DISCRETE MARKOV DECISION PROCESS – INVENTORY CONTROL IN A SERVICE FACILITY SYSTEM

C. Selvakumar, P. Maheswari and C. Elango
Research Department of Mathematical Sciences,
Cardamom Planters' Association College,
Bodinayakanur - 625 513, Tamil Nadu, India

ABSTRACT
This article addresses the problem of inventory control in a service facility system in which ordering level is controlled by MDP. A discrete time service facility system is studied in which demand process follows a Bernoulli process. If a customer finds a free server, then he/ she enter the server immediately, otherwise join the waiting space with finite capacity say N. Inventory pool with maximum level S is maintained at service station to satisfy the customers. The (s, S) type policy is adopted to the inventory the lead times and service times are assumed to follow a geometric distribution. Average cost criteria based MDP is implemented to fix the optimal policy to be implemented. The results are illustrated numerically.

KEYWORDS: Discrete time service facility system, discrete inventory system, (s, S) ordering policy, Markov Decision Processes (MDP).

MSC: 90C40

RESUMEN
Este articulo trata el problema del control de inventarios en un sistema de facilidades en el cual el nivel de órdenes es controlado por un MDP. Un sistema de facilidades tiempo discreto es estudiado en el cual el proceso de demanda sigue un proceso Bernoulli. Si un cliente halla libre un server, entonces accede inmediatamente, en otro caso se une a un espacio de espera con capacidad finita digamos N. El pool de inventarios con una nivel máximo S es mantenidos en una estación de servicios para satisfacer los La política de inventarios adoptada es del tipo (s, S) los tiempos se asumen siguen una distribución geométrica. El costo promedio basada en MDP se implementa para fijar la política optimal a implementar. Los resultados son ