \mathbf{x}(1), \dots, \dots) \tag{2.5}
$$

Fro[m e](#page-8-7)quation (2.4) we can obtain

$$
x(0)Q_1 + x(1)Q_2 = \mathbf{0}
$$

$$
x(i-1)Q_0 + x(i)Q_1 + x(i+1)Q_2 = \mathbf{0}, i \ge 1
$$

We see that *x*, under the assumption that the stability condition (2.2) holds, is obtained as (see Neuts [8])

$$
\mathbf{x}(n) = \mathbf{x}(0)R^n, \quad n \ge 1,\tag{2.6}
$$

where R is the minimal non-negative solution to the matrix quadratic equation:

$$
R^2 Q_2 + R Q_1 + Q_0 = 0,\t\t(2.7)
$$

(see Neuts [8]). Under normalizing condition

$$
x(0)(I - R)^{-1} e = 1.
$$
\n(2.8)

Then using equations (2.6) and (2.8), we find $x(i)$, $i \ge 0$.

Once the rate matrix R is obtained, the vector x can be computed using logarithmic reduction algorithm. For full details on the logarithmic reduction algorithm we refer the reader to [9].

4 System Performance Measures

In this section we present the main system performance measures to bring out the qualitative aspects of the model under study. These are listed below along with their formula for computation. We further partition the vectors, $\mathbf{x}(n)$, $n \geq 0$, into $\mathbf{x}(n) = (\mathbf{x}^*(n), \mathbf{x}_1(n), \mathbf{x}_2(n))$, $n \geq 0$, where

$$
\mathbf{x}^*(n) = (x_0^*(n), \cdots, x_K^*(n))
$$

$$
\mathbf{x}_1(n) = (\mathbf{x}_{1,0}(n), \mathbf{x}_{1,1}(n), \cdots, \mathbf{x}_{1,K}(n)), n \ge 0,
$$

$$
\mathbf{x}_2(n) = (\mathbf{x}_{2,0}(n), \mathbf{x}_{2,1}(n), \cdots, \mathbf{x}_{2,K}(n)), n \ge 0.
$$

Note that $x_{j,r}(n)$, *j* = 1, 2, 0 ≤ *r* ≤ *K*, *n* ≥ 0, is of dimension *K* + 1 − *r*.

- The probability that the server is idle : $P_{idle} = \sum_{i=1}^{\infty}$ $\sum_{n=0}^{\infty}$ *x*^{*}(*n*)*e***.**
- *•* The probability that the server is busy with a primary/orbital customer : $P_{bsyo} = \sum^{\infty}$ $\sum_{n=0}$ $x_1(n)e = x(0)(I - R)^{-1}(e_2(3) \otimes e).$
- *•* The probability that the server is busy with an interrupted customer: $P_{bsys} = \sum_{i=1}^{\infty}$ $\sum_{n=0}$ $x_2(n)e = x(0)(I - R)^{-1}(e_3(3) \otimes e).$
- *•* The probability that an interrupted customer is lost: $P_{loss} = \frac{\theta}{\theta + \mu} \sum_{i=1}^{\infty}$ $\sum_{n=0}$ $x_{1,0,K}(n)$.
- *•* The expected number of customers in the orbit : $E_{orbit} = x(0)R(I - R)^{-2}e$.
- *•* The expected number of interrupted customers in the *BIC* buffer : $E_{BIC} = \sum_{i=1}^{\infty}$ *n*=0 $\sum_{ }^{ }$ *i*=1 ∑*K j*=0 *^K*∑*[−]^j* $\sum_{m=0}$ *j* $x_{i,j,m}(n)$ *.*
- *•* The expected number of interrupted customers in the *BIP* buffer : $E_{BIP} = \sum_{i=1}^{\infty}$ *n*=0 $\sum_{i=1}^{2}$ *i*=1 ∑*K j*=0 *^K*∑*[−]^j* $\sum_{m=0}^{\infty} m x_{i,j,m}(n)$.
- *•* The rate at which the orbiting customer successfully reach the server is given by $\sigma_1^* = \sigma \, \stackrel{\infty}{\sum}$ $\sum_{n=0}^{\infty}$ *x*^{*}(*n*)*e***.**
- *•* The overall rate of retrials at which the orbiting customers request service is given by $\sigma_2^* = \sigma \mu_{orbit}$.
- The fraction, *FSR*, of successful rate of retrials is given by $FSR = \frac{\sigma_1^*}{\sigma_2^*}$.

