Article

Retrial Queue with a Finite Number of Orbits for Failed Attempts to Access the Server and Search of Customers from Last Orbit

Ambily P. Mathew , Varghese C. Joshua , Achyutha Krishnamoorthy , Agassi Melikov [*] , Janos Sztrik [*]

Article

Retrial Queue with a Finite Number of Orbits for Failed Attempts to Access the Server and Search of Customers from Last Orbit

Ambily P. Mathew [1], **Varghese C. Joshua** [1], **Achyutha Krishnamoorthy** [2], **Agassi Melikov** [3] **and Janos Sztrik** [4,*]

1   Department of Mathematics, CMS College Kottayam, Kottayam 686001, India
2   Centre for Research in Mathematics, CMS College Kottayam, Kottayam 686001, India
3   Baku Engineering University, Baku AZ 0101, Khirdalan, Hasan Aliyev str. 120, Azerbaijan
4   Department of Informatics Systems and Networks, Faculty of Informatics, University of Debrecen, 4032 Debrecen, Hungary
*   Correspondence: sztrik.janos@inf.unideb.hu

Abstract: Customers arrive at a single server queue with a finite sequence M of orbits for retrial customers based on the number of retrials made by each customer. Of these M orbits, the first $M-1$ are of finite capacities and the last one (marked M) is of infinite capacity. Primary arrivals who are unable to access the server upon arrival, go to an orbit of finite capacity. From this orbit customers retry; those who fail to access the server join a second orbit of size larger than the first one and retry to access the server. Finally, they join the Mth orbit (after $M-1$ repeated trials). Retrials from all the orbits are assumed to be linear. Retrials occur from the Mth orbit and if those were unsuccessful, each customer will come back to the Mth orbit itself with probability q whereas with probability $1-q$, leaves the system. Also after each service completion epoch, the server goes in search of customers from the Mth orbit with probability p. With probability $1-p$ the server remains idle. Steady-state analysis has been done. Various measures of performance are calculated. We analyze this retrial queuing model graphically as well as numerically. We also analyze the model for the optimum value of search probability. Some probability distributions of interest related to the model, have been derived.

Keywords: retrial queue; multi-orbits;transfer of orbital customers to service station; abandonment of last orbit; search

MSC: 60J28; 60K25; 90B05; 90B22

---

1. Introduction

The Retrial Queueing (RTQ) model under consideration is a single server model with a finite number of orbits arranged in hierarchical order, labeled as $1, 2, \ldots, M$, and arranged in that order. Their capacities are respectively $K, 2K, \ldots (M-1)K$ and unlimited for the orbit-M. Further, the retrial rate of customers from these orbits has individual rates $\theta, 2\theta, \ldots, M\theta$. These two facts together almost ensure minimum loss of customers, consequent to failed retrials, for a reason to be indicated in the descriptions to follow. If the server is busy at the time of arrival of a primary customer he/she will join the first orbit. Retrials are made from the first orbit and if those are unsuccessful, customers move to the next orbit and so on until they reach the last orbit. All the orbits except orbit-M are of finite capacity. The last orbit, orbit-M has infinite capacity and the customers in that orbit may either return to the same orbit or leaves the system forever, after each unsuccessful retrial attempt. Retrials from all the orbits follow linear retrial policy. Thus the model under consideration takes into account the impatience of customers in such queueing systems. On completion of service, the server transfers a customer from orbit M with probability p for the next service or stays idle with probability $1-p$. The time taken for transfer is negligible. Thus, customers in that orbit (who have already completed at least $M-1$ retrials) will get higher priority over all customers from other orbits and also the primary arrivals.

In the mathematical modeling of real-life problems, the theory of retrial queues plays a pivotal role. The monograph by Falin and Templeton [11], gives an introduction to the theory of retrial queues.

Over the years, the theory of retrial queues has been developing and the literature covers various fields in which the theory can practically be applicable. The book by Artalejo and Gomez-Corral [4] includes comprehensive coverage of techniques for the computational analysis of retrial queues. It includes motivating examples in telephone and computer networks and a comparative analysis of the retrial queues versus standard queues with waiting lines and queues with losses. A modern treatment for the study of retrial queues can be seen in the books by Dudin et al. [9] and Chakravarthy [5,6]. Detailed and classified bibliography of articles on retrial queues can be found in [1,2]. The survey on retrial queues in [13,20] includes literature with algorithmic and computational aspects of retrial queues.

When we analyze most of the retrial queueing models, we can find that the server remains idle even if there are customers in the corresponding orbit/orbits. This happens just because of the fact that the retrying customers are unaware of the status of the server. Retrying customers will succeed in getting service only if the server is idle at the retrying epoch. To overcome the difficulties posed by such situations and to improve the utilization of the server the concept of orbital search in negligible time to the server at the end of a service completion epoch was introduced. In the classical queueing models the concept of searching of customers was introduced by Neuts and Ramalhoto [16]. It was extended to the case of retrial queues in 2002 by Artalejo et al. in [3]. Krishnamoorthy et.al.[14] extended the search mechanism to a queueing model in which the blocked customers leave the system forever. Orbital search in a wider sense is used in [7,8,12,15]. In these papers, the authors assume different distributions for the search time of customers from the orbit. However, in this, there is no guarantee that the customer so taken will be given the next service. This is because, during the search time, a primary or a retrial customer can access the server thereby failing the very objective of effectively reducing the idle time of the server and the waiting time of orbital customers.

The expanding(increasing) sizes of successive orbits compared to the previous one are to minimize customer loss due to retrials. Abandonment of the system due to failed retrials together with the additional feature of limited (finite) capacities and consequent loss of retrial customers from orbits 1, 2, $\ldots$, $M-2$ to orbits 2, 3, $\ldots$, $M-1$ lead to a more stable system. The increasing rate of retrials from successive orbits coupled with increasing rate of service of customers belonging to successive orbits and transfer of customers from orbit-M for service at a service completion epoch, ensure " almost sure" completion of service of customers reaching the final orbit within a specified time from the epoch of arrival of such customers.

A salient feature of this model is that we get a complete picture of customers in each of the orbits, namely the number of retrials each customer has made so far. 1, 2, $\ldots$, $M-2$ to orbits 2, 3, $\ldots$, $M-1$. This feature is not seen in any RTQ models discussed in the literature so far with the exception that multiple hierarchical orbits are considered in [19].

Highlights

- Information about the number of retrials made by customers up to $(M-1)$th orbit.
- Customers from orbit- M are picked for service by the server immediately on completion of service, with a positive probability. If this probability is taken as one, then it represents the case of a queue of customers who have completed $M-1$ retrials
- This model considers separate orbits for customers who failed to access the server as a primary customer (orbit-1), completed one retrial (orbit-2, ..., completed $M-2$ retrials (orbit-$(M-1)$ and those who have completed $M-1$ or more retrials (orbit-M). Because of finite capacities of orbits$1, 2, ..., M-1$, a few primary customers, a few customers who retried once, ..., retried $M-2$ times might have been lost. However, this is not the case with the Mth orbit. It has infinite capacity. The actual loss of customers from this orbit is due to failed retrial $((M-1)^{th}$or more retrials). We could have modified the last orbit customers' status as those having completed $M-1$ retrials by assuming that customers who fail to access the server from this orbit NOT returning to orbit with probability 1. This is essentially a particular case of the model that we discuss in this paper.

- The increase in capacities in the ascending order of orbits $1, 2, ...M - 1, M$, ensures minimum loss of retrial customers from the immediately preceding orbits.
- The increase in the rates of retrials (in their ascending order) also ensures minimum loss of customers to orbit-j from orbit-$j - 1$ for $j = 1, 2, .., M$ Here $j = 1$ means orbit zero which "holds primary customers".
- Computation of certain probabilities (distributions) of significance is done in this paper.
- Because of abandonment of the system by customers in orbit-M upon a retrial, the RTQ under consideration is stable by Tweedie's theorem
- The increasing nature of retrial rates from orbits can ensure minimum loss of retrial customers due to finite capacity restriction of orbits $1, 2, . . . , M - 1$.
- The Idle time of the server also gets reduced due to increasing retrial rates.

Notations and Abbreviations

The following abbreviations and notations are used in this manuscript:

RTQ	Retrial Queue
CTMC	Continuous Time Markov chain
QBD process	Quasi-Birth–Death Process
LDQBD process	Level Dependent Quasi-birth–Death process
MAM	Matrix Analytic Method
P{a}	Probabilty of a
e	All one vector with appropriate dimension
I_n	Identity matrix of order $n \times n$
A'	Transpose of matrix A
0	Matrix whose entries are 0, of appropriate size
$A \otimes B$	Kronecker product; if $A = [a_{ij}]$ is a matrix of order $m \times n$ and if B is a matrix of order $p \times q$, then $A \otimes B = [a_{ij}B]$ will denote a matrix of order $mp \times nq$
λ	arrival rate
μ	Service rate
θ	retrial rate
p	probability with which search has been done
q	probability with which a customer returns to orbit M

The remainder of this paper is as follows. In Section 2, the description of the problem is given. In Section 3, the mathematical modeling of the problem and its analysis has been done. Steady-state analysis has been done in Section 4 and the steady-state probabilities were found. Section 5 contains performance measures. Section 6 contains some probability distributions of interest that are related to the model. Numerical and graphical illustrations that provide insight into the operation of the system are included in Section **??**. Conclusions are provided in Section 7.

2. Model Description

The model considered in this paper is a single server retrial queue in which M orbits are available for retrial customers. These orbits are arranged in sequential order $1, 2, ...,M$. The first one is for primary customers who fail to access the server on their arrival; the second one is for failed retrials from orbit-1, the third one is for failed retrials from orbit-2, and so on. Finally, the Mth orbit is for failed retrials from orbit-$(M - 1)$. If a retrial attempt from orbit-M is not successful, the customer returns to the same orbit with some probability q. If a retrial from the Mth orbit is not successful, with probability $(1 - q)$ customers leave the system forever. The capacities of orbits $1, 2, ..., M - 1$ are finite: $K, 2K, ..., (M - 1)K$ respectively. This gives an exact picture of the number of retrials made by retrial customers up to orbit-$(M - 1)$. The finite capacity restriction of these orbits may lead to loss of retrial customers. The retrial rates from these orbits are in increasing order. For easy exposition, we assume these to be $\theta, 2\theta, . . . , M\theta$. The inter retrial time is exponentially distributed with parameter $r\theta$ for customers in the rth orbit for $r = 1, 2, . . . , M$. Retrial rates from all orbits are

according to the linear retrial policy. The service rates of all customers are independent and identically distributed exponential random variables with parameter μ. This can be made to depend on the nature of customers, for example, for primary customers, the parameter is μ, for customers from orbits $1, 2, 3, \ldots, M$ the parameters are $2\mu, 3\mu, \ldots, (M+1)\mu$, respectively. This will lead to an increase in the dimension of the resulting CTMC because we have to include the additional information on whether the customer in service is primary or from the orbit-j, $j = 1, 2, \ldots, M$. This is what we consider in this manuscript. After each service completion, with probability p the server searches for customers from orbit-M, and with probability $(1 - p)$, it remains idle. The search time is assumed to be negligible. The diagram given in Figure 1 gives a representation of the process of primary customer arrival, retrial by orbit customers, and abandonment/return to orbit-M.
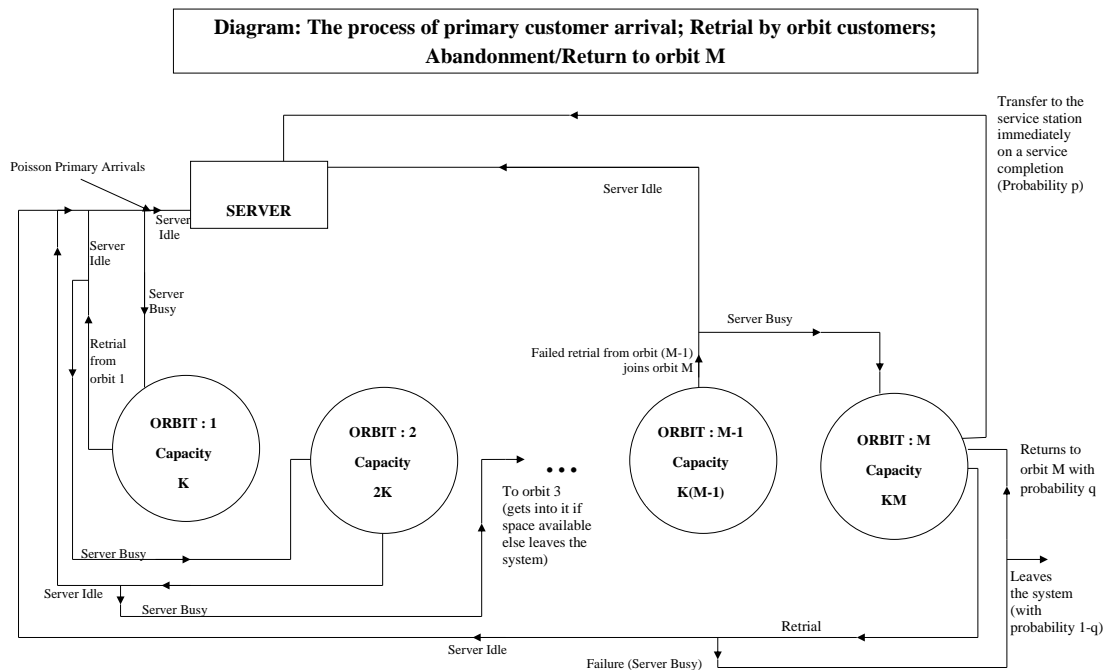

Figure 1. Diagrammatic Representation of the Model.

3. Mathematical Formulation

Let

- $N_i(t)$ be the number of customers in the orbit-i at time t for $i = 1, 2, \ldots, M$;
- $C(t)$ be the status of the server at time t

$$
C(t) = \begin{cases} 0, & \text{if the server is idle} \\ 1, & \text{if the server is busy with a primary customer} \\ j + 1, & \text{if the server is busy with customers from orbit-}j\text{, j=1,2\ldots,M} \end{cases}
$$

$$
\phi(t) = \{(N_M(t), N_{(M-1)}(t), \ldots, N_1(t), C(t)); t \geq 0\}
$$

is an irreducible continuous-time Markov chain and it describes the process under consideration.

The state space is Ω where

$$
\Omega = \{(i_M, i_{(M-1)}, \ldots, i_1, j)\}
$$

where

$$
i_M \geq 0; 0 \leq i_r \leq rK
$$

Let the ordering of the elements of Ω be lexicographical. We form the so called states $(i_M, i_{(M-1)}, \ldots, i_1, j)$ of the Markov chain $\phi(t)$. We analyze the transitions of the Markov chain $\phi(t)$ during an interval having an infinitesimal duration. We can form the matrices defining the transition rates of this chain. The infinitesimal generator Q of the LDQBD describing the model under consideration is of the form

$$
Q = \begin{pmatrix}
A_1^0 & A_0 & & & \\
A_2^1 & A_1^1 & A_0 & & \\
& A_2^2 & A_1^2 & A_0 & \\
& & \cdots & \cdots & \cdots
\end{pmatrix}
\tag{1}
$$

where A_0 represents the rate matrix corresponding to the arrival of a customer to the Mth orbit, that is the transition from level $i_M \to i_M + 1$ where $i_M \geq 0$ and it is independent of i_M,

A_2^i represents the rate matrix corresponding to the loss of a customer from the Mth orbit as a result of the retrial, that is, transitions from level $i_M \to i_M - 1$; for $i_M = 1, 2, \ldots,$ and

A_1 describes all transitions in which the level does not change (transitions within levels i_M).

The structure of the A_1^i, A_0, A_2^i for $i \geq 0$ can all be defined in terms of transition rates corresponding to the transitions of the states given in the Table 1 The first column defines the state from which a transition can occur, the second column defines a state to which a transition can occur, the third column describes the condition when this transition occurs and the last column contains the rate of the corresponding transition.

Table 1. Rate of Transitions.

From	To	Description	Transition Rate
$(i_M, .., i_1, 0)$	$(i_M, .., i_1, 1)$	arrival of a primary customer to an idle server	λ
$(i_M, .., i_1, 1)$	$(i_M, .., i_1 + 1, 1)$	entry of a customer to the orbit-1	λ
$(i_M, .., i_r, \ldots, i_1, 0)$	$(i_M, .., i_r - 1, \ldots, i_1, 1)$	successful retrial from the orbit-r	$i_r \theta$
$(i_M, .., j_r \neq \Omega, i_1, j)$;	$(i_M, ..i_r - 1, \ldots, i_{(r+1)} + 1, \ldots, i_1, j); j \neq 0$	entry in to orbit-(r+1) from the orbit-r, $r \neq M$	$i_r \theta$
$(i_M, .., i_r, \ldots, i_1, j)$	$(i_M - 1, .., i_r, \ldots, i_1, M + 1)$	successful search from orbit-M	$pj\mu$
$(i_M, .., i_r, \ldots, i_1, j)$	$(i_M, .., i_r, \ldots, i_1, 0)$	unsuccessful search from orbit-M if $j \neq 0$	$(1 - p)j\mu$
$(i_M, .., i_r, \ldots, i_1, j)$	$(i_M - 1, .., i_r, \ldots, i_1, j)$	customers leave the system from orbit-M if $j \neq 0$	$(1 - q)i_M \theta$
$(i_M, .., i_r, \ldots, i_1, j)$	$(i_M - 1, .., i_r, \ldots, i_1, j)$	customers leave the system from orbit-r if $i_r = rK$; $j \neq 0$	$i_r \theta$
$(i_M, .., i_r, \ldots, i_1, j)$	$(i_M, .., i_r, \ldots, i_1, j)$	customers return to orbit-M if $j \neq 0$ from orbit-M if $j \neq 0$	$qi_M \theta$

The retrial queueing model considered here can be studied as a level dependent Quasi Birth Death process and we apply Neuts-Rao truncation for the analysis of the model. We assume that when the number of customers in the Mth orbit exceeds a certain limit, say N, retrial from that orbit occurs at constant rates $N\mu$. In that situation the matrices A_2^i becomes A_2^N and A_1^i becomes A_1^N for $i \geq N$. The infinitesimal generator Q^1 of the modified model becomes

$$Q^1 = \begin{pmatrix} A_1^0 & A_0 & & & & \\ A_2^1 & A_1^1 & A_0 & & & \\ & A_2^2 & A_1^2 & A_0 & & \\ & & \cdots & \cdots & \cdots & \\ & & & A_2 & A_1 & A_0 \\ & & & & \cdots & \cdots & \cdots \end{pmatrix} \tag{2}$$

where

$$A_1 = A_1^N, A_2 = A_2^N$$

.

