RAIRO-Oper. Res. 54 (2020) 81–99
https://doi.org/10.1051/ro/2018118

RAIRO Operations Research
www.rairo-ro.org

STOCHASTIC DECOMPOSITION IN RETRIAL QUEUEING-INVENTORY SYSTEM

Dhanya Shajin[1] and A. Krishnamoorthy[2,*]

Abstract. The purpose of this paper is to obtain product form solution for retrial – queueing – inventory system. We study an $M/M/1$ retrial queue with a storage system driven by an (s, S) policy. When server is idle, external arrivals enter directly to an orbit. Inventory replenishment lead time is exponentially distributed. The interval between two successive retrials is exponentially distributed and only the customer at the head of the orbit is permitted to access the server. No customer is allowed to join the orbit when the storage system is empty and also when the serer is busy. We first derive the stationary joint distribution of the queue length and the on-hand inventory in explicit product form. Using the joint distribution, we investigate long-run performance measures such as distribution of number of customers served, number of arrivals, number of customers lost during an interval of random duration and a cost function. The optimal pair (s, S) is numerically investigated.

Mathematics Subject Classification. 90B05, 90B22, 90B50.

Received September 6, 2016. Accepted December 9, 2018.

1. Introduction

In this paper we produce a stochastic decomposition (and even more, a product form solution) for a retrial – queueing – inventory model. To this end we construct a (partially) blocking set (for the arrival process).

Blocking sets (more aptly partial blocking sets) have been discussed in Krenzler and Daduna [6], among several other researchers to produce product form solution (see for example [3–5, 11, 14–16]). A discussion on optimal blocking set could be found in Krishnamoorthy *et al.* [8]. Schwarz *et al.* [16] are the first to produce form solution in queueing-inventory models with positive lead time (in this paper they assumed exponential distribution). Krishnamoorthy and Viswanath [11] extended it to the production inventory set up and Saffari *et al.* [15] to the case of arbitrarily distributed lead time.

Devi et al. [7] considered the problem of controlling the selection rates of the pooled customer of a single commodity inventory system with postponed demands. The demands arrive according to a Poisson process. The ordering policy is (s, S) policy that is as and when the inventory level drops to s an order for $Q(= S - s)$ items is placed. The ordered items are received after a random time, which is distributed as exponential. They assume that the demands that occur during stock out period either enter a pool of finite size or leave the system

Keywords. queueing-inventory, stochastic decomposition, retrial.

[1] Department of Mathematics Amrita School of Arts and Sciences Amrita Vishwa Vidyapeetham, Cochin 682024, India.

[2] Department of Mathematics, CMS College, Kottayam 686001, India.

*Corresponding author: achyuthacusat@gmail.com

Article published by EDP Sciences

according to a Bernoulli distribution. Whenever the on-hand inventory level is positive, customers are selected one-by-one and the selection rate can be chosen from a given set. The problem is to determine a decision rule that specifies the rate of these selections as a function of the on-hand inventory level and the number of customers waiting in the pool at each instant of time to minimize the long-run total expected cost rate. The problem is modeled as a semi-Markov decision problem. The optimal policy is computed using Linear Programming algorithm and the results are illustrated numerically.

A single-server queueing system with a marked Markovian arrival process of heterogeneous customers is considered in Baek *et al.* [2]. Type-1 customers have limited preemptive priority over type-2 customers. There is an infinite buffer for type-2 customers and no buffer for type-1 customers. There is also a finite buffer for consumable additional items which arrive according to the Markovian arrival process. Service of a customer requires a fixed number of consumable additional items depending on the type of the customer. The service time has a phase-type distribution depending on the type of the customer. Customers in the buffer are impatient and may leave the system without service after an exponentially distributed amount of waiting time. Aiming to minimize the loss probability of type-1 customers and maximize throughput of the system, a threshold strategy of admission to service of type-2 customers is offered. Service of type-2 customer can start only if the server is idle and the number of consumable additional items in the stock exceeds the fixed threshold. Stationary distributions of the system states and the waiting time are computed. Dhanya *et al.* [9] extended findings in Baek *et al.* [2] to retrial of low priority customers.

Investigation on stochastic decomposition of retrial – queueing – inventory had not produced the desired result for the past one decade. This could basically be attributed to the fact that an appropriate blocking set was evading the researchers. The main objective of this paper is to produce such a blocking set and thereby achieve the desired result. In this process we also provide a geometric distribution, except for the multiplicative constants, for the 'modified' retrial queueing process that we introduced to produce stochastic decomposition.

Attached inventory is controlled by the (s, S) policy. Neverthless, we can extend the results obtained here to the control policies such as $(s, Q), (S-1, S)$ and random order cases. The (s, S) policy is as follows: Assume that each customer demands exactly one unit of the item at the end his service. So the inventory depletion is by exactly one unit. From S, when $(S-s)$ items are served out, the inventory level reaches s which triggers an order placement for replenishment. As and when order materialization takes place (lead time is exponentially distributed (with parameter γ)), the quantity replenished is that much to bring the inventory level to S.

As an example for the model under study we may think of a polling model where the gate is closed the moment the server starts service at a node.

Primary arrivals (which form a Poisson process of rate λ) are directed to an orbit of infinite capacity (see [13]) where from they try to access the server according to FIFO discipline (that is a queue of customers is formed). However, we impose a restriction on the entry to orbit of these primary customers – they are not allowed to join the orbit when a service is proceeding. Only when server is idle do primary customers join orbit (they do not walk into the counter straight from outside) and form a queue. Further, the head of the queue in the orbit alone retries which is according an exponentially distributed inter-occurrence time with parameter θ. On retrial, if the server is found to be occupied, the customer returns to orbit as the head of the queue. Service time duration of customers are i.i.d $\exp(\mu)$ random variables. Another crucial assumption that we make, as in Schwarz *et al.* [16], Saffari *et al.* [15], Krishnamoorthy and Viswanath [11], is that when inventory level is zero, no primary customer joins orbit or orbital customers retry (even if they do so, they return to orbit).

For a survey of investigation on queueing – inventory process one may refer to Krishnamoorthy *et al.* [10].

The rest of this paper is arranged as follows. A brief description of the problem along with its modelling and analysis are provided in Section 2. The long run system state distribution is discussed in Section 3. It also provides the blocking set that helped to develop the stochastic decomposition. Section 4 discusses in detail important measures of performance of the system. It also contains the expected waiting time of a customer in the orbit. Extensive numerical investigation is provided in Section 5. That section also contains numerical investigation of a cost function providing the minimum running cost and optimal (s, S) pair.

Notations and abbreviation used in the sequel are

e column vector of 1's with appropriate order
0 vector consisting of 0's with appropriate dimension
O zero matrix with appropriate order
$CTMC$ continuous time Markov chain
$LIQBD$ level-independent quasi-birth and death process
$N(t)$ number of customers in the system at time t
$I(t)$ number of items in the inventory at time t
$C(t)$ status of server at time $t = \begin{cases} 0 \text{ server is idle} \\ 1 \text{ server is busy} \end{cases}$

2. Model description

Consider an infinite capacity retrial queueing-inventory system with a single server to which customers arrive according to a Poisson process of rate λ. The service time is exponentially distributed with parameter μ. The (s, S)- control policy is adopted in which whenever the inventory level falls to the reorder level s, order for replenishment is placed to bring back to maximum level S at the time of replenishment. The lead time for replenishment follows an exponential distribution with parameter γ. For the purpose of producing a stochastic decomposition of the system state, we restrict the arrival of customers as follows: All primary arrivals (external customers) must join in orbit of infinite capacity on arrival, from where, through retrial alone, they can access the server (as in [13]). Further if the server is busy at the time an external arrival takes place, then that external customer does not join the system. The customer at the head of the orbit alone is permitted to access the server. The interval between two successive repeated attempts is exponentially distributed with parameter θ (see [1]). If there is no item in the inventory, all arriving (external) customers are lost. Then $\Omega = \{(N(t), C(t), I(t)), t \geq 0\}$ forms a $CTMC$ with state space $\{(n, 0, i); n \geq 0, 0 \leq i \leq S\} \bigcup \{(n, 1, i); n \geq 0, 1 \leq i \leq S\}$.

Now we describe the transition rates in the Markov chain $\{(N(t), C(t), I(t)), t \geq 0\}$ are

(i) $(n, 0, i) \rightarrow (n+1, 0, i)$: rate is λ for $n \geq 0, 1 \leq i \leq S$.
(ii) $(n, 0, i) \rightarrow (n-1, 1, i)$: rate is θ for $n \geq 1, 1 \leq i \leq S$.
(iii) $(n, 1, i) \rightarrow (n, 0, i-1)$: rate is μ for $n \geq 0, 1 \leq i \leq S$.
(iv) $(n, k, i) \rightarrow (n, k, S)$: rate is γ for $n \geq 0, 0 \leq i \leq s, k = 0, 1$.

Other transitions have rate 0.

Write

$$Pr.(N(t) = n, C(t) = 0, I(t) = i) = P_{n,0,i}(t),\ n \geq 0, 0 \leq i \leq S,$$

$$Pr.(N(t) = n, C(t) = 1, I(t) = i) = P_{n,1,i}(t),\ n \geq 0, 1 \leq i \leq S.$$

These satisfy the system of difference − differential equations:

$$P'_{n,0,0}(t) = -\gamma P_{n,0,0}(t) + \mu P_{n+1,1,1}(t), n \geq 0, \tag{2.1}$$

$$P'_{0,0,i}(t) = -(\gamma + \lambda) P_{0,0,i}(t) + \mu P_{0,1,i+1}(t), 1 \leq i \leq s, \tag{2.2}$$

$$P'_{0,0,i}(t) = -\lambda P_{0,0,i}(t) + \mu P_{0,1,i+1}(t), s+1 \leq i \leq S-1, \tag{2.3}$$

$$P'_{0,0,S}(t) = -\lambda P_{0,0,S}(t) + \gamma \sum_{i=0}^{s} P_{0,0,i}(t), \tag{2.4}$$

$$P'_{n,0,i}(t) = -(\gamma + \lambda + \theta) P_{n,0,i}(t) + \mu P_{n,1,i+1}(t) + \lambda P_{n-1,0,i}(t), n \geq 1, 1 \leq i \leq s, \tag{2.5}$$

$$P'_{n,0,i}(t) = -(\lambda + \theta) P_{n,0,i}(t) + \mu P_{n,1,i+1}(t) + \lambda P_{n-1,0,i}(t), n \geq 1, s+1 \leq i \leq S-1, \tag{2.6}$$

$$P'_{n,0,S}(t) = -(\lambda + \theta) P_{n,0,S}(t) + \gamma \sum_{i=0}^{s} P_{n,0,i}(t) + \lambda P_{n-1,0,S}(t), n \geq 1, \tag{2.7}$$

$$P'_{n,1,i}(t) = -(\gamma+\mu)P_{n,1,i}(t) + \theta P_{n+1,0,i}(t), n \geq 0, 1 \leq i \leq s, \tag{2.8}$$
$$P'_{n,1,i}(t) = -\mu P_{n,1,i}(t) + \theta P_{n+1,0,i}(t), n \geq 0, s+1 \leq i \leq S-1, \tag{2.9}$$
$$P'_{n,1,S}(t) = -\mu P_{n,1,S}(t) + \gamma \sum_{i=1}^{s} P_{n,1,i}(t) + \theta P_{n+1,0,S}(t), n \geq 0. \tag{2.10}$$

In the steady state (the condition for its existence is $\lambda < \theta$ which will be proved subsequently) time derivative is equated to zero.

Write

$$\lim_{t\to\infty} P_{n,0,i}(t) = p_n(0,i),\ \ n \geq 0, 0 \leq i \leq S,$$

and

$$\lim_{t\to\infty} P_{n,1,i}(t) = p_n(1,i),\ \ n \geq 0, 1 \leq i \leq S.$$

Thus the above set of equations (2.1) to (2.10) become

$$-\gamma p_n(0,0) + \mu p_{n+1}(1,1) = 0, n \geq 0, \tag{2.11}$$
$$-(\gamma+\lambda)p_0(0,i) + \mu p_0(1,i+1) = 0, 1 \leq i \leq s, \tag{2.12}$$
$$-\lambda p_0(0,i) + \mu p_0(1,i+1) = 0, s+1 \leq i \leq S-1, \tag{2.13}$$
$$-\lambda p_0(0,S) + \gamma \sum_{i=0}^{s} p_0(0,i) = 0, \tag{2.14}$$
$$-(\gamma+\lambda+\theta)p_n(0,i) + \mu p_n(1,i+1) + \lambda p_{n-1}(0,i) = 0, n \geq 1, 1 \leq i \leq s, \tag{2.15}$$
$$-(\lambda+\theta)p_n(0,i) + \mu p_n(1,i+1) + \lambda p_{n-1}(0,i) = 0, n \geq 1, s+1 \leq i \leq S-1, \tag{2.16}$$
$$-(\lambda+\theta)p_n(0,S) + \gamma \sum_{i=0}^{s} p_n(0,i) + \lambda p_{n-1}(0,S) = 0, n \geq 1, \tag{2.17}$$
$$-(\gamma+\mu)p_n(1,i) + \theta p_{n+1}(0,i) = 0, n \geq 0, 1 \leq i \leq s, \tag{2.18}$$
$$-\mu p_n(1,i) + \theta p_{n+1}(0,i) = 0, n \geq 0, s+1 \leq i \leq S-1, \tag{2.19}$$
$$-\mu p_n(1,S) + \gamma \sum_{i=1}^{s} p_n(1,i) + \theta p_{n+1}(0,S) = 0, n \geq 0. \tag{2.20}$$

Solving equations (2.11) to (2.20). We have the following:

Theorem 2.1. *Under the condition $\lambda < \theta$, we get*

$$p_n(k,i) = \begin{cases} \left(1-\frac{\lambda}{\theta}\right)\left(\frac{\lambda}{\theta}\right)^n \psi(k,i),\ k=0, 0 \leq i \leq S, \\ \left(1-\frac{\lambda}{\theta}\right)\left(\frac{\lambda}{\theta}\right)^n \psi(k,i),\ k=1, 1 \leq i \leq S \end{cases} \tag{2.21}$$

where

$$\psi(k,i) = \begin{cases} \frac{\gamma}{\lambda}\left(\frac{\gamma+\mu}{\mu}\right)^i \left(\frac{\gamma+\lambda}{\lambda}\right)^{i-1} \psi(0,0), & k=0, 1 \leq i \leq s, \\ \frac{\gamma}{\lambda}\left(\frac{\gamma+\mu}{\mu}\right)^s \left(\frac{\gamma+\lambda}{\lambda}\right)^{s} \psi(0,0), & k=0, s+1 \leq i \leq S-1, \\ \left[\frac{\gamma}{\lambda}\left(\frac{\gamma+\mu}{\mu}\right)^s \left(\frac{\gamma+\lambda}{\lambda}\right)^{s} + \frac{\gamma}{\gamma+\lambda+\mu}\left[1-\left(\frac{\gamma+\mu}{\mu}\right)^s\left(\frac{\gamma+\lambda}{\lambda}\right)^s\right]\right]\psi(0,0), & k=0, i=S, \\ \frac{\gamma}{\mu}\left(\frac{\gamma+\mu}{\mu}\right)^{i-1} \left(\frac{\gamma+\lambda}{\lambda}\right)^{i-1} \psi(0,0), & k=1, 1 \leq i \leq s, \\ \frac{\gamma}{\mu}\left(\frac{\gamma+\mu}{\mu}\right)^s \left(\frac{\gamma+\lambda}{\lambda}\right)^{s} \psi(0,0), & k=1, s+1 \leq i \leq S \end{cases} \tag{2.22}$$

with

$$\psi(0,0)=\left\{\left(\frac{\gamma+\mu}{\mu}\right)^{s}\left(\frac{\gamma+\lambda}{\lambda}\right)^{s}(\lambda+\mu)\left[\frac{1}{\gamma+\lambda+\mu}+(S-s)\frac{\gamma}{\lambda\mu}\right]+\frac{\gamma}{\gamma+\lambda+\mu}\right\}^{-1}. \tag{2.23}$$

That equation (2.21) is the solution to (2.11) to (2.20) can also be verified by substituting the $p_n(k,i)$ values in the infinitesimal generator

$$\mathcal{Q}=\begin{bmatrix} A_{00} & A_0 & & & \\ A_2 & A_1 & A_0 & & \\ & A_2 & A_1 & A_0 & \\ & & \ddots & \ddots & \ddots \end{bmatrix} \tag{2.24}$$

such that $\mathbf{p}\mathcal{Q}=0$ and $\mathbf{pe}=1$. Each matrix A_{00}, A_0, A_1, A_2 is a square matrix of order $(2S+1)$.

Entries of A_0 are given in (i); then of A_2 are given in (ii) and that in $A_{0,0}$ and A_1 correspond to infinitesimal rates given by (iii) and (iv). In addition diagonal entries in A_{00} and A_1 are non-positive, having numerical value equal to the sum of other elements of the same row found in A_{00}, A_0, A_1 and A_2. All other transitions have rate zero.

2.1. Stability condition

We examine the requirement for system stability. Define

$$A(=A_0+A_1+A_2)=\begin{bmatrix}\mathcal{A}^{(00)} & \mathcal{A}^{(01)}\\ \mathcal{A}^{(10)} & \mathcal{A}^{(11)}\end{bmatrix}$$

where

$$\mathcal{A}^{(00)}=\begin{bmatrix} -\gamma & & & & & & \gamma\\ & -(\gamma+\theta) & & & & & \gamma\\ & & \ddots & & & & \vdots\\ & & & -(\gamma+\theta) & & & \gamma\\ & & & & -\theta & & \\ & & & & & \ddots & \\ & & & & & & -\theta \end{bmatrix}_{S+1\times S+1},$$

$$\mathcal{A}^{(11)}=\begin{bmatrix} -(\gamma+\mu) & & & & & \gamma\\ & \ddots & & & & \vdots\\ & & -(\gamma+\mu) & & & \gamma\\ & & & -\mu & & \\ & & & & \ddots & \\ & & & & & -\mu \end{bmatrix}_{S\times S},$$

$$\mathcal{A}^{(01)}=\begin{bmatrix}\theta & & \\ & \ddots & \\ & & \theta\end{bmatrix}_{(S+1)\times S}, \mathcal{A}^{(10)}=\begin{bmatrix}\mu & & \\ & \ddots & \\ & & \mu\end{bmatrix}_{S\times(S+1)}.$$

A is the infinitesimal generator of the finite state $CTMC$ $\Omega'=\{(C(t),I(t)),t\geq 0\}$ corresponding to the state space $\{(0,i),0\leq i\leq S\}\bigcup\{(1,i),1\leq i\leq S\}$. Let $\boldsymbol{\pi}=(\boldsymbol{\pi}_0,\boldsymbol{\pi}_1)$ be the steady state probability vector of A where

$$\boldsymbol{\pi}_0=(\pi(0,0),\pi(0,1),...,\pi(0,S)) \text{ and } \boldsymbol{\pi}_1=(\pi(1,1),...,\pi(1,S)).$$

Then

$$\boldsymbol{\pi} A = 0, \quad \boldsymbol{\pi}\mathbf{e} = 1. \tag{2.25}$$

From (2.25) we have

$$\pi(k,i) = \begin{cases} \frac{\gamma}{\theta}\left(\frac{\gamma+\mu}{\mu}\right)^{i}\left(\frac{\gamma+\theta}{\theta}\right)^{i-1}\pi(0,0), & k=0, 1\le i\le s,\\ \frac{\gamma}{\theta}\left(\frac{\gamma+\mu}{\mu}\right)^{s}\left(\frac{\gamma+\theta}{\theta}\right)^{s}\pi(0,0), & k=0, s+1\le i\le S-1,\\ \left[\frac{\gamma}{\theta}\left(\frac{\gamma+\mu}{\mu}\right)^{s}\left(\frac{\gamma+\theta}{\theta}\right)^{s}+\frac{\gamma}{\gamma+\theta+\mu}\left[1-\left(\frac{\gamma+\mu}{\mu}\right)^{s}\left(\frac{\gamma+\theta}{\theta}\right)^{s}\right]\right]\pi(0,0), & k=0, i=S,\\ \frac{\gamma}{\mu}\left(\frac{\gamma+\mu}{\mu}\right)^{i-1}\left(\frac{\gamma+\theta}{\theta}\right)^{i-1}\pi(0,0), & k=1, 1\le i\le s,\\ \frac{\gamma}{\mu}\left(\frac{\gamma+\mu}{\mu}\right)^{s}\left(\frac{\gamma+\theta}{\theta}\right)^{s}\pi(0,0), & k=1, s+1\le i\le S \end{cases} \tag{2.26}$$

where

$$\pi(0,0) = \left\{\left(\frac{\gamma+\mu}{\mu}\right)^{s}\left(\frac{\gamma+\theta}{\theta}\right)^{s}(\theta+\mu)\left[\frac{1}{\gamma+\theta+\mu}+(S-s)\frac{\gamma}{\theta\mu}\right]+\frac{\gamma}{\gamma+\theta+\mu}\right\}^{-1}. \tag{2.27}$$

The Markov chain is stable if and only if (see Neuts [12]) the left drift rate exceeds the right drift rate. That is,

$$\boldsymbol{\pi} A_0\mathbf{e} < \boldsymbol{\pi} A_2\mathbf{e}. \tag{2.28}$$

From the matrices A_0, A_2 we have $\boldsymbol{\pi} A_0\mathbf{e} = \lambda\sum_{i=1}^{s}\pi(0,i)$ and $\boldsymbol{\pi} A_2\mathbf{e} = \theta\sum_{i=1}^{s}\pi(0,i)$. Using relation (2.28) we obtain the stability condition as $\lambda<\theta$. Thus we have the following lemma.

Lemma 2.2. *The system under study is stable if and only if* $\lambda<\theta$.

3. Steady state analysis

For finding the steady state probability vector of the *CTMC* Ω, we first consider the system with arriving customers instantaneously taken for service whenever the server is idle with positive inventory. This has infinitesimal generator given by

$$\tilde{A} = \begin{bmatrix} \mathcal{A}'^{(00)} & \mathcal{A}'^{(01)} \\ \mathcal{A}^{(10)} & \mathcal{A}^{(11)} \end{bmatrix}$$

where

$$\mathcal{A}'^{(00)} = \begin{bmatrix} -\gamma & & & & & & \gamma \\ & -(\gamma+\lambda) & & & & & \gamma \\ & & \ddots & & & & \vdots \\ & & & -(\gamma+\lambda) & & & \gamma \\ & & & & -\lambda & & \\ & & & & & \ddots & \\ & & & & & & -\lambda \end{bmatrix}, \mathcal{A}'^{(01)} = \begin{bmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{bmatrix}$$

Let $\boldsymbol{\psi} = (\boldsymbol{\psi}_0, \boldsymbol{\psi}_1)$ be the steady state vector of $\tilde{A}$. Then $\boldsymbol{\psi}$ satisfies the equations

$$\boldsymbol{\psi}\tilde{A} = 0, \quad \boldsymbol{\psi}\mathbf{e} = 1. \tag{3.1}$$

From $\boldsymbol{\psi}\tilde{A} = 0$ we have

$$\begin{aligned} -\gamma\psi(0,0) + \mu\psi(1,1) &= 0, \\ -(\gamma+\lambda)\psi(0,i) + \mu\psi(1,i+1) &= 0, 1 \le i \le s \\ -\lambda\psi(0,i) + \mu\psi(1,i+1) &= 0, s+1 \le i \le S-1 \\ \gamma\sum_{i=0}^{s}\psi(0,i) - \lambda\psi(1,S) &= 0 \\ \lambda\psi(0,i) - (\gamma+\mu)\psi(1,i) &= 0, 1 \le i \le s \\ \lambda\psi(0,i) - \mu\psi(1,i) &= 0, s+1 \le i \le S-1 \\ \lambda\psi(0,S) + \gamma\sum_{i=1}^{s}\psi(1,i) - \mu\psi(1,S) &= 0 \end{aligned}$$

and $\boldsymbol{\psi}_k$ can be obtained as

$$\psi(k,i) = \begin{cases} \frac{\gamma}{\lambda}\left(\frac{\gamma+\mu}{\mu}\right)^{i}\left(\frac{\gamma+\lambda}{\lambda}\right)^{i-1}\psi(0,0), & k=0, 1 \le i \le s, \\ \frac{\gamma}{\lambda}\left(\frac{\gamma+\mu}{\mu}\right)^{s}\left(\frac{\gamma+\lambda}{\lambda}\right)^{s}\psi(0,0), & k=0, s+1 \le i \le S-1, \\ \left[\frac{\gamma}{\lambda}\left(\frac{\gamma+\mu}{\mu}\right)^{s}\left(\frac{\gamma+\lambda}{\lambda}\right)^{s} + \frac{\gamma}{\gamma+\lambda+\mu}\left[1-\left(\frac{\gamma+\mu}{\mu}\right)^{s}\left(\frac{\gamma+\lambda}{\lambda}\right)^{s}\right]\right]\psi(0,0), & k=0, i=S, \\ \frac{\gamma}{\mu}\left(\frac{\gamma+\mu}{\mu}\right)^{i-1}\left(\frac{\gamma+\lambda}{\lambda}\right)^{i-1}\psi(0,0), & k=1, 1 \le i \le s, \\ \frac{\gamma}{\mu}\left(\frac{\gamma+\mu}{\mu}\right)^{s}\left(\frac{\gamma+\lambda}{\lambda}\right)^{s}\psi(0,0), & k=1, s+1 \le i \le S. \end{cases} \tag{3.2}$$

The unknown probability $\psi(0,0)$ can be found from the normalizing condition

$$\psi(0,0) = \left\{\left(\frac{\gamma+\mu}{\mu}\right)^{s}\left(\frac{\gamma+\lambda}{\lambda}\right)^{s}(\lambda+\mu)\left[\frac{1}{\gamma+\lambda+\mu} + (S-s)\frac{\gamma}{\lambda\mu}\right] + \frac{\gamma}{\gamma+\lambda+\mu}\right\}^{-1}. \tag{3.3}$$

Now using the vector $\boldsymbol{\psi}$, we proceed to compute the steady state probability vector of the original system. Let $\mathbf{p}$ be the steady state probability vector of the generator $\mathcal{Q}$. Then $\mathbf{p}$ must satisfy the set of equations

$$\mathbf{p}\mathcal{Q} = 0, \qquad \mathbf{pe} = 1. \tag{3.4}$$

Partition $\mathbf{p}$ as $\mathbf{p} = (\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2, \ldots)$. Then the above system of equations reduces to:

$$\mathbf{p}_0 A_{00} + \mathbf{p}_1 A_2 = 0, \tag{3.5}$$

$$\mathbf{p}_{n-1}A_0 + \mathbf{p}_n A_1 + \mathbf{p}_{n+1}A_2 = 0, \quad n \ge 1. \tag{3.6}$$

