RESEARCH PAPER

Adaptive Control Framework for Cost-Efficient Operations in Retrial Service Facility Systems

P. Maheswari*[1] and S.D. Jeniffer[2]
*[1]Department of Mathematics, Mepco Schlenk Engineering College,
Sivakasi, Tamil Nadu, India, E-mail: maheswari23061993@gmail.com
[2]Department of Mathematics, Mepco Schlenk Engineering College,
Sivakasi, Tamil Nadu, India, Email- jenifferdavid1996@gmail.com

Abstract.
This research investigates cost-efficient inventory and service control in a retrial service facility system with inventory maintenance. Traditional queueing–inventory models often assume static operating conditions and limited customer behavior, restricting their applicability in real-world service environments characterized by repeated customer retrials and stochastic demand patterns. To address these challenges, the proposed model integrates inventory management and service control within a unified decision-making framework. The system considers finite inventory capacity, limited waiting space, reorder levels, and stochastic lead times, enabling a comprehensive analysis of service and replenishment operations.
The problem is formulated as a Semi-Markov Decision Process (SMDP), where optimal ordering and service control decisions are determined at various decision epochs. A linear programming approach is employed to derive the optimal policy that minimizes the long-run expected operating cost while maintaining satisfactory service performance.
Keywords: Retrial service facility, Inventory management, Semi-Markov Decision Process (SMDP), Cost-efficient operations, Service control, Queueing-inventory systems, Stochastic demand, Optimal ordering policy.
How to cite this article: Maheswari P, Jeniffer SD. Adaptive Control Framework for Cost-Efficient Operations in Retrial Service Facility Systems. Int J Drug Deliv Technol. 2026;16(58s): 426-432. DOI: 10.25258/ijddt.16.58s.41
Source of support: Nil.
Conflict of interest: None.

1. Introduction

The integration of inventory management and service operations has become increasingly important in many practical systems, including manufacturing facilities, repair centers, healthcare services, and communication networks. In such systems, customers may not always receive immediate service because of server congestion or resource limitations. When service is unavailable, customers often reattempt service after a certain period, leading to the formation of retrial queues. The analysis and optimization of these systems are essential for improving operational efficiency, reducing costs, and enhancing customer satisfaction.
This study focuses on a retrial service facility system with inventory support, where service completion requires the availability of inventory items. The primary objective is to determine optimal state-dependent control policies that govern service and inventory decisions. Unlike fixed-control mechanisms, the proposed approach allows the service rate to vary according to the current state of the system, particularly the number of customers waiting in the retrial orbit. Such adaptive control policies provide greater flexibility and can significantly improve system performance under varying demand conditions.
Retrial queueing theory has evolved into an important area of operations research. Initially motivated by applications in telephone communication systems, retrial models are now widely used to represent a variety of real-world processes in which unsuccessful customers return for service after a random delay. Examples include computer networks with repeated access requests, maintenance systems with repeated repair demands, and service centers with call-back mechanisms. The ability of retrial queues to capture repeated customer behavior makes them particularly useful for analyzing congestion and service accessibility.
A considerable body of literature has examined queueing systems that incorporate inventory considerations. Researchers have developed numerous models to investigate the interaction between inventory replenishment and customer service activities. These studies have demonstrated that inventory-related decisions can significantly influence service quality, operational costs, and resource utilization. Various optimization techniques have been employed to identify effective inventory policies under different demand and replenishment environments.
Over the past two decades, considerable attention has been devoted to service facility systems integrated with inventory management. Early contributions by Berman and Sapna (1998), Arivarignan and Elango (2002), and Selvakumar et al. (2019) examined various aspects of inventory-controlled service systems and developed analytical models to evaluate their operational performance.
Retrial queueing systems, in which customers who fail to receive immediate service reattempt after a certain period, were initially studied in the context of

*Author for Correspondence: : maheswari23061993@gmail.com

telecommunication networks. Subsequently, their applicability expanded to computer and communication systems, where repeated access requests are common. For instance, a group of computers competing to access a shared database can be effectively represented through a retrial queueing framework, since only one user may access the resource at a given time.