4. Steady-State Analysis

4.1. Stability Condition

Theorem 1. *The system under consideration is always stable.*

Proof. Let Ψ be a Lyapunov test function defined by

$$\Psi(u) = i \tag{3}$$

where u denotes a state in the ith level. The mean drift for the state u belonging to the ith level is given by

$$y_u = \sum_{r \neq u} q_{ur}(\Psi(r) - \Psi(u)) \tag{4}$$

$$= \sum_{u'} q_{uu'}(\Psi(u') - \Psi(u)) + \sum_{u''} q_{uu''}(\Psi(u'') - \Psi(u)) + \sum_{u'''} q_{uu'''}(\Psi(u''') - \Psi(u)) \tag{5}$$

where u' is a state in the $(i-1)^{th}$ level, u'' is a state in the ith level and u''' is a state in the $(i+1)^{th}$ level, q_{ij} is the $(i,j)^{th}$ entry of the infinitesimal generator. So $\Psi(u) = i, \Psi(u') = i - 1, \Psi(u'') = i, \Psi(u''') = i + 1$.

$$y_u = \sum_{u'} q_{uu'}(i - 1 - i) + \sum_{u''} q_{uu''}(i - i) + \sum_{u'''} q_{uu'''}(i + 1 - i) \tag{6}$$

$$= \sum_{u'} q_{uu'}(-1) + \sum_{u''} q_{uu''}(0) + \sum_{u'''} q_{uu'''}(1) \tag{7}$$

$$= -\sum_{u'} q_{uu'} + \sum_{u'''} q_{uu'''} \tag{8}$$

We have, $\sum_{u'''} q_{uu'''} = c_1, \sum_{u'} q_{uu'} = c_2 + i\theta$, where c_1 and c_2 are constants. So the mean drift

$$y_u = c_1 - c_2 - i\theta \tag{9}$$

which depends only on the level i. Thus for any $\epsilon > 0$, we can find N' large enough so that

$$y_u < -\epsilon \tag{10}$$

for any u belonging to level $i \geq N'$. Hence by Tweedie's result [18], the theorem follows.

$\square$

4.2. Stationary Distribution

The stationary distribution of the Markov process under consideration is obtained by solving the set of equations

$$\mathbf{x}\mathbf{Q^1} = \mathbf{0}, \mathbf{x}\mathbf{e} = \mathbf{1}.$$

Let $\mathbf{x}$ be the steady- state probability vector of Q^1.

Partition this vector in conformity with Q^1 as follows:

$$\mathbf{x} = (\mathbf{x(0)}, \mathbf{x(1)}, \mathbf{x(2)}, \dots,)$$

$$\mathbf{x}(i) = (\mathbf{x}(i,0), \mathbf{x}(i,1), \dots, \mathbf{x}(i, (M-1)K)$$

$\mathbf{x}(i,j) = (\mathbf{x}(i,j,0), \mathbf{x}(i,j,1), \dots, \mathbf{x}(i,j,(M-2)K))$ and so on. Finally we get elements of the form $x(i_M, i_{(M-1)}, i_{(M-2)}, \dots, i_1, j)$ and it is the probability of being in state $(i_M, i_{(M-1)}, i_{(M-2)}, \dots, i_1, j)$ $i_M \geq 0; 0 \leq i_r \leq rK; j = 0, 1, 2, \dots, M+1$

The steady-state probability vector is obtained as

$$\mathbf{x}_{(N-1)+i} = \mathbf{x}_{(N-1)} R^i, i \geq 0 \tag{11}$$

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

$$R^2 A_2 + R A_1 + A_0 = 0 \tag{12}$$

and the vectors $\mathbf{x_0}, \dots, \mathbf{x_{(N-1)}}$ are obtained by solving

$$\mathbf{x_0} A_1^0 + \mathbf{x_1} A_2^1 = 0 \tag{13}$$

$$\mathbf{x_{(i-1)}} A_0 + \mathbf{x_i} A_1^i + \mathbf{x_{(i+1)}} A_2^{(i+1)} = 0 \tag{14}$$

For $1 \leq i \leq (N-2)$

$$\mathbf{x_{(N-2)}} A_0 + \mathbf{x_{(N-1)}} \left[A_1^{(N-1)} + A_2 R \right] = 0 \tag{15}$$

subject to the normalizing condition

$$\sum_{i=0}^{(N-2)} \mathbf{x_i} + \mathbf{x_{(N-1)}} (I - R)^{-1} \mathbf{e} = 1. \tag{16}$$

5. Performance Measures

Some measures of performance, which help the operators of the system to make decisions concerning the optimal values of the number M of hierarchical orbits to be maintained and the number K which decides customers to be allowed in those hierarchical orbits are evaluated. Following are some performance measures which help us to make a detailed study about the problem under consideration. For the evaluation of the performance measures we assume that summation w.r.t each i_r is up to rk

1. Expected number of customers in the Mth orbit

$$E[N_M] = \sum_{i=0}^{\infty} i\mathbf{x}(\mathbf{i})\mathbf{e} \tag{17}$$

2. Expected number of customers in the rth orbit

$$E[H(r)] = \sum_{i_M=1}^{MK} \sum_{i_{(M-1)}=1}^{(M-1)K} \ldots, \sum_{i_r=1}^{rK} \mathbf{x}(i_M, i_{M-1}, i_{M-2}, \ldots, i_r)\mathbf{e} \tag{18}$$

3. Probability that the server is busy with customers from orbit-r

$$P_r = \sum_{i_M=1}^{MK} \sum_{i_{(M-1)}=1}^{(M-1)K} \ldots, \sum_{i_1=1}^{K} \mathbf{x}(i_M, i_{M-1}, i_{M-2}, \ldots, i_1, r+1)\mathbf{e} \tag{19}$$