5 Numerical Illustrations

In this section we present some numerical examples to show the effect of various parameters of the system when other parameters are fixed. The correctness and accuracy of the code are verified by a number of accuracy check.

Example 1 : Here we fix the parameters $(K, \lambda, \mu, \eta) = (5, 6, 8, 5)$.

Looking to the Fig. 1.(a) and Fig. 1.(b), we can see the influence of the parameter θ on the measures E_{BIC} and P_{bsys} for various values of the retrial rate σ . As θ increases both the measures are also increases. This is because increase of interruption rate results in more customers are getting interrupted and moves to *BIP* buffer results in the increase of customers in *BIC* (note that $\eta =$

5). Also, due to the non-preemption priority policy of interrupted customer over primary/orbital customers, the server is getting busy with these interrupted customers. So P_{bsws} increases as θ increases. Again, for a fixed θ , from the Figure 1(a), we observe that E_{BIC} is a non-decreasing function of σ .

Fig. 1. Fixing the parameters $(K, \lambda, \mu, \eta) = (5, 6, 8, 5)$

Fig. 2. Fixing the parameters $(K, \lambda, \theta, \eta) = (5, 1, 10, 8)$

Example 2 : In this example we fix the parameters $(K, \lambda, \theta, \eta) = (5, 1, 10, 8)$

In Fig. 2, we observed that the measures *Eorbit* and *Pidle* are a non-increasing function of *σ* and for every values of μ .

The reason is that, as retrial rate increases, more orbital customers can get into service and so the idle probability of the server getting reduced. From Fig. 2.(b), for a fixed σ , E_{orbit} is a nonincreasing function of μ . This is due to the fast clearance of customers from the system.

Also from Fig. 2.(b), for a fixed σ , P_{idle} is non-decreasing function of μ . The reason is that for higher value of μ , the service completion will be at a faster rate results in the system can stay in an idle state if there is no successful retrial.

Example 3 : Here we fix the parameters $(K, \lambda, \theta, \eta) = (5, 1, 10, 8)$.

Looking at the Table 1, we summarize the following.

- The measure ρ , is a non-increasing function of σ for every values of μ . This is due to the fact that an increase in σ will cause more customers leaving the system after service completion, so the traffic intensity getting reduced.
- The measures P_{bsys} , E_{BIC} , σ_1^* and σ_2^* are increases as σ increases whereas the measures E_{BIF} is decreases when σ increases for every values of μ . P_{bsys} and E_{BIC} increases due to the fact that (note that $\theta = 10$, $\eta = 8$) so the server being busy with interrupted customers due to our non- priority assumption of *BIC* customers. Again, *σ ∗* ¹ increases as more customers retry for service from the orbit results in a successful retrial due to non-preemption assumption.
- Again, from table we can see that, for a higher value of σ , all the measures except E_{BIC} are decreases when μ increases. The rate of increase/decrease is low for higher values of σ . This is because for smaller values of σ , the number of customer leaving the system after getting service increases ($\eta = 8$). For smaller values of σ , as rate of service increases, the number of customers in the *BIP* buffer completes their interruption and moved to *BIC* buffer. So the probability that the server is busy with *BIC* customers decreases.