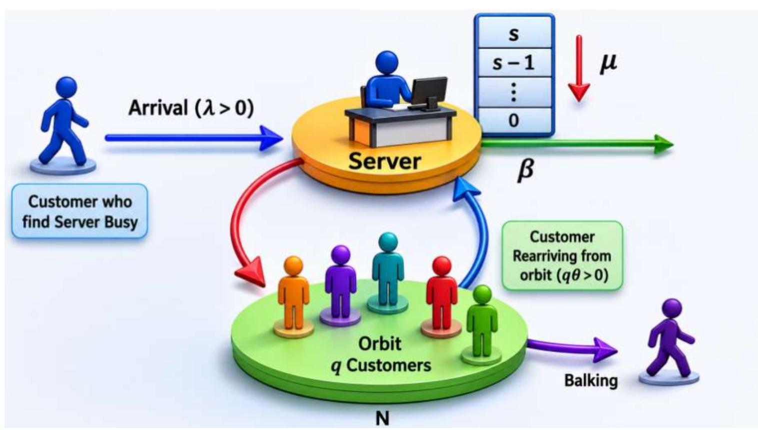
During the last three decades, retrial queueing theory has evolved significantly through the contributions of numerous researchers. Studies by Elcan (1999), Arivudainambi et al. (2009), Dragieva (2013), Dudin et al. (2015), Artalejo et al. (2000, 2000a, 2008), Krishnakumar et al. (2018), and Selvakumar et al. (2023) investigated various single-server retrial queueing models incorporating customer-return behavior, balking mechanisms, and Bernoulli vacation policies. These works employed analytical approaches such as matrix-analytic methods, generating function techniques, truncation procedures, and level-dependent quasi-birth-and-death (LDQBD) processes to derive system characteristics and solution algorithms.

Most of these studies assumed customer arrivals to follow a Poisson process and service times to be exponentially distributed. Their analyses focused on obtaining important performance measures, including system stability, queue lengths, waiting times, and server utilization. Furthermore, several researchers developed cost models to evaluate the long-run operating performance of the systems and determined optimal parameter settings by minimizing the expected total cost rate.

Motivated by these developments, the present study investigates a service facility system coupled with inventory control under a retrial queueing environment. Customers who encounter a busy server enter an orbit and retry for service after a random interval. The model assumes finite capacities for both the orbit and the inventory source. In addition, the service rate is regulated by a decision-making agent, introducing a semi-Markov decision process (semi-MDP) framework for system control and optimization.

The rest of this article is organized as follows. Section 2 presents the formulation of the Semi-Markov Decision Process (Semi-MDP) model. The analytical development of the model is provided in Section 3. Finally, Section 6 describes the procedure for implementing the long-run expected cost rate criterion and determining the optimal policy for the infinite-horizon system under consideration.

2. Model Description



We consider a service facility that maintains an inventory to provide service to customers. Customers arriving when the server is busy may join a finite-capacity orbit and retry later. The system operates under the following assumptions:

- Customer arrivals follow a Poisson process with arrival rate λ (> 0).
- If an arriving customer finds the server idle, the customer immediately enters service, receives the required item, and then departs from the system.
- If the server is busy upon arrival, the customer joins a retrial orbit with a maximum capacity of N.
- Whenever the inventory level is zero, the arriving customer enters into the orbit.
- Whenever the server is busy and the number customer in the orbit is N (the orbit is full), the arriving customer is forced to balk the system.
- Each customer in the orbit reattempts for getting service after an exponential time at rate $q\theta$ (> 0). (q - number of customer in the orbit)
- Service times of customers are independent of each other and have a common exponential distribution with parameter β_i $(0 \le i \le K)$.
- Each β_i $(0 \le i \le K)$.can be chosen from given set of N values $\{\beta_1, \beta_2, \beta_3, \dots . \beta_K\}$ where i depends on number of customer in orbit, when the orbit is empty, the service rate become $\beta_0 = 0$, after service completion of the customer inside the server.
- One (unit) of item is served to each customer during service. The maximum capacity of the inventory is fixed as S. Whenever the inventory level reaches to a prefixed level $s(0 \le s<S)$, a decision for ordering fixed Q = S-s items such that Q > s or non - ordering is taken. Consequently, at levels 1,2,3,…s-1,s decision is

taken for ordering or non - ordering and at level 0 compulsory order for S items is placed. Lead time for the inventory is assumed to be exponentially distributed with parameter $\mu > 0$.