4. Probability that the server is busy

$$P_{busy} = \sum_{r=1}^{M+1} P_r \tag{20}$$

5. Probability that the server is idle

$$P_{idle} = \sum_{i_M=1}^{MK} \sum_{i_{(M-1)}=1}^{(M-1)K} \ldots, \sum_{i_1=1}^{K} \mathbf{x}(i_M, i_{M-1}, i_{M-2}, \ldots, i_1, 0)\mathbf{e} \tag{21}$$

6. Expected number of customers in the system

$$E[N_s] = \sum_{r=1}^{M} E[H(r)] + P_{busy} \tag{22}$$

7. Expected number of customers entering the service facility as a result of orbital search

$$E[S] = \sum_{i_M=0}^{\infty} \sum_{i_{(M-1)}, \ldots, i_1} p * i_M * \mathbf{x}(i_{(M-1)}, \ldots, i_1)\mathbf{e} \tag{23}$$

8. Expected number of customers leaving the system from orbit -M as a result of unsuccessful retrial

$$E[L_M] = \sum_{i=0}^{\infty} (1-q)i * \mathbf{x}(i)\mathbf{e} \tag{24}$$

6. Some Probability Distributions

Theorem 2. *The probability of a primary customer accessing an idle server*

$$P[R_p] = P_{idle} = \sum_{i_M=1}^{MK} \sum_{i_{(M-1)}=1}^{(M-1)K} \ldots, \sum_{i_1=1}^{K} \mathbf{x}(i_M, i_{M-1}, i_{M-2}, \ldots, i_1, 0)\mathbf{e} \tag{25}$$

Theorem 3. *The probability of a primary customer being able to get into orbit-1= $(1 - P_{idle})$ *P{(orbit-1 is not full)}*

Theorem 4. *Probability of a customer in orbit-j, successfully accessing the server on retrial*
$P[R_p] = P_{idle} * \sum_{k=1}^{jK} kj\theta$* *P{ k customers in orbit -j} j = 1, 2, M - 1*

Theorem 5. *Probability of a customer retrying to access the server $M - 1$ times with the last retrial turning out to be a success= P_{busy}^{M-1} *P_{idle}*

Proof. In the primary arrival and subsequent $M - 2$ retrials the customer encounters a busy server and in the $M - 1$ retrial he is able to access the server $\square$

6.1. A Lower Bound for the Time Taken to Access the Server

It is extremely hard to compute the exact distribution of the time until a customer is able to access the server. Statements of theorems 2 to 5 given above provide certain probabilities in this connection; however those do not say anything about the distribution of time until a customer, starting from the time he gets into the system until either served out or leaves the system because of the orbit being full or could be because of impatience. We provide bounds on the distributions of the time to access the server or leave the system without taking service.

The following lemmas provide distribution of time till accessing the server

Lemma 1. *The distribution of the time until accessing the server upon reaching the system immediately on arrival ie; $\tau=0$ is given by Prob{server seen idle upon primary arrival}=$P\{\tau = 0\}$ =P_{idle}*

Lemma 2. *The distribution of the time until accessing the server upon reaching the system conditional on accessing the server at the first retrial is $Prob\{\tau = T_1^{(n)}\}$ where $T_1^{(n)}$ is the random variable representing the time to access the server conditional on n customers in that orbit. This has exponential distribution with parameter $2n\theta$ given by =*

$$\sum_{n=1}^{K} (1 - e^{-2n\theta t}) * P_{idle} \tag{26}$$

Extending the above result to orbit-j, conditional on accessing the server by a retrial customer in orbit-j, the distribution of τ is given by the recursive relation

$$\tau^{(j)} = \tau^{(j-1)} + T_j$$

where $\tau^{(j-1)}$ is the random duration of time a customer spent until reaching orbit-j and T_j is the duration of time he spends in that orbit. T_j has the exponential distribution with parameter $nj\theta$ where n is the number of customers in orbit-j. Thus

$$\tau^{(1)} = T_1$$

where $T_1 \sim \sum_{n=1}^{K}(1 - P_{idle})exp(\lambda)$, $\tau^{(2)} = \tau^{(1)} + T_2$ where $T_2 \sim exp(nj\theta)$ when there are n customers in orbit-j ; for $j = 1, 2, \ldots, M - 1$. When $j = 1$, $\tau^{(j-1)} = \tau^{(0)}$ is defined as zero. Finally, for $j = M$, $\tau^{(M)} = \tau^{(M-1)} + T_M$ where $T_M \sim exp(nj\theta)$ when n customers are in orbit-M. Since we are looking for the distribution of the minimum time a customer spends in the system, conditional on meeting an idle server as a primary customer, or failing which access from orbit-1, failing which accessing the server from orbit-2,... and if all these retrials turn out to be failures, accessing the server from orbit-M in the first retrial on getting into that orbit or through the search by the system(negligible time, immediately on completion of service)

Important Note: It is to be noted that T_1 should be multiplied by the Probability that the server is idle at the time when the retrial customer tries to reach the server and $\tau^{(1)}$ has to be multiplied by the probability that the server busy when the customer tries to access the server. In general distribution, of $\tau^{(j-1)}$ in

$$\tau^{(j)} = \tau^{(j-1)} + T_j$$

has to be multiplied probability that the server is busy when the customer retries from orbit-$j - 1$ and distribution T_j has to be multiplied by probability that server idle when the customer retries from orbit-j for $j = 1, 2, \ldots, M - 1$

A customer on reaching orbit -M , will be either chosen for service through a search made at the end of a service completion or by repeated retrials. Assuming that the customer does not leave the system before accessing the server we get the following results:

Theorem 6. *Denote by τ the random duration the customer spends in orbit before successfully accessing the server. Then with $\phi = P\{serverbusy\}$,*

1.

$$P\{\tau \le t\} = \sum_{j=1}^{\infty}$$

P{j-1 unsuccessful retrials followed by successful one before time t}

$$= \sum_{n_1,\dots,n_j=1}^{\infty} \phi^{(j-1)}(1-\phi)n_1 n_2 \dots n_{(j-1)}\theta^{(j-1)}M^{j-1}q^{j-1}\int_0^{u_1}\dots\int_{u_{j-2}}^{u_{j-1}}\int_{u_{j-1}}^{t} e^{-n_1\theta Mu_1}du_1 e^{-n_2\theta Mu_2}du_2 \dots e^{-n_j\theta Mu_j}du_j$$

2. *Having failed j retrials($j = 0, 1, \dots$), the customer is transferred to the service station at the end of a service:*

$$P\{\tau \le t\} = \sum_{j=0}^{\infty}\phi^j\int_0^t\int\int\cdots\int_{u_{j-1}}^{u_j}\sum_{n_1,\dots,n_j=1}^{\infty}e^{-n_1\theta Mu_1}du_1\dots e^{-n_j\theta Mu_j}du_j n_1 n_2\dots n_{(j}\theta^{(j)}M^j q^j\Big(\int_{u_j}^t\sum_{k=0}^{\infty}k\theta e^{-k\theta v}dv\Big)pdu_1\dots du_j$$

.

Here the last integral stands for the service completion of a customer in $(v, v + dv)$ and immediately the customer in orbit is taken for service through search(probability p). The other integrals represent unsuccessful retrials by the customer j.

7. Numerical Examples

The following example gives an illustration of the effect of search probability on the performance measures of the system. Here we assume that the arrival follows Poisson distribution with parameter λ.. Also, the retrial rates are fixed for the customers from all the orbits, and we take θ as the retrial rate. The service rates for all customers are taken as μ. In this example, the following values are kept fixed:

$$\lambda = 2; \theta = 3; \mu = 18; K = 3; M = 2.$$

From Table 2 and the corresponding graphical illustrations, the following conclusions can be drawn:

- Figure 2 it can be seen that the expected number of customers in orbit-M decreases as p increases.
- From Figure 3 we can conclude that the expected number of customers in the first orbit decreases as p increases.
- Figure 4 shows that the expected number of customers getting service as a result of search increases as p increases.

- Figure 5 shows that, P_{busy} increases with an increase in value of p.
- Figure 6 shows that P_0 decreases with an increase in value of p.
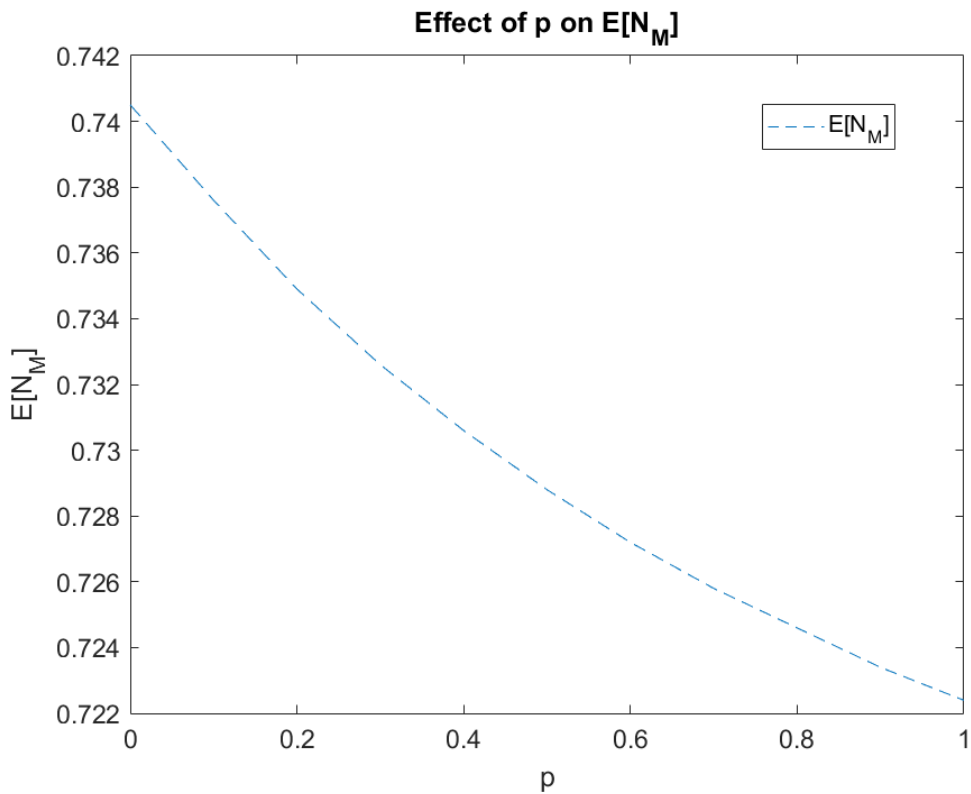- Maximum server utilization is guaranteed if search is done after each service completion.
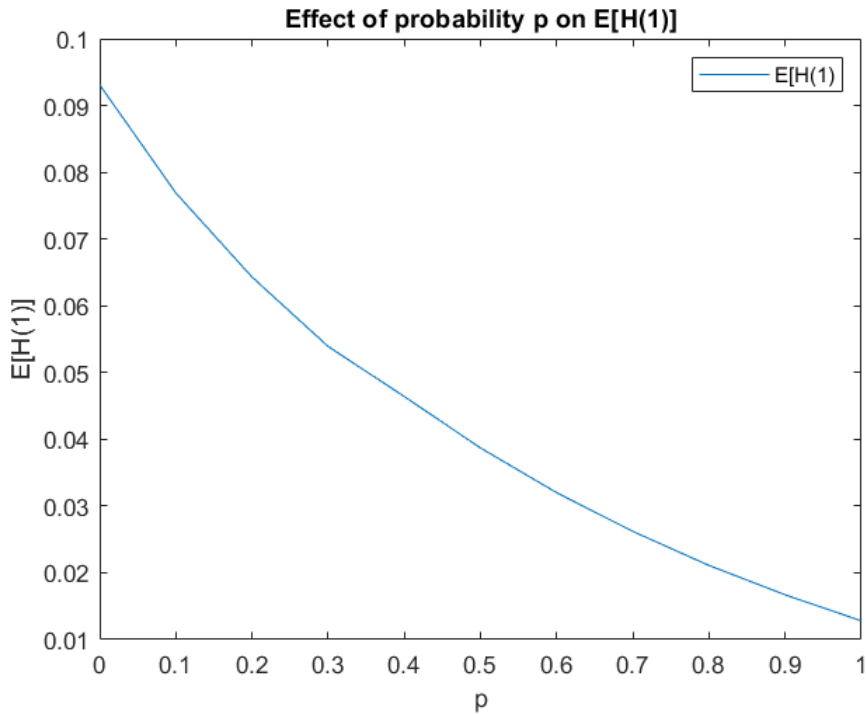
Figure 2. Effect of p on $E[N_M]$.
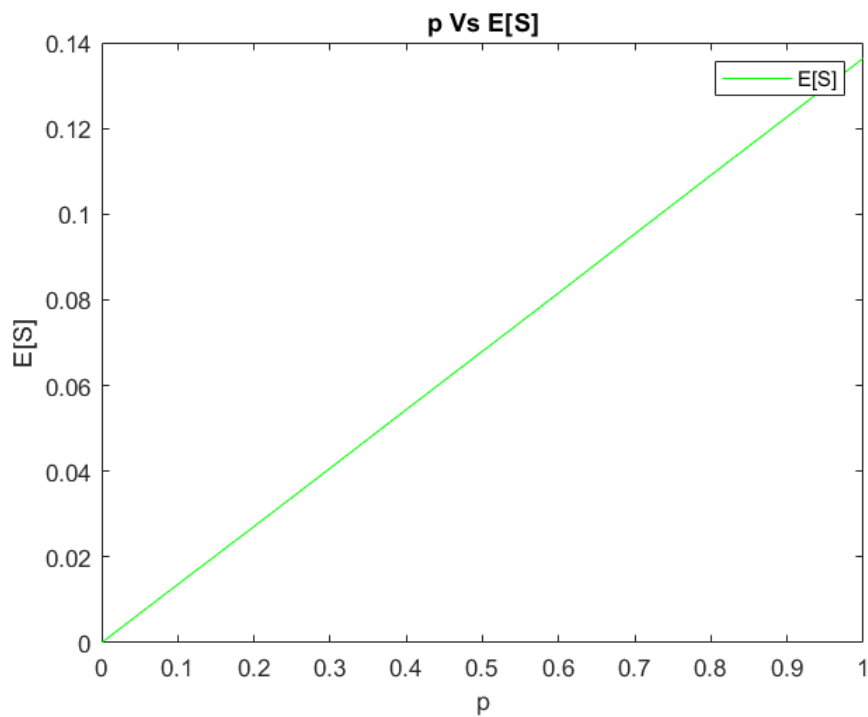

Figure 3. Effect of p on number of customers in orbit-1.
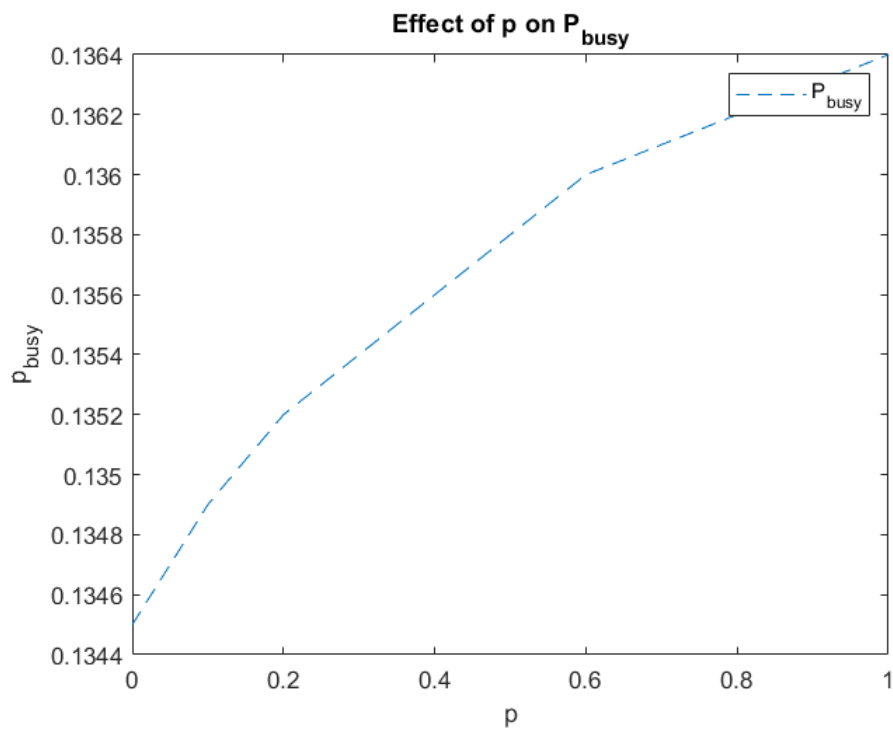
Figure 4. Effect of p on $E[S]$.

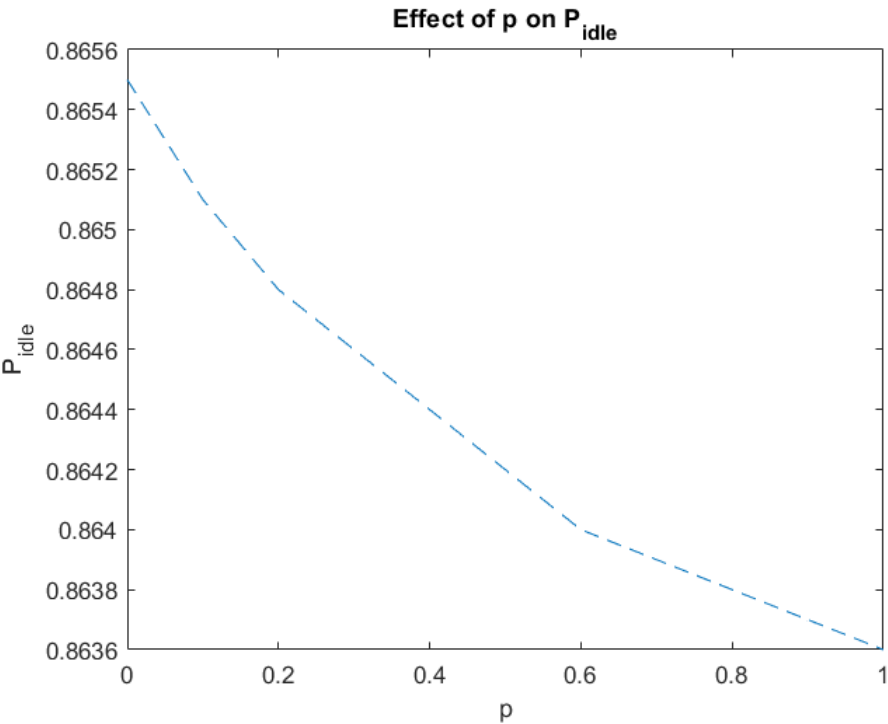**Figure 5.** Effect of p on P_{busy}.