σ	ρ	P_{bsys}	E_{BIC}	E_{BIP}	σ_1^*	σ_2^*
$\mu = 6$						
3	0.55555	0.54952	0.93059	0.06718	1.16395	2.04988
5	0.40000	0.65601	1.16035	0.06144	1.40743	3.12524
7	0.33333	0.71642	1.29420	0.05693	1.54754	4.18179
9	0.29628	0.75538	1.38253	0.05330	1.63906	5.22841
10	0.28332	0.77010	1.41644	0.05174	1.67395	5.74935
11	0.27271	0.78261	1.44551	0.05031	1.70375	6.26914
14	0.24996	0.81094	1.51239	0.04669	1.77185	7.82351
$\mu = 7$						
3	0.52381	0.53045	0.92605	0.06232	1.23219	1.75508
5	0.37143	0.64021	1.15968	0.05656	1.50484	2.67914
7	0.30612	0.70348	1.29716	0.05204	1.66390	3.58385
9	0.26983	0.74467	1.38829	0.04841	1.76857	4.47863
10	0.25713	0.76031	1.42332	0.04685	1.80862	4.92369
11	0.24674	0.77364	1.45338	0.04543	1.84288	5.36763
14	0.22446	0.80394	1.52252	0.04184	1.92134	6.69462
$\mu = 8$						
3	0.50000	0.51265	0.91459	0.05815	1.29540	1.53973
5	0.35000	0.62507	1.15149	0.05244	1.59685	2.35191
7	0.28571	0.69087	1.29242	0.04797	1.77505	3.14460
9	0.24999	0.73409	1.38632	0.04439	1.89317	3.92735
10	0.23749	0.75059	1.42250	0.04285	1.93853	4.31640
11	0.22726	0.76468	1.45357	0.04145	1.97740	4.70435
14	0.20533	0.79683	1.52513	0.03793	2.06663	5.86345

Table 1. Fixing the parameters $(K, \lambda, \theta, \eta) = (5, 1, 10, 8)$

6 Conclusion

In this paper, our objective is to analyse an infinite *M/M/*1 retrial queueing model with constant retrial policy and customer induced interruption while in service using matrix geometric method. All underlying distributions are assumed to be exponential that are independent of each other.

There is a finite buffer for self interrupted customers to wait for completion of interruption and another buffer of the same capacity for those who have completed their interruption. The system steady-state (long run behaviour) are analyzed by using matrix analytic method. We derived several performance measures of the system under study. The effect of various parameters on the system performance are also investigated with the help of numerical examples.

Acknowledgements

The Author is very thankful to two anonymous reviewers and an editor, for the valuable comments and constructive criticism which resulted in improvement in the presentation of the paper.

Competing Interests

Author has declared that no competing interests exist.

References

- [1] Artalejo JR. Accessible bibliography on retrial queues: Progress in 2000-2009. Mathematical and Computer Modelling. 2010;51:1071-1081.
- [2] Krishnamoorthy A, Pramod PK, Chakravarthy SR. A survey on queues with interruptions. TOP-Spanish journal of Statistics & Operations Research. 2014;22(1):290-320.
- [3] Jacob V, Chakravarthy SR, Krishnamoorthy A. On a customer induced interruption in a service system. Abstract of International Workshop on Retrial Queues (8*th* WRQ). Beijing, China; 2010.
- [4] Jacob V, Chakravarthy SR, Krishnamoorthy A. On a customer induced interruption in a service system. Stochastic Analysis and Applications. 2012;30:949-962.
- [5] Krishnamoorthy A, Jacob V. Analysis of customer induced interruption in a multi server systeml. Neural, Paralle and Scientific Computations. 2012;20:153-172.
- [6] Jacob V, Krishnamoorthy A. Analysis of customer induced interruption and retrial of interrupted customers. American Journal of Mathematical and Management Sciences. 2015;34(4):343-366.
- [7] Punalal J, Babu S. A single server queueing model with self-generation of priorities, customer induced interruption and retrial of customers. International Journal of Research in Advent Technology. 2019;7(3):1299-1303.
- [8] Neuts MF. Matrix-Geometric solutions in stochastic models - an algorithmic approach. 2nd ed., Dover Publications Inc., New York; 1994.
- [9] Latouche G, Ramaswami V. A logarithmic reduction algorithm for quasi-birth-and-death processes. Journal of Applied Probability. 1993;30:650-674.

 $\mathcal{L}=\{1,2,3,4\}$, we can consider the constant of the con *⃝*c *2020 Jacob; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.*

Peer-review history:

The peer review hist[ory for this paper can be accessed here \(Pleas](http://creativecommons.org/licenses/by/2.0)e copy paste the total link in your browser address bar)

http://www.sdiarticle4.com/review-history/56995