- The size of the order is adjusted at the time of replenishment so that immediately after replenishment the inventory level becomes S.

3. Analysis

Let X(t), M(t) and I(t) denotes the status of the server, number of customers in the *orbit* and inventory level at time t, respectively.

Then {(X (t), M (t), I (t)): t ≥ 0} is a three dimensional continuous time Markov process with state space $E_1 \times E_2 \times E_3$, where $E_1=\{0,1\}$ (0 denotes the idle server and 1 denotes the busy server), $E_2= \{0,1,2,\ldots,N\}$and $E_3= \{1,2,\ldots,S\}$

The infinitesimal generator A of the Markov process has entry of the form $\left(a_{(p,q,r)}^{(l,m,n)}\right)$.

Some of the state transitions are given below:

From state (0,q,r) only transitions into the following states are possible:

(i) To the state (1,q,r) with rate λ for $0 \le q \le N; 1 \le r \le S$ (customer arrival)

(ii) (1,q-1,k) with rate $q\theta$ for q=1, 2,…N; $1 \le r \le S$ (Customer arrive from an orbit).

From state (1,q,r) only transitions into the following states are possible:

(i) To the state (1,q+1,r) with rate λ for $0 \le q \le N-1; 1 \le r \le S$ (customer arrival).

(ii) (0,q-1,r-1) with rate β_i for $1 \le q \le N;\ 1 \le r \le S$ (Service completion)

From states (0, q, 0) transitions are possible only to the states (0, q+1, 0) (primary customer arrival with rate λ for $0 \le q \le N-1$.

From state (0, q, r) transitions are possible to the states (0,q, S) for $0 \le q \le N$ and $0 \le r \le s$ (replenishment order Q items is placed and adjusted).

From state (1, j, k) transitions are possible to the states (1, j, S) for $0 \le q \le N$ and $0 \le r \le s$ (replenishment order Q items is placed and adjusted).

No impatient customer. Customer leaves from system (or) orbit. The service of the system is finite, Arrival source to the system is finite. ie) The orbit size is always finite.

Now, we have to convert this Markov process into continuous time MDP by considering the following five components,

4. MDP Formulation

(i) **Decision epochs:** The decision epochs are random points as time line at each service Completion.

(ii) **State space**:

E=

$$\{e=(p,q,r): p=0,1; 0 \le q \le \mathrm{N}, \mathrm{N}<\infty; 1 \le r \le \mathrm{S}\}$$

(iii) **Action set**:

$$A=\{a: 0 \le a=i \le k\}$$

$$A=\bigcup_{e \in E} A_e \ .$$

A decision rule from the class of rules Π is equivalent to a function π: $E \to A$ and is given by

$$\pi(p,q,r)=\{a:(p,q,r) \in \mathrm{E}, \mathrm{a} \in \mathrm{A}\} .$$

(iv) **Transition probability**:

$p_{(p,q,r)}^{(l,m,n)}(a)$-a transition probability from state (p,q,r) to the state (l,m,n).

(v) **Cost**:

Cost occurred when action 'a' is taken at state (p,q,r) is given by $C^{\pi}\left((l,m,n) \mid (p,q,r), a\right)$.

The long-run expected (average) cost rate when policy Π is adopted is given by

$$C^{\pi}=c_1 \overline{I}^{\pi}+c_2 \overline{W}^{\pi}+c_3 \overline{\alpha}^{\pi}+c_4 \overline{r}^{\pi}+g \overline{B}^{\pi} \tag{1}$$

when in the steady state , for given policy Π is given by r^{π} is the expected reorder rate, $\overline{W}^{\pi}$ is the average waiting time for a customer, $\overline{\alpha}^{\pi}$ is the service expected cost due to different service rate due to d rate, c_1 denotes the holding cost / unit time/ unit item, c_2 denotes the waiting cost /customer / unit time , c_3 denotes the service cost / customer, c_4 denotes the reorder cost / unit time/ unit item and g denotes the balking cost/customer associated with using the different service rates.

5. **Steady state Analysis:**

Let f denote the stationary policy, which is deterministic time invariant and Markovian Policy (MD). From the assumptions made in our system, it can be seen that $\{(X(t), \mathrm{M}(t), \mathrm{I}(t)): t \ge 0\}$ is the controlled process $\{(X^f(t), \mathrm{M}^f(t), \mathrm{I}^f(t)): t \ge 0\}$ when policy f is adopted. Since the process $\{(X^f(t), \mathrm{M}^f(t), \mathrm{I}^f(t)): t \ge 0\}$ is a Markov Process with finite state space E. The process is completely Ergodic, if every stationary policy gives rise to an irreducible Markov chain. It can be seen that for every stationary policy π the Markov process is completely Ergodic and also the optimal stationary policy π^* exists, because the state and action spaces are finite

Our objective is to find an optimal policy π^* for which $C^{\pi^*} \le C^{\pi}$ for every MD policy in Π^{MD}.

For any fixed MD policy Π^{MD} and (p,q,r),(l,m,n) $\in E$, define

$$P_{pqr}^{\pi}(\mathrm{l,m,n,t})=prob\{\mathrm{X}^{\pi}(\mathrm{t})=l, \mathrm{M}^{\pi}(\mathrm{t})=\mathrm{m}, \mathrm{I}^{\pi}(\mathrm{t}) / \mathrm{X}^{\pi}(0)=p, \mathrm{M}^{\pi}(0)=q, \mathrm{I}^{\pi}(0)=r\};$$

$(p,q,r),(\mathrm{l,m,n}) \in \mathrm{E}$

Now $p_{pqr}^{\Pi}(\mathrm{l,m,t})$ satisfies the Kolmogorov forward differential equation. P'(t)=P(t)A, where A is an infinitesimal generator of the Markov process $\{(\mathrm{X}^R(\mathrm{t}), \mathrm{Z}^R(\mathrm{t}), \mathrm{I}^R(\mathrm{t})): \mathrm{t} \ge 0\}$

For each MD policy Π, we get a Markov chain with state space E and action set A which are finite, $P^{\pi}(l,m,n)=\lim_{t\to\infty}P^{\pi}_{pqr}(l,m,n;t)$ exists and is independent of initial state (p,q,r)conditions.
The balance equations are obtained by using the fact that transition out of a state is equal to transition into a state

$$(\lambda+q\theta)\mathrm{P}^{\pi}(0,q,\mathrm{S})=\mu.\mathrm{P}^{\pi}(0,q,\mathrm{s}),\ 0\le q\le N \quad (2)$$

$$(\lambda+\beta)\mathrm{P}^{\pi}(1,0,\mathrm{Q}+r)=\lambda P^{\pi}(0,0,\mathrm{Q}+r)+\theta P^{\pi}(0,1,\mathrm{Q}+r)+\mu\mathrm{P}^{\pi}(1,0,r),\ 1\le r\le s,\ Q=S-s \quad (3)$$

$$(\lambda+\beta)\mathrm{P}^{\pi}(1,q,\mathrm{Q}+r)=\lambda\sum_{i=0,1}P^{\pi}(p,q-p,\mathrm{Q}+r)+(q+1)\theta P^{\pi}(0,q+1,\mathrm{Q}+r)+\mu\mathrm{P}^{\pi}(1,q,r),\ 1\le q\le N-1,1\le r\le s \quad (4)$$

$$\beta\mathrm{P}^{\pi}(1,\mathrm{N},\mathrm{Q}+r)=\lambda\sum_{p=0}^{1}P^{\pi}(p,\mathrm{N}-p,\mathrm{Q}+r)+\mu\mathrm{P}^{\pi}(1,\mathrm{N},r),\ 1\le r\le s, \quad (5)$$

$$(\lambda+q\theta)\mathrm{P}^{\pi}(0,q,\mathrm{Q}+r)=\beta\mathrm{P}^{\pi}(1,\mathrm{q},\mathrm{Q}+r+1)+\mu\mathrm{P}^{\pi}(0,\mathrm{q},\mathrm{r}),\ 0\le q\le \mathrm{N},0\le r\le s-1 \quad (6)$$

$$(\lambda+q\theta)\mathrm{P}^{\pi}(0,\mathrm{q},\mathrm{r})=\beta\mathrm{P}^{\pi}(1,\mathrm{q},\mathrm{r}+1),\ 0\le q\le \mathrm{N},\ s+1\le r\le Q-1 \quad (7)$$

$$(\lambda+\beta)\mathrm{P}^{\pi}(1,0,\mathrm{r})=\lambda\mathrm{P}^{\pi}(0,0,\mathrm{r})+\theta\mathrm{P}^{\pi}(0,1,\mathrm{r}),\ s+1\le r\le Q \quad (8)$$

$$(\lambda+\beta)\mathrm{P}^{\pi}(1,\mathrm{q},\mathrm{r})=\lambda\sum_{p=0}^{1}P^{\pi}(\mathrm{p},\mathrm{q}-p,\mathrm{r})+(\mathrm{q}+1)\theta\mathrm{P}^{\pi}(0,\mathrm{q}+1,\mathrm{r}),\ 1\le q\le \mathrm{N}-1,s+1\le r\le Q \quad (9)$$

$$\beta\mathrm{P}^{\pi}(1,\mathrm{N},\mathrm{r})=\lambda\sum_{p=0}^{1}P^{\pi}(\mathrm{p},\mathrm{N}-p,\mathrm{r}),\ s+1\le r\le Q \quad (10)$$

$$(\lambda+q\theta+\mu)\mathrm{P}^{\pi}(0,\mathrm{q},\mathrm{r})=\beta\mathrm{P}^{\pi}(1,\mathrm{q},\mathrm{r}+1),0\le q\le N,\ 1\le r\le s \quad (11)$$

$$(\lambda+\beta+\mu)\mathrm{P}^{\pi}(1,0,\mathrm{r})=\lambda P^{\pi}(0,0,\mathrm{r})+\theta\mathrm{P}^{\pi}(0,1,\mathrm{r}),\ 1\le r\le s \quad (12)$$

$$(\lambda+\beta+\mu)\mathrm{P}^{\pi}(1,\mathrm{q},\mathrm{r})=\lambda\sum_{p=0}^{1}P^{\pi}(\mathrm{p},\mathrm{q}-p,\mathrm{r})+(\mathrm{q}+1)\theta\mathrm{P}^{\pi}(0,\mathrm{q}+1,\mathrm{r}),\ 1\le q\le N-1,1\le r\le s \quad (13)$$

$$(\beta+\mu)\mathrm{P}^{\pi}(1,\mathrm{N},\mathrm{r})=\lambda\sum_{p=0}^{1}P^{\pi}(\mathrm{p},\mathrm{N}-p,\mathrm{r}),\ 1\le r\le s \quad (14)$$

$$(\lambda+\mu)\mathrm{P}^{\pi}(0,0,0)=\beta\mathrm{P}^{\pi}(1,0,1), \quad (15)$$

$$(\lambda+\mu)\mathrm{P}^{\pi}(0,\mathrm{q},0)=\beta\mathrm{P}^{\pi}(1,\mathrm{q},1)+\lambda\mathrm{P}^{\pi}(0,\mathrm{q}-1,0),\ 1\le q\le N-1 \quad (16)$$

$$\mu.\mathrm{P}^{\pi}(0,\mathrm{N},0)=\lambda P^{\pi}(0,N-1,0)+\beta\mathrm{P}^{\pi}(1,\mathrm{N},1) \quad (17)$$

Together with the above set of equations, the total probability condition

$$\sum_{(\mathrm{p},\mathrm{q},\mathrm{r})\in E}P^{\Pi}(\mathrm{p},\mathrm{q},\mathrm{r})=1,$$

gives steady state probabilities $\{P^{\Pi}(p, q, r), (p, q, r) \in E\}$ uniquely.

6. System Performance Measures.

(i) Mean inventory level in the system is given by

$$\overline{I}^{\pi}=\sum_{p=0}^{1}\sum_{r=1}^{S}r\sum_{q=0}^{N}P^{\pi}(\mathrm{p},\mathrm{q},\mathrm{r}) \quad (19)$$

(ii) Mean waiting time in the system is given by

$$\overline{W}^{\pi}=\sum_{q=1}^{N}\sum_{r=0}^{S}\left(\frac{\lambda+q\theta}{\beta^{\pi}}\right)P^{\pi}(1,\mathrm{q},\mathrm{r}). \quad (20)$$

(iii) Mean reorder rate is given by

$$\overline{r}^{\pi}=\sum_{q=1}^{N}\beta^{\pi}P^{\pi}(1,\mathrm{q},\mathrm{s}+1) \quad (21)$$

(iv)The expected service rate is given by

$$\overline{\alpha}^{\pi}=\sum_{q=0}^{N}\sum_{r=1}^{S}\beta^{\pi}P^{\pi}(1,\mathrm{q},\mathrm{r}) \quad (22)$$

(v) The mean Balking rate is given by

$$\overline{B}^{\pi}=\lambda\sum_{k=1}^{S}P^{\pi}(1,\mathrm{N},\mathrm{r}) \quad (23)$$

The long run expected cost rate is given by

$$C^{\Pi}=c_1\sum_{p=0}^{1}\sum_{r=1}^{S}r\sum_{q=0}^{N}P^{\pi}(\mathrm{p},\mathrm{q},\mathrm{r})+c_2\sum_{q=1}^{N}\sum_{r=1}^{S}\left(\frac{\lambda+q\theta}{\beta^{\pi}}\right)P^{\pi}(1,\mathrm{q},\mathrm{r})+\mathrm{c}_3\sum_{q=0}^{N}\sum_{r=1}^{S}\beta^{\pi}P^{\pi}(1,\mathrm{q},\mathrm{r})+c_4\sum_{q=0}^{N}\beta^{\pi}P^{\pi}(1,\mathrm{q},\mathrm{s}+1)+g(\lambda)\sum_{r=1}^{S}P^{\pi}(1,\mathrm{N},\mathrm{r}) \quad (24)$$

7. Linear Programming Problem:

7.1 Formulation of LPP

In this section we propose a LPP model within a MDP framework. First we define the variables D(p,q,k,a) as a conditional probability expression

$$D(\mathrm{p},\mathrm{q},\mathrm{r},\mathrm{a})=\mathrm{Prob}\left\{decisionis\,'a'/\,\mathrm{stateis}\,(\mathrm{p},\mathrm{q},\mathrm{r})\right\}$$

Since $0 \le D(p,q,r,a)\le 1$, this is compatible with Randomized time invariant Markovian policies. Here, the Semi-Markovian decision problem can be formulated as a linear programming problem.

Hence

$$0 \le D(\mathrm{p,q,r}) \le 1 \ \ and \ \sum_{a \in A} D(\mathrm{p,q,r},a) = 1, \mathrm{p} = 0,1;$$

$$0 \le q \le N; 0 \le r \le S$$

For the reformulation of the MDP as LPP, we define another variable y(p,q,r,a) as follows.

$$\mathrm{y(p,q,r},a) = D(\mathrm{p,q,r,a})P^{\Pi}(\mathrm{p,q,r}) \qquad (25)$$

From the above definition of the transition probabilities

$$P^{\pi}(\mathrm{p,q,r}) = \sum_{a \in A} y(\mathrm{p,q,r,a}), (\mathrm{p,q,r}) \in \mathrm{E}, \mathrm{a} \in \mathrm{A} \qquad (26)$$

Expressing P^{Π}(p,q,k) in terms of y(p,q,k,a), the expected total cost rate functions(24) is obtained and the LPP formulation is of the form

Minimize

$$C^{\Pi} = c_1 \sum_{p=0}^{1}\sum_{r=1}^{S} r \sum_{q=0}^{N} P^{\pi}(\mathrm{p,q,r}) + c_2 \sum_{q=1}^{N}\sum_{r=1}^{S}\left(\frac{\lambda + q\theta}{\beta^{\pi}}\right)$$

$$P^{\pi}(1,\mathrm{q,r}) + \mathrm{c}_3 \sum_{q=0}^{N}\sum_{r=1}^{S} \beta^{\pi} P^{\pi}(1,\mathrm{q,r}) + c_4 \sum_{q=0}^{N} \beta^{\pi}$$

$$P^{\pi}(1,\mathrm{q,s+1}) + g(\lambda)\sum_{r=1}^{S} P^{\pi}(1,\mathrm{N,r}) \qquad (27)$$