Figure 6. Effect of p on P_{idle}.

Table 2. Effect of Search probability p on System Performance measures.

p	$E[N_M]$	$E[H(1)]$	$E[S]$	P_{busy}	P_{idle}
0	0.7405	0.0931	0	0.1345	0.8655
0.1	0.7376	0.0769	0.0136	0.1349	0.8651
0.2	0.7349	0.0643	0.0271	0.1352	0.8648
0.3	0.7326	0.0539	0.0407	0.1354	0.8646
0.4	0.7306	0.0464	0.0544	0.1356	0.8644
0.5	0.7288	0.0387	0.0680	0.1358	0.8642
0.6	0.7272	0.0320	0.0816	0.1360	0.8640
0.7	0.7258	0.0262	0.0954	0.1361	0.8639
0.8	0.7246	0.0211	0.1091	0.1362	0.8638
0.9	0.7234	0.0167	0.1227	0.1363	0.8637
1	0.7224	0.0128	0.1364	0.1364	0.8636

Conclusion and Future Work

In this paper, we considered a retrial queueing system with more than one orbit $(M \leq \infty)$ in which the first $M - 1$ orbits have finite capacities with increasing waiting space from orbit-1 to orbit-2, then to orbit-3 and so on, from orbit-$M - 2$ to orbit- $M - 1$, for retrial customers to move in the order of the number of retrials each one has completed. Orbit- M is of unbounded capacity. Customers in this orbit are transferred to the service station (one at a time) at a service completion epoch with probability p for the next service. Because of the finiteness of the capacities of the first M-1 orbits a large number of customers are bound to be lost to the system. The only check to this is the increasing retrial rates (linear). Further, the fact that customers from orbit abandon the system consequent to a retrial (with positive probability), the system turns out to be stable in the long run. We derive the long-run system distribution; several performance metrics are evaluated; the least upper bounds of certain distributions associated with the model are also derived; numerical illustrations are provided. We see Telecommunication as an avenue for the application of this model. This is so because, messages which are not transmitted within a specified time lose their information values. As for future work, there are umpteen possibilities:

1. The arrival process constitutes a Markovian arrival process (MAP) and service times follow distinct phase type distributions based on the customer category—primary (external arrival) or retrial customer (based on the number of retrials (valid only up to the first M-1 orbits).
2. Constant retrial rate of orbital customers (only by the head of the orbital queues) with distinct parameters for distinct orbits.
3. A very hard problem is to give all orbits infinite capacity and to assume that failed retrial customers from these orbits abandon the system.

References

1. Artalejo, J.R.: Accessible bibliography on retrial queues: Progress in 2000-2009. *Math.and Comput. Modell. 2010, 51*, 1071–1081
2. Artalejo, J.R.: A classified bibliography of research on retrial queues: Progress in 1990-1999. *Top 7,1999* 187–211
3. Artalejo, J.R., Joshua, V.C., Krishnamoorthy, A.: An M/G/1 Retrial Queue with Orbital Search by Server. *Advances in Stoch. Modell.2002* 41–54. Notable Publications, New Jersey
4. Artalejo,Jesús R., Gómez-Corral,Antonio.:*Retrial Queueing Systems:A computational approach,2008,*Springer-Verlag Berlin Heidelberg
5. Chakravarthy, S.R; *Introduction to Matrix Analytic Methods in Queues12022,* Wiley ISTE
6. Chakravarthy, S.R; *Introduction to Matrix Analytic Methods in Queues22022,* Wiley ISTE
7. Dudin, A.N., Krishnamoorthy, A., Joshua, V.C., Tsarenkov,G.: Analysis of BMAP/G/1 retrial system with search of customers from the orbit. *Eur. J. of Oper. Res.2004,157*, 169–179
8. Dudin, A.N., Deepak, T.G., Krishnamoorthy, A., Joshua, V.C., Vishnevsky, V.: On a BMAP/G/1 retrial system with two types of search of customers from the orbit. *Int. Conference on ITMM, 1–12. Springer 2017*
9. Dudin, A.N., Klimenok, V.I., Vishnevsky. V.M; *The Theory of Queueing systems with correlated flows,2020,* Springer
10. Falin, G.I., A survey of Retrial queues. *Queueing syst.19907*, 127–167
11. Falin, G.I., Templeton J.G.C.: *Retrial Queues*. Chapman and Hall, London (1997)
12. Joshua,V.C., Ambily P. Mathew.: Krishnamoorthy,A.: A Retrial queueing System in which the Server Searches to Accumulate Customers for Optimal Bulk Serving *Springer LNCS Series .2020* vol 12563.Springer, cham. pp 311-321
13. Kim, J., Kim, B.:*A survey of retrial queueing systems. (2016)*Annals of Operations Research 247, 3–36 . https://doi.org/10.1007/s10479-015-2038-7
14. Krishnamoorthy, A., Deepak, T.G., Joshua, V.C.: An M/G/1 Retrial Queue with Non persistent Customers and Orbital Search.*Stoch. Anal. and Appl.200523*, 975–997
15. Krishnamoorthy, A., Joshua, V.C., Ambily, P.M.: A retrial queueing system with abandonment and search for prioritycustomers. *Springer CCIS, 2017700,*98–107.
16. Neuts, M.F., Ramalhoto.: A service model in which the server is required to search for customers. *J. of Appl. Prob.1984 21*, 57–166
17. Neuts, M.F., Rao.,B.M.: Numerical investigation of a multiserver retrial model.*Queueing syst.,1990,7*, 169–189
18. R.L.Tweedie:, Sufficient condition for regularity, recurrence and ergodicity of Markov processes. *Proc. Camb. Phil. Soc.,1975 78*:Part I

19. Krishnamoorthy,A., V.C. Joshua and Ambily P. Mathew, A Retrial Queueing System with Multiple Hierarchical Orbits and Orbital Search,, *Springer, CCIS, 2018, 919,* 224-233
20. Phung-Duc, Tuan. *Retrial queueing models: A survey on theory and applications.* arXiv preprint arXiv:1906.09560 (2019).