We assume a solution of the form

$$\mathbf{p}_n = \mathcal{K}\left(\frac{\lambda}{\theta}\right)^{n}\boldsymbol{\psi} \text{ for } n \ge 0 \tag{3.7}$$

where $\mathcal{K}$ is a constant to be determined. From (3.5), we have

$$\mathbf{p}_0 A_{00} + \mathbf{p}_1 A_2 = \mathcal{K}\boldsymbol{\psi}\left(A_{00} + \frac{\lambda}{\theta}A_2\right) = \mathcal{K}\boldsymbol{\psi}\tilde{A} = 0 \tag{3.8}$$

and from relation (3.6), we have

$$\begin{aligned} \mathbf{p}_{n-1}A_0 + \mathbf{p}_n A_1 + \mathbf{p}_{n+1}A_2 &= \mathcal{K}\left(\frac{\lambda}{\theta}\right)^{n}\boldsymbol{\psi}\left(\frac{\theta}{\lambda}A_0 + A_1 + \frac{\lambda}{\theta}A_2\right) \\ &= \mathcal{K}\left(\frac{\lambda}{\theta}\right)^{n}\boldsymbol{\psi}\left(\frac{\theta}{\lambda}A_0 + A_{00} - \frac{\theta}{\lambda}A_0 + \frac{\lambda}{\theta}A_2\right) \\ &= \mathcal{K}\left(\frac{\lambda}{\theta}\right)^{n}\boldsymbol{\psi}\tilde{A} = 0. \end{aligned} \tag{3.9}$$

Thus (3.7) satisfies (3.5) and (3.6). Now applying the normalizing condition $\mathbf{pe} = 1$, we get

$$\mathcal{K}\boldsymbol{\psi}\left[1+\left(\frac{\lambda}{\theta}\right)+\left(\frac{\lambda}{\theta}\right)^2+\ldots\right]\mathbf{e}=1.$$

Hence under the condition that $\lambda < \theta$, we have

$$\mathcal{K} = 1 - \frac{\lambda}{\theta}. \tag{3.10}$$

Thus we arrive at our main result:

Theorem 3.1. *Under the necessary and sufficient condition $\lambda < \theta$ for stability, the components of the steady-state probability vector of the CTMC Ω, with generator $\mathcal{Q}$, is given by (3.7) and (3.10). That is,*

$$\boldsymbol{p}_n = \left(1-\frac{\lambda}{\theta}\right)\left(\frac{\lambda}{\theta}\right)^n \boldsymbol{\psi} \text{ for } n \geq 0. \tag{3.11}$$

Now suppose that the material for service is abundantly available. In this case we ignore the inventory status and examine only the server status and the number of customers in the orbit. As in Neuts and Rao [13], we assume that primary customers do not access the server directly; instead they first join an orbit of infinite capacity, from where they access the server according to the FIFO discipline. Primary customers do not join orbit when server is busy. On retrial by the head of orbital queue, if the server is found busy, the customer returns to orbit. With these assumptions we get the following important corollary for the retrial queue under consideration.

Corollary 3.2. *For the retrial queue under consideration we deduce from (3.11) the following system state distribution:*

$$Pr.(i\text{customers in the orbit and server is busy}) = \left(1-\frac{\lambda}{\theta}\right)\left(\frac{\lambda}{\theta}\right)^i\left(\frac{\lambda}{\lambda+\mu}\right) \text{ for } i \geq 0$$

and

$$Pr.(i \text{ customers in the orbit and server is idle}) = \left(1-\frac{\lambda}{\theta}\right)\left(\frac{\lambda}{\theta}\right)^i\left(\frac{\mu}{\lambda+\mu}\right) \text{ for } i \geq 0.$$

Proof. Consider the classical retrial queue where all primary customers must first join the orbit, from where, through retrial alone, they can access the server. If the server is busy at the time an external arrival takes place, then that external customer does not join the system. Also we assume the constant retrial rate and exponentially distributed service time. Then $\{(N(t), C(t)), t \geq 0\}$ forms a $CTMC$ with state space $\{(n, i), n \geq 0, i = 0, 1\}$. The infinitesimal generator $\tilde{\mathcal{Q}}$ of this system is similar to $M/M/1$ retrial inventory system (see (2.24)) but with

$$A_{00} = \begin{bmatrix} -\lambda & 0 \\ \mu & -\mu \end{bmatrix}, A_0 = \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix}, A_1 = \begin{bmatrix} -(\lambda+\theta) & 0 \\ \mu & -\mu \end{bmatrix}, A_2 = \begin{bmatrix} 0 & \theta \\ 0 & 0 \end{bmatrix}.$$

The inventory component is taken out from the phase of the process. Let $\tilde{\boldsymbol{\pi}} = (\tilde{\pi}_0, \tilde{\pi}_1)$ be the steady-state probability vector of the generator $\mathcal{A} = A_0 + A_1 + A_2$. Then $\tilde{\boldsymbol{\pi}}\mathcal{A} = 0, \tilde{\boldsymbol{\pi}}\mathbf{e} = 1$.

The steady-state equation $\tilde{\boldsymbol{\pi}}\mathcal{A} = 0$ reduce to $-\theta\tilde{\pi}_0 + \mu\tilde{\pi}_1 = 0$ subject to the normalizing condition $\tilde{\pi}_0 + \tilde{\pi}_1 = 1$ we have

$$\tilde{\pi}_i = \begin{cases} \dfrac{\mu}{\theta+\mu} \text{ for } i = 0 \\[2ex] \dfrac{\theta}{\theta+\mu} \text{ for } i = 1 \end{cases}$$

The system under study with the generator given in (2.24) is stable (see Neuts [12]) if and only if $\tilde{\boldsymbol{\pi}} A_0 \mathbf{e} < \tilde{\boldsymbol{\pi}} A_2 \mathbf{e} \Rightarrow \lambda < \theta$. Let $\mathbf{q} = (\mathbf{q}_0, \mathbf{q}_1, \mathbf{q}_2, \ldots)$ be the steady-state probability vector of $\tilde{\mathcal{Q}}$. That is, $\mathbf{q}$ satisfies

$$\mathbf{q}\tilde{\mathcal{Q}} = 0, \quad \mathbf{qe} = 1. \tag{3.12}$$

Each $\mathbf{q}_i = (q_i(0), q_i(1))$ for $i \geq 0$. From $\mathbf{q}\tilde{\mathcal{Q}} = 0$ we have

$$\begin{aligned} -\lambda q_0(0) + \mu q_0(1) &= 0, \\ -\mu q_0(1) + \theta q_1(0) &= 0, \\ \lambda q_{i-1}(0) - (\lambda + \theta) q_i(0) + \mu q_i(1) &= 0, \quad i \geq 1 \\ -\mu q_i(1) + \theta q_{i+1}(0) &= 0. \end{aligned} \tag{3.13}$$

The solution to these is subject to the normalizing condition

$$\sum_{i=0}^{\infty} (q_i(0) + q_i(1)) = 1.$$

Solving these set of equations we get

$$q_i(j) = \begin{cases} \left(1 - \frac{\lambda}{\theta}\right) \left(\frac{\lambda}{\theta}\right)^i \left(\frac{\mu}{\lambda + \mu}\right) \text{ for } i \geq 0, j = 0 \\ \left(1 - \frac{\lambda}{\theta}\right) \left(\frac{\lambda}{\theta}\right)^i \left(\frac{\lambda}{\lambda + \mu}\right) \text{ for } i \geq 0, j = 1 \end{cases} \tag{3.14}$$

which is geometric distribution, except for the multiplicative factor $\frac{\mu}{\lambda + \mu}$ and $\frac{\lambda}{\lambda + \mu}$. □

We extend the above corollary (Cor. 3.2) to the case when retrial rate is linear. This mean that the retrial rate is $n\theta$ when n customers are in the orbit (it may be FIFO discipline or all customers trying to access the server). Then we have the following result.

Theorem 3.3. *In the case of linear retrial rate of customers in orbit for our queueing model, the long run system state probability is given by*

$$\frac{1}{n!} \left(\frac{\lambda}{\theta}\right)^n e^{-\lambda/\theta} \left(1 + \frac{\lambda}{\mu}\right)^{-1} \text{ for } n \text{ customers in the orbit with server idle}$$

and

$$\frac{1}{n!} \frac{\lambda}{\mu} \left(\frac{\lambda}{\theta}\right)^n e^{-\lambda/\theta} \left(1 + \frac{\lambda}{\mu}\right)^{-1} \text{ for } n \text{ customers in the orbit with server busy.}$$

4. Performance measures

- Probability mass function of the number of customers in the orbit = Pr[n customers in the orbit] is given by $\mathcal{P}_n = \mathbf{p}_n \mathbf{e} = \left(1 - \frac{\lambda}{\theta}\right) \left(\frac{\lambda}{\theta}\right)^n, \quad n \geq 0.$
- Probability that the server is idle

$$\begin{aligned} P_0 &= \sum_{n=0}^{\infty} \sum_{j=0}^{S} p_n(0, j) \\ &= \left\{ \frac{\gamma + \lambda}{\lambda} \left[1 + \frac{\gamma + \mu}{\gamma + \mu + \lambda} \left(\left(\frac{\gamma + \mu}{\lambda}\right)^s \left(\frac{\gamma + \lambda}{\mu}\right)^s - 1\right)\right] \right. \\ &\quad \left. + \frac{\gamma}{\lambda} \left(\frac{\gamma + \mu}{\lambda}\right)^s \left(\frac{\gamma + \lambda}{\mu}\right)^s (S - s - 1) \right\} \psi(0, 0). \end{aligned}$$

- Probability that the server is busy

$$\begin{aligned} P_1 &= \sum_{n=0}^{\infty}\sum_{j=1}^{S} p_n(1,j) \\ &= \left\{ \frac{\lambda}{\lambda+\mu+\gamma} \left[\left(\frac{\gamma+\mu}{\lambda}\right)^s \left(\frac{\gamma+\lambda}{\mu}\right)^s - 1 \right] + \frac{\gamma}{\mu} \left(\frac{\gamma+\mu}{\lambda}\right)^s \left(\frac{\gamma+\lambda}{\mu}\right)^s (S-s) \right\} \psi(0,0). \end{aligned}$$

- Mean number of customers in the retrial orbit

$$\begin{aligned} N_O &= \sum_{n=1}^{\infty} n \left[\sum_{j=0}^{S} p_n(0,j) + \sum_{j=1}^{S} p_n(1,j) \right] \\ &= \frac{\lambda}{\theta-\lambda} \left\{ \gamma \left(\frac{\gamma+\mu}{\lambda}\right)^s \left(\frac{\gamma+\lambda}{\mu}\right)^s \left[\frac{(S-s-1)}{\lambda} + \frac{(S-s)}{\mu} \right] + \frac{\gamma+\lambda}{\lambda} \right. \\ &\quad \left. + \left[\left(\frac{\gamma+\mu}{\lambda}\right)^s \left(\frac{\gamma+\lambda}{\mu}\right)^s - 1 \right] \left[\frac{(\gamma+\lambda)(\gamma+\mu)}{\lambda(\gamma+\lambda+\mu)} + \frac{\lambda}{\gamma+\lambda+\mu} \right] \right\} \psi(0,0). \end{aligned}$$

- Mean number of customers in the system $N_S = N_O + P_1$.
- Expected loss rate

$$\begin{aligned} E_L &= \lambda \sum_{n=0}^{\infty} \left[p_n(0,0) + \sum_{j=1}^{S} p_n(1,j) \right] \\ &= \lambda \left\{ 1 + \frac{\lambda}{\lambda+\mu+\gamma} \left[\left(\frac{\gamma+\mu}{\lambda}\right)^s \left(\frac{\gamma+\lambda}{\mu}\right)^s - 1 \right] \right. \\ &\quad \left. + \frac{\gamma}{\mu} \left(\frac{\gamma+\mu}{\lambda}\right)^s \left(\frac{\gamma+\lambda}{\mu}\right)^s (S-s) \right\} \psi(0,0). \end{aligned}$$

- Successful rate of retrial

$$\begin{aligned} E_{SR} &= \theta \sum_{n=1}^{\infty}\sum_{j=1}^{S} p_n(0,j) \\ &= \frac{\theta\lambda}{\theta-\lambda} \left\{ \frac{\gamma+\mu}{\gamma+\mu+\lambda} \left[\left(\frac{\gamma+\mu}{\lambda}\right)^s \left(\frac{\gamma+\lambda}{\mu}\right)^s - 1 \right] \right. \\ &\quad \left. + \frac{\gamma}{\lambda} \left(\frac{\gamma+\mu}{\lambda}\right)^s \left(\frac{\gamma+\lambda}{\mu}\right)^s (S-s-1) \right\} \psi(0,0). \end{aligned}$$

- Expected number of items in the inventory $E_I = \sum_{n=0}^{\infty}\sum_{j=1}^{S} j[p_n(0,j) + p_n(1,j)]$.
- Expected reorder rate $E_R = \mu \sum_{n=0}^{\infty} p_n(1, s+1) = \gamma \left(\frac{\gamma+\mu}{\lambda}\right)^s \left(\frac{\gamma+\lambda}{\mu}\right)^s \psi(0,0)$.
- Expected replenishment rate $E_{\mathrm{RR}} = \gamma \sum_{n=0}^{\infty} \left[\sum_{j=0}^{s} p_n(0,j) + \sum_{j=1}^{s} p_n(1,j) \right]$.
- Probability that the inventory level is zero is given by

$$Pr[I=0] = \sum_{n=0}^{\infty} p_n(0,0) = \psi(0,0).$$

- Probability that the inventory level is greater above safety stock is given by

$$\begin{aligned} Pr[I > s] &= \sum_{n=0}^{\infty}\sum_{j=s+1}^{S}[p_n(0,j)+p_n(1,j)] \\ &= \gamma\left(\frac{\gamma+\mu}{\lambda}\right)^s\left(\frac{\gamma+\lambda}{\mu}\right)^s\left[\frac{S-s}{\mu}+\frac{S-s-1}{\lambda}\right]\psi(0,0). \end{aligned}$$

4.1. Expected waiting time of a customer in the orbit

For computing expected waiting time of a tagged customer in the orbit in the r^{th} position $r > 0$. Consider the Markov chain $\{(N'(t), C(t), I(t)), t \geq 0\}$ where $N'(t)$ is the rank of the tagged customer, $C(t)$ is the status of server and $I(t)$ is the number of items in the inventory. We arrange the state space as $\{r, r-1, \ldots, 1\} \times (\{(0,0), \ldots (0,S)\} \bigcup \{(1,1), \ldots (1,S)\}) \bigcup \{\Delta\}$ where $\{\Delta\}$ is the absorbing state denoting that the tagged customer is selected for service. Thus the infinitesimal generator $\tilde{W}$ is of the form

$$\tilde{W} = \begin{bmatrix} T_r & T_r^{\mathbf{0}} \\ \mathbf{0} & 0 \end{bmatrix}$$

where

$$T_r = \begin{bmatrix} \mathcal{B}_1 & \mathcal{B}_2 & & \\ & \ddots & \ddots & \\ & & \mathcal{B}_1 & \mathcal{B}_2 \\ & & & \mathcal{B}_1 \end{bmatrix}, T_r^{\mathbf{0}} = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathcal{B} \end{bmatrix}$$

with

$$\mathcal{B}_1 = \begin{bmatrix} B_{00} & O \\ B_{10} & B_{11} \end{bmatrix}, \mathcal{B}_2 = \begin{bmatrix} O & N \\ O & O \end{bmatrix}, \mathcal{B} = \begin{bmatrix} N\mathbf{e} \\ \mathbf{0} \end{bmatrix}$$

$$B_{00} = \begin{bmatrix} -\gamma & & & & & & \gamma \\ & -(\gamma+\theta) & & & & & \gamma \\ & & \ddots & & & & \vdots \\ & & & -(\gamma+\theta) & & & \gamma \\ & & & & -\theta & & \\ & & & & & \ddots & \\ & & & & & & -\theta \end{bmatrix}, N = \begin{bmatrix} \theta & & \\ & \ddots & \\ & & \theta \end{bmatrix}$$

$$B_{11} = \begin{bmatrix} -(\gamma+\mu) & & & & & \gamma \\ & \ddots & & & & \vdots \\ & & -(\gamma+\mu) & & & \gamma \\ & & & -\mu & & \\ & & & & \ddots & \\ & & & & & -\mu \end{bmatrix}, B_{10} = \begin{bmatrix} \mu & & \\ & \ddots & \\ & & \mu \end{bmatrix}.$$

Now, the waiting time of r^{th} customer in the orbit is the time until absorption of the Markov chain $\{(N'(t), C(t), I(t)), t \geq 0\}$. Thus the expected waiting time of the r^{th} customer in the orbit is given by

$$E_W = -\alpha_r T_r^{-1}\mathbf{e}$$

where $\alpha_r = (\boldsymbol{\eta}, \mathbf{0}, \mathbf{0}, \ldots, \mathbf{0})$ is the initial probability vector with
$\boldsymbol{\eta} = (\psi(0,0), \ldots, \psi(0,S), \psi(1,1), \ldots, \psi(1,S))$.

4.2. Some important system characteristics

Under condition of stability, given $\epsilon > 0$ there exists a K sufficiently large such that

$$\sum_{n=0}^{K}\left(1-\frac{\lambda}{\theta}\right)\left(\frac{\lambda}{\theta}\right)^{n} > 1-\epsilon$$

Except for heavy traffic, the above approximation is very close to exact value.

In this section we consider some important measures such as distribution of number of customers served, number of arrivals, number of customers lost during an interval of random duration. For this first we choose a random clock of duration T which is exponentially distributed with parameter δ.

4.2.1. Distribution of the number of customers served during an interval of duration T

We have the following lemma, about the expected number of customers served during an interval.

Lemma 4.1. *Expected number of customers served during an interval of duration T is given by* $\sum_{m=0}^{\infty} m\boldsymbol{x}_m\boldsymbol{e}$.

Proof. To calculate the expected number of customers served during an interval, we consider the Markov Chain $\{(M(t), N(t), C(t), I(t)), t \geq 0\}$ where $M(t)$ represents the number of customers served in $(0, t]$ which is contained in the random duration T, $N(t)$ is the number of customers in the orbit, $C(t)$ is the status of server and $I(t)$ is the number of items in the inventory at time t with state space $\{\Delta\}\bigcup\{(m, n, 0, i), m \geq 0, 0 \leq n \leq K, 0 \leq i \leq S\}\bigcup\{(m, n, 1, i), m \geq 0, 0 \leq n \leq K, 1 \leq i \leq S\}$ where $\{\Delta\}$ is the absorbing state denoting the realization of random clock. The infinitesimal generator is of the form

$$\mathcal{N}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ \delta\mathbf{e} & N_1 & N_0 & & \\ \delta\mathbf{e} & & N_1 & N_0 & \\ \vdots & & & \ddots & \ddots \end{pmatrix}.$$

Thus the transition rates are

$$\begin{array}{ll}
(m,n,1,i) \xrightarrow{\mu} (m+1,n,0,i-1) & \text{for } m \geq 0, 0 \leq n \leq K, 1 \leq i \leq S \\
(m,n,0,i) \xrightarrow{\lambda} (m,n+1,0,i) & \text{for } m \geq 0, 0 \leq n \leq K-1, 1 \leq i \leq S \\
(m,n,0,i) \xrightarrow{\theta} (m,n-1,1,i) & \text{for } m \geq 0, 1 \leq n \leq K, 1 \leq i \leq S \\
(m,n,0,i) \xrightarrow{\gamma} (m,n,0,S) & \text{for } m \geq 0, 0 \leq n \leq K, 0 \leq i \leq s \\
(m,n,1,i) \xrightarrow{\gamma} (m,n,1,S) & \text{for } m \geq 0, 0 \leq n \leq K, 1 \leq i \leq s \\
(m,n,0,i) \xrightarrow{\delta} \{\Delta\} & \text{for } m \geq 0, 0 \leq n \leq K, 0 \leq i \leq S \\
(m,n,1,i) \xrightarrow{\delta} \{\Delta\} & \text{for } m \geq 0, 0 \leq n \leq K, 1 \leq i \leq S
\end{array}$$

Let $\mathbf{x}_m$ be the probability that the number of customers served before realization of the random clock is m. Then $\mathbf{x}_m$ is given by

$$\mathbf{x}_0 = -\delta\boldsymbol{\beta}N_1^{-1}\mathbf{e}, \tag{4.1}$$

and

$$\mathbf{x}_m = (-1)^m\delta\boldsymbol{\beta}\left(N_0N_1^{-1}\right)^m N_1^{-1}\mathbf{e},\ \text{for } m \geq 1 \tag{4.2}$$

where $\boldsymbol{\beta}$ is the initial probability vector is of the form $(\boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_K)$ with

$$\boldsymbol{\beta}_n = \frac{\left(1-\frac{\lambda}{\theta}\right)\left(\frac{\lambda}{\theta}\right)^n(\psi(0,0),\ldots,\psi(0,S),\psi(1,1),\ldots,\psi(1,S))}{1-\left(\frac{\lambda}{\theta}\right)^{K+1}}. \tag{4.3}$$

Therefore, the expected number of customers served during an interval of duration T is $\sum_{m=0}^{\infty} m\mathbf{x}_m\mathbf{e}$.

□

4.2.2. Distribution of the number of arrivals during an interval of duration T

To get the distribution of number of arrivals, we consider the Markov chain $\{(M(t), N(t), C(t), I(t)), t \geq 0\}$ where $M(t)$ is the number of primary customers arriving in $(0, t]$ which is contained in the random duration T, $N(t)$ is the number of customers in the orbit, $C(t)$ is the status of server and $I(t)$ is the number of items in the inventory at time t with state space $\{\Delta\} \bigcup \{(m, n, 0, i), m \geq 0, 0 \leq n \leq K, 0 \leq i \leq S\} \bigcup \{(m, n, 1, i), m \geq 0, 0 \leq n \leq K, 1 \leq i \leq S\}$ where $\{\Delta\}$ is the absorbing state denoting the realization of random clock. The infinitesimal generator is of the form

$$\mathcal{N}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ \delta\mathbf{e} & M_1 & M_0 & & \\ \delta\mathbf{e} & & M_1 & M_0 & \\ \vdots & & & \ddots & \ddots \end{pmatrix}.$$

Thus the transition rates are

$$\begin{array}{ll} (m,n,0,i) \xrightarrow{\lambda} (m+1,n+1,0,i) & \text{for } m \geq 0, 0 \leq n \leq K-1, 1 \leq i \leq S \\ (m,n,1,i) \xrightarrow{\mu} (m,n,0,i-1) & \text{for } m \geq 0, 0 \leq n \leq K, 1 \leq i \leq S \\ (m,n,0,i) \xrightarrow{\theta} (m,n-1,1,i) & \text{for } m \geq 0, 1 \leq n \leq K, 1 \leq i \leq S \\ (m,n,0,i) \xrightarrow{\gamma} (m,n,0,S) & \text{for } m \geq 0, 0 \leq n \leq K, 0 \leq i \leq s \\ (m,n,1,i) \xrightarrow{\gamma} (m,n,1,S) & \text{for } m \geq 0, 0 \leq n \leq K, 1 \leq i \leq s \\ (m,n,0,i) \xrightarrow{\delta} \{\Delta\} & \text{for } m \geq 0, 0 \leq n \leq K, 0 \leq i \leq S \\ (m,n,1,i) \xrightarrow{\delta} \{\Delta\} & \text{for } m \geq 0, 0 \leq n \leq K, 1 \leq i \leq S \end{array}$$

Let $\mathbf{y}_m$ be the probability that the number of arrivals is m before realization of the random clock. Then $\mathbf{y}_m$ is given by

$$\mathbf{y}_0 = -\delta\widetilde{\boldsymbol{\beta}}M_1^{-1}\mathbf{e}, \tag{4.4}$$

and

$$\mathbf{y}_m = (-1)^m\delta\widetilde{\boldsymbol{\beta}}\left(M_0M_1^{-1}\right)^m M_1^{-1}\mathbf{e}, \text{ for } m \geq 1 \tag{4.5}$$

where $\widetilde{\boldsymbol{\beta}}$ is the initial probability vector of the form $\left(\widetilde{\boldsymbol{\beta}}_0, \widetilde{\boldsymbol{\beta}}_1, \ldots, \widetilde{\boldsymbol{\beta}}_K\right)$ with

$$\widetilde{\boldsymbol{\beta}}_n = \frac{\left(1 - \frac{\lambda}{\theta}\right)\left(\frac{\lambda}{\theta}\right)^n (\psi(0,0), \ldots, \psi(0,S), \psi(1,1), \ldots, \psi(1,S))}{1 - \left(\frac{\lambda}{\theta}\right)^{K+1}}. \tag{4.6}$$

Hence we have the following lemma

Lemma 4.2. *Expected number of arrivals during an interval of duration T is given by* $\sum_{m=1}^{\infty} m\boldsymbol{y}_m\boldsymbol{e}$.

4.2.3. Distribution of the number of customers lost during an interval of duration

To find the expected number of customers lost during an interval, we consider the Markov chain $\{(M(t), N(t), C(t), I(t)), t \geq 0\}$ where $M(t)$ is the number of customers lost in $(0, t]$ belonging to the random duration of length T, $N(t)$ is the number of customers in the orbit, $C(t)$ is the status of server and $I(t)$ is the number of items in the inventory at time t with state space $\{\Delta\} \bigcup \{(m, n, 0, i), m \geq 0, 0 \leq n \leq K, 0 \leq i \leq$

$S\} \bigcup \{(m,n,1,i), m \geq 0, 0 \leq n \leq K, 1 \leq i \leq S\}$ where $\{\Delta\}$ is the absorbing state denoting the realization of random clock. The infinitesimal generator is of the form

$$\mathcal{N}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ \delta\mathbf{e} & L_1 & L_0 & & \\ \delta\mathbf{e} & & L_1 & L_0 & \\ \vdots & & & \ddots & \ddots \end{pmatrix}.$$

Thus the transition rates are

$$\begin{array}{lll}
(m,n,0,0) \xrightarrow{\lambda} (m+1,n,0,0) & \text{for } m \geq 0, 0 \leq n \leq K \\
(m,n,1,i) \xrightarrow{\lambda} (m+1,n,1,i) & \text{for } m \geq 0, 0 \leq n \leq K, 1 \leq i \leq S \\
(m,n,0,i) \xrightarrow{\lambda} (m,n+1,0,i) & \text{for } m \geq 0, 0 \leq n \leq K-1, 1 \leq i \leq S \\
(m,n,1,i) \xrightarrow{\mu} (m,n,0,i-1) & \text{for } m \geq 0, 0 \leq n \leq K, 1 \leq i \leq S \\
(m,n,0,i) \xrightarrow{\theta} (m,n-1,1,i) & \text{for } m \geq 0, 1 \leq n \leq K, 1 \leq i \leq S \\
(m,n,0,i) \xrightarrow{\gamma} (m,n,0,S) & \text{for } m \geq 0, 0 \leq n \leq K, 0 \leq i \leq s \\
(m,n,1,i) \xrightarrow{\gamma} (m,n,1,S) & \text{for } m \geq 0, 0 \leq n \leq K, 1 \leq i \leq s \\
(m,n,0,i) \xrightarrow{\delta} \{\Delta\} & \text{for } m \geq 0, 0 \leq n \leq K, 0 \leq i \leq S \\
(m,n,1,i) \xrightarrow{\delta} \{\Delta\} & \text{for } m \geq 0, 0 \leq n \leq K, 1 \leq i \leq S
\end{array}$$