Subject to the constraints,

(i) $\mathrm{y(p,q,r,a)} \ge 0; (\mathrm{p,q,r}) \in \mathrm{E}, \mathrm{a} \in \mathrm{A}$

$$ii)\sum_{l=0,1}\ \sum_{(\mathrm{p,q,r}) \in \mathrm{E}}\ \sum_{a \in A} y(\mathrm{p,q,r,a}) = 1,$$

And the balance equation (2)-(17) obtained by replacing

$$P^{\pi}(\mathrm{p,q,r}) \text{ with } \sum_{a \in A} y(\mathrm{p,q,r,a}).$$

7.1 Lemma:

The optimal solution of the above Linear Programming Problem yield a deterministic policy.

Proof:

From the equations (25) and (26)

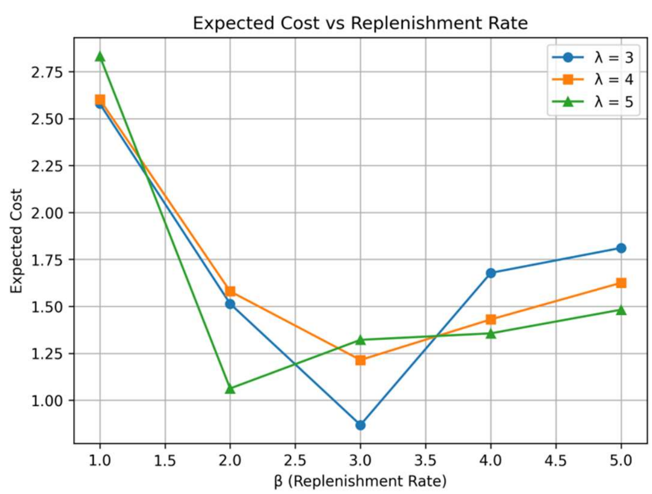
$$\mathrm{D(p,q,r,a)} = \frac{y(\mathrm{p,q,r,a})}{\sum_{r=0}^{N} y(\mathrm{p,q,r},\beta_i)}, a = \beta_i, \mathrm{i} = 0,1,2,\ldots\mathrm{k} \qquad (28)$$

$$P^{\pi}(\mathrm{p,q,r}) = \sum_{a \in A} y(\mathrm{p,q,r,a}), (\mathrm{p,q,r}) \in \mathrm{E} \qquad (29)$$

Since the decision problem is completely Ergodic, every feasible solution to above linear programming problem has the property that for each $(p,q,r)\epsilon E$, $y(p,q,r,a)$>0for exactly one $a\epsilon E$.
Hence for each $(p,q,r)\epsilon E$, $y(p,q,r,a)$=1, for atleast one value of a and zero for all other values of a. Thus, given the number of customers in the orbit, we have to choose the service rate β for which $y(p,q,r,a)$=1. Hence any basic feasible solution of the linear programming yields a deterministic policy.

8. Numerical illustration and Discussion:

In this section we consider a service facility system to illustrate the method described in section 7, through numerical examples. We implemented TORA software to solve LPP by simplex algorithm.
The following table describes the solution for LPP problem by varying the arrival (Poisson) rates from 2 to 5 and an exponential service rates from 1 to 7. The expected cost is computed by taking waiting cost per customer is 0.5 and the service cost per customer is 0.8.



Arrival rate →	=3	=4	=5

Service rate ↓			
β=1	2.5789	2.6019	2.8331
β=2	2 1.5137	1.5798	1.0623
β=3	0.8679	1.2142	1.3214
β=4	1.6781	1.4302	1.3560
β=5	1.8112	1.6252	1.4821

Table 1: The Expected total cost

From the above table,
(a) The minimum expected cost for arrival rate 2 will be obtained by adjusting the service rate as β = 3 per unit time.
(b) The minimum expected cost for arrival rate 4 and 5 will be obtained at service rate 4.

(waiting cost per customer, the service cost per customer) (c_1,c_2)	Expected total cost
(0.5, 0.8)	0.8679
(1, 0.8)	1.3126
(1.5, 0.8)	1.9135
(2, 0.8)	2.3651

Table 2: The Expected total cost for varying waiting cost per customer

(waiting cost per customer, the service cost per customer) (c_1,c_2)	Expected total cost
(0.5, 0.8)	0.8679
(0.5, 1.6)	1.2013
(0.5, 2.4)	1.4862
(0.5, 3.2)	1.7607

Table 3: The Expected total cost for varying service cost per customer

9. Conclusions and future research

This work investigates the problem of service-rate control in a retrial service facility with inventory support. Unlike many existing studies that primarily focus on inventory replenishment policies or the evaluation of performance measures, the present study develops a decision-making framework for selecting appropriate service rates. The system is formulated as a Semi-Markov Decision Process (SMDP), and the optimal policy is obtained using a linear programming approach. The objective is to determine service-rate decisions that minimize the long-run average operating cost while considering the effects of customer retrials, inventory availability, and system congestion.

The proposed model demonstrates how service control and inventory management can be integrated within a unified optimization framework. The results indicate that appropriate service-rate selection can significantly influence overall system performance and operating costs, providing useful guidance for the efficient management of service facilities.

Several opportunities exist for further research. A natural extension is to relax the assumption of Poisson arrivals and consider more general arrival processes, which may better represent practical situations. The corresponding optimization problem can still be explored within the Semi-Markov Decision Process framework. Future studies may also examine systems with multiple servers, heterogeneous customers, finite waiting times, server interruptions, or alternative inventory replenishment strategies. Such extensions would broaden the scope of the model and increase its relevance to a wider range of real-world service and inventory systems.

REFERENCES

1. Arivaringan, G., Elango, C. & Arumugam, N., (2002), *A Continuous Review Perishable*
2. *Inventory Control System at Service Facilities*, pp. 19-40, Notable Publications Inc, New Jersey.
3. Arivudainambi, D., Averbakh, I., Berman, O., 2009), *Stationary analysis of a single server retrial queue with priority and vacation*, International Journal of Operational Research 5, 26-47.
4. Artalejo, J.R., Gomez-Corral, A., (2008), *Retrial Queueing Systems: A Computational Approach,*, Springer, Berlin.
5. Artalejo, J.R., Lopez-Herrero, M.J., (2000), *On the single server retrial queue with balking*, Infor 38, 33-50.
6. Artalejo, J.R., Rajagopalan, V., Sivasamy, R., (2000), *On finite Markovian queues with*
7. *repeated attempts*, Investigacion Operativa 9, 83-94.
8. Berman, O. & Sapna, K. P., (1998), *Inventory management at service facilities for systems with arbitrarily distributed service times*, Working

paper, Faculty of Management, University of Toronto.

9. Dragieva, V. I., *A finite source retrial queue: number of retrials*, Communications in Statistics - Theory and Methods, 42, (2013), 812-829.
10. Dudin, A. N., Manzo, R., and Piscopo, R., (2015), *Single server retrial queue with group admission of customers*, Computers & Operations Research, 61, 89-99.
11. Elcan, A., (1999), *Asymptotic Bounds for an Optimal State-Dependent Retrial Rate of the M/M/l Queue with Returning Customers*, Mathematical and Computer Modelling, 30, 129-140.
12. Krishnakumar, S., Selvakumar, C. and Elango, C., (2018*), Inventory Ordering Control for a Retrial Service Facility System - Semi- MDP*, International Journal of Engineering Science Invention, ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726, Volume 7, Issue 6, Ver I, PP 14-20.
13. Selvakumar, C., Maheswari, P. and Elango, C., (2019), *Discrete MDP Problem in Service Facility Systems with Inventory Management,* International Journal of Computational Systems Engineering, ISSN: 2046 - 3391 (P), 2046 - 3405 (online), volume 5, No. 4, 243 - 250.
14. Selvakumar, C., Maheswari, P., Gowsalya, V. and Krishnakumar, S., (2023), *Service control in a Retrial service facility System with Two type of customer Semi-MDP*, International Journal on Recent and Innovation Trends in Computing and Communication, ISSN: 2321 - 8169, volume 11, Issue. 9, 3515 - 3523.