Let $\mathbf{z}_m$ be the probability that the number of customer lost is m before random clock realization. Then $\mathbf{z}_m$ is given by

$$\mathbf{z}_0 = -\delta\widehat{\boldsymbol{\beta}}L_1^{-1}\mathbf{e}, \tag{4.7}$$

and

$$\mathbf{z}_m = (-1)^m \delta\widehat{\boldsymbol{\beta}} \left(L_0 L_1^{-1}\right)^m L_1^{-1}\mathbf{e}, \text{ for } m \geq 1 \tag{4.8}$$

where $\widehat{\boldsymbol{\beta}}$ is the initial probability vector is of the form $\left(\widehat{\boldsymbol{\beta}}_0, \widehat{\boldsymbol{\beta}}_1, \ldots, \widehat{\boldsymbol{\beta}}_K\right)$ with

$$\widehat{\boldsymbol{\beta}}_n = \frac{\left(1-\frac{\lambda}{\theta}\right)\left(\frac{\lambda}{\theta}\right)^n (\psi(0,0),\ldots,\psi(0,S),\psi(1,1),\ldots,\psi(1,S))}{1-\left(\frac{\lambda}{\theta}\right)^{K+1}}. \tag{4.9}$$

Hence we have the following lemma

Lemma 4.3. *Expected number of customer lost during an interval of duration T is given by* $\sum_{m=1}^{\infty} m\boldsymbol{z}_m\boldsymbol{e}$.

4.2.4. Analysis of cycle length

Cycle length T_c is the time length from idle state with no customer in the orbit to idle state with no customer

We have the following lemma, about the expected cycle length.

Lemma 4.4. *Expected cycle length T_c is given by* $-\boldsymbol{\gamma}_c W_c^{-1}\boldsymbol{e}$.

Proof. To compute the expected cycle length, we consider the Markov Chain $\{(N(t),C(t),I(t)), t \geq 0\}$ where $N(t)$ is the number of customers in the orbit, $C(t)$ is the status of server and $I(t)$ is the number of items in the inventory at time t with state space $\{(n,0,i), 0 \leq n \leq K, 0 \leq i \leq S\} \bigcup \{(n,1,i), 0 \leq n \leq K, 1 \leq i \leq S\} \bigcup \{\Delta\}$ where $\{\Delta\}$ is the absorbing state which means the idle system with no customer in the orbit.

Let $\boldsymbol{\gamma}_c = (\boldsymbol{\gamma}_c^{(0)}, \mathbf{0}, \dots, \mathbf{0})$ where

$$\boldsymbol{\gamma}_c^{(0)} = \frac{(\psi(0,0), \dots, \psi(0,S), 0, \dots, 0)}{\sum_{i=0}^{S} \psi(0,i)}.$$

This can be taken as the initial probability vector of cycle length T_c. The first passage time from the state $(0,0)$ to $(0,0)$ follows a Phase type distribution with representation $(\boldsymbol{\gamma}_c, W_c)$. Corresponding infinitesimal generator is of the form $\begin{pmatrix} W_c & \mathbf{w}_c^0 \\ \mathbf{0} & 0 \end{pmatrix}$ where

$$W_c = \begin{pmatrix} H_1^0 & H_0 & & & \\ H_2 & H_1 & H_0 & & \\ & \ddots & \ddots & \ddots & \\ & & H_2 & H_1 & H_0 \\ & & & H_2 & H_1^{00} \end{pmatrix}, \mathbf{w}_c^0 = \begin{pmatrix} H \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

Thus the transition rates are

$$\begin{array}{lll} (n,1,i) \xrightarrow{\mu} (n,0,i-1) & \text{for } 1 \le n \le K, 1 \le i \le S \\ (n,0,i) \xrightarrow{\lambda} (n+1,0,i) & \text{for } 0 \le n \le K-1, 1 \le i \le S \\ (n,0,i) \xrightarrow{\theta} (n-1,1,i) & \text{for } 1 \le n \le K, 1 \le i \le S \\ (n,0,i) \xrightarrow{\gamma} (n,0,S) & \text{for } 0 \le n \le K, 0 \le i \le s \\ (n,1,i) \xrightarrow{\gamma} (n,1,S) & \text{for } 0 \le n \le K, 1 \le i \le s \\ (0,1,i) \xrightarrow{\mu} \Delta & \text{for } 1 \le i \le S \end{array}$$

Therefore, the expected cycle length is $-\boldsymbol{\gamma}_c W_c^{-1} \mathbf{e}$.

□

5. Numerical illustrations

In this section we provide numerical illustration of the system performance with variation in values of underlying parameters.

Effect of the arrival rate λ

In Figure 1, we investigate the effect of arrival rate for $(s, S, \mu, \gamma, \theta) = (20, 45, 17, 3, 15)$. The outcome is on expected lines.

Effect of the service rate μ

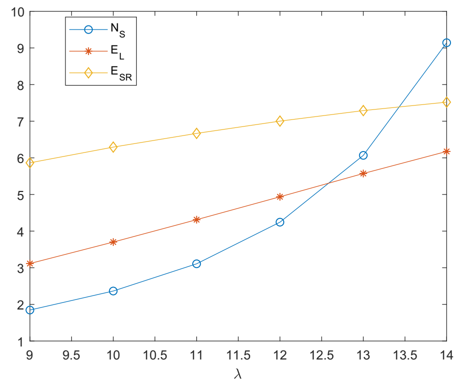
In Figure 2, we notice that as service rate μ is varied, keeping $(s, S, \lambda, \gamma, \theta) = (20, 45, 8, 2, 9)$, the probability of server being idle increases with increase in service rate. This could be attributed to the fact that, server becomes idle with faster rate of service. The trend in other measures are on expected lines.

Effect of the retrial rate θ

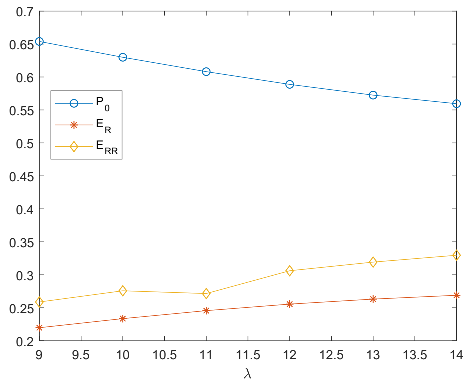
Figure 3 provides results of a numerical experiment with variation in retrial rate with $(s, S, \lambda, \gamma, \mu) = (20, 45, 7, 2, 12)$. With increase in value of θ, server idle probability decreases. Except for the vales of expected replenishment rate, the expected outcome supports reality. We are not able to give an interpretation for the behavior of E_{RR} in this case.

Effect of the replenishment rate γ

The outcome in Figure 4 also supports our guess.

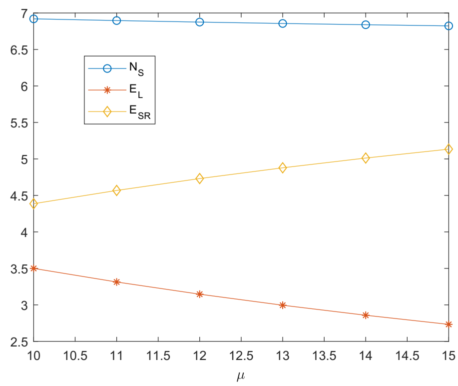


(a) Effect of λ

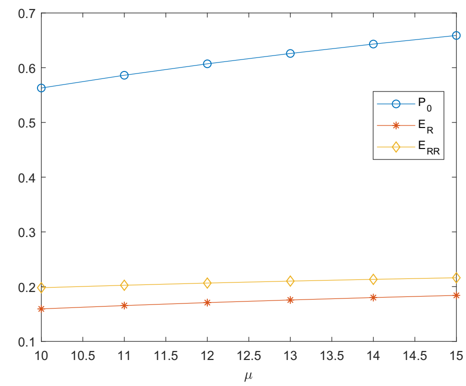


(b) Effect of λ

FIGURE 1. Effect of λ for $(s, S, \mu, \gamma, \theta) = (20, 45, 17, 3, 15)$.

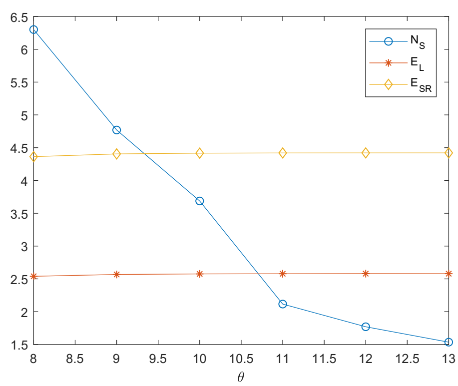


(a) Effect of μ

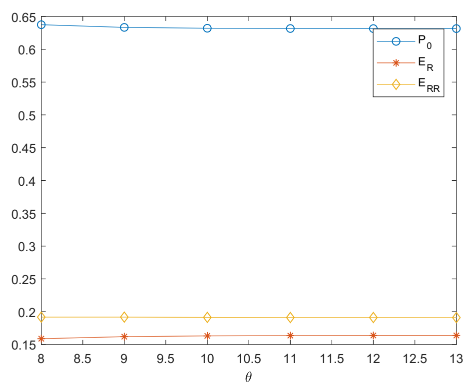


(b) Effect of μ

FIGURE 2. Effect of μ for $(s, S, \lambda, \gamma, \theta) = (20, 45, 8, 2, 9)$.

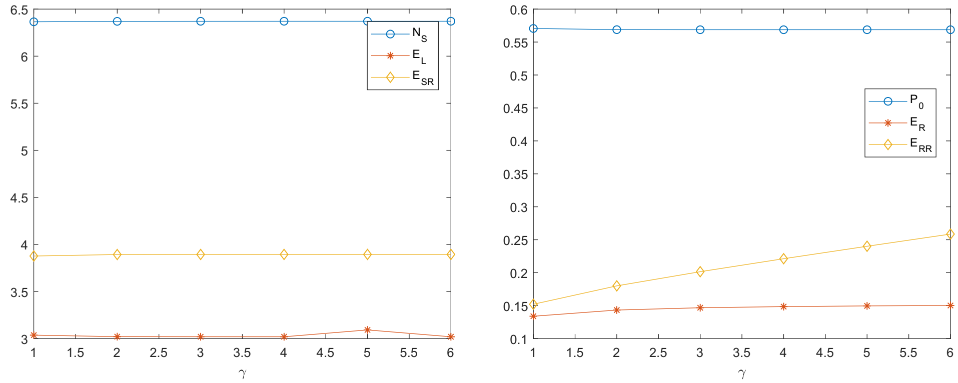


(a) Effect of θ



(b) Effect of θ

FIGURE 3. Effect of θ with $(s, S, \lambda, \gamma, \mu) = (20, 45, 7, 2, 12)$.



(a) Effect of γ (b) Effect of γ

FIGURE 4. Effect of γ with $(s, S, \lambda, \theta, \mu) = (15, 40, 7, 8, 9)$.

TABLE 1. Optimal (s, S) pair and minimum cost.

Optimum (s, S) pair	(1,23)	(2,22)	**(3,22)**	(4,22)	(5,23)	(6,23)
Minimum cost	27.204	25.325	**24.902**	25.546	28.864	31.135

5.1. Cost Analysis

Now for computing the minimal cost for the system under study, we introduce a cost function $\mathcal{F}(s, S)$ defined by

$$\mathcal{F}(s, S) = C_1 E_{\text{RR}} + C_2 N_S + C_3 E_I + C_4 P_0 + C_5 E_L - C_6 E_{\text{SR}}$$

where
C_1 = Cost of inventory procurement per item
C_2 = Cost of holding customers for one unit of time
C_3 = Cost of holding inventory for one unit of time
C_4 = Cost per unit time due to an idle server
C_5 = Cost per unit time due to loss of customer
C_6 = Revenue per unit time due to successful retrial

In order to study the cost function we first fix $(C_1, C_2, C_3, C_4, C_5, C_6, \lambda, \mu, \theta, \gamma) = (\$100, \$2, \$3, \$10, \$20, \$30, 8, 10, 9, 2)$.

We notice from our computational experiments that the cost function constructed behaves like a convex function in the pair (s, S). For the input values of parameters indicated, Table 1 indicates that the optimal (s, S) pair is $(3, 22)$ and the minimum cost is \$24.902 (given in bold letters in Table 1). In this model it is difficult to prove analytically that the cost function is convex in both s and S because of high non-linearity of the function. Nevertheless, all numerical experiments (see Fig. 5) we have performed indicate that this cost function first decreases in S (or in s), attains a minimum and then starts going up.

Conclusion:

In this paper we produced explicit product form solution for a retrial – queueing – inventory problem. For the analysis (s, S) inventory control policy was followed. Other control policies also can be shown to yield the

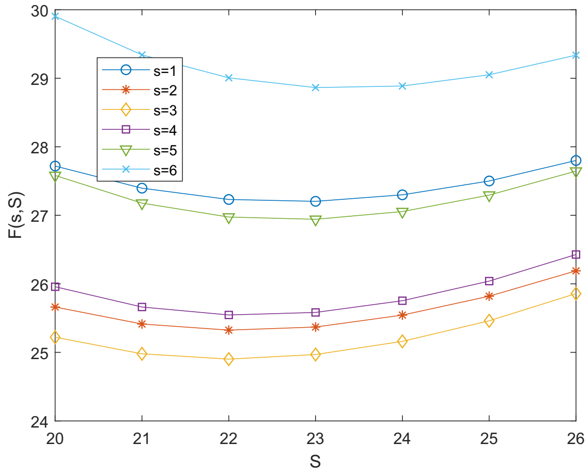


FIGURE 5. Cost function.

product form solution. Because of the highly non linear nature of the cost function in s and S, it is extremely difficult to prove that it is convex in both variables. However, our computational experience indicates that the cost function constructed is convex.

We conjecture that for the arbitrary distributed lead time case (see [15]) our model should yield product form solution.

Acknowledgements. The authors thank the Prof. Soren Asmussen for valuable suggestions which helped in improving the presentation of the paper. The authors also thank the reviewers for their critical comments and suggestions that lead to vast improvement in the presentation of the paper. Research supported by Kerala State Council for Science, Technology & Environment (No.001/KESS/2013/CSTE) in Department of Mathematics, Cochin University of Science and Technology, Cochin-682022, India. Research supported by UGC No.F.6-6/2017-18/EMERITUS-2017-18-GEN-10822 (SA-II).

REFERENCES

[1] J.R. Artalejo and A. Gomez-Corral, Retrial Queueing Systems. Springer Verlag, Berlin Heidelberg (2008).

[2] J. Baek, O. Dudina and C. Kim, A queueing system with heterogeneous impatient customers and consumable items. *Int. J. Appl. Math. Comput. Sci.* **27** (2017) 367–384.

[3] J.W. Baek and S.K. Moon, The $M/M/1$ queue with a production-inventory system and lost sales. *Appl. Math. Comput.* **233** (2014) 534–544.

[4] J.W. Baek and S.K. Moon, A production-inventory system with a Markovian service queue and lost sales. *J. Korean Stat. Soc.* **45** (2016) 14–24.

[5] J.W. Baek, Y.H. Bae, H.W. Lee and S. Ahn, Continuous-type (s,Q)-inventory model with an attached $M/M/1$ queue and lost sales. *Perform. Eval.* **125** (2018) 68–79.

[6] R. Krenzler and H. Daduna, Loss systems in a random environment: steady-state analysis. *Queueing Syst.* **80** (2014) 127–153.

[7] P.C. Devi, B. Sivakumar and A. Krishnamoorthy, Optimal control policy of an inventory system with postponed demand. *RAIRO: OR* **50** (2016) 145–155.

[8] Dhanya Shajin, A. Krishnamoorthy and A.Z. Melikov, On partial and complete blocking set of sets of states in Queueing-inventory model. *Appl Comput. Math.* **18** (2019) 189–201.

[9] Dhanya Shajin, A.N. Dudin, O. Dudina and A. Krishnamoorthy, A two-priority single server retrial queue with additional items. To appear in: *J. Ind. Manage. Optim.* Doi: 10.3934/jimo.2019085 (2019).

[10] A. Krishnamoorthy, B. Lakshmy and R. Manikandan, A survey on inventory models with positive service time. *OPSEARCH* **48** (2011) 153–169.

[11] A. Krishnamoorthy and C.N. Viswanath, Stochastic decomposition in production inventory with service time. *Eur. J. Oper. Res.* **228** (2013) 358–366.

[12] M.F. Neuts, Matrix-geometric Solutions in Stochastic Models: an Algorithmic Approach. The Johns Hopkins University Press, Baltimore [1994 version is Dover Edition] (1981).

[13] M.F. Neuts and B.M. Rao, Numerical investigation of a multi-server retrial model. *Queueing Syst.* **7** (1990) 169–190.

[14] S. Otten, R. Krenzler and H. Daduna, Models for integrated production-inventory systems: steady state and cost analysis. *Int. J. Prod. Res.* **54** (2016) 6174–6191.

[15] M. Saffari, S. Asmussen and R. Haji, The $M/M/1$ queue with inventory, lost sale and general lead times. *Queueing Syst.* **75** (2013) 65–77.

[16] M. Schwarz, C. Sauer, H. Daduna, R. Kulik and R. Szekli, $M/M/1$ queueing systems with inventory. *Queueing Syst.* **54** (2006) 55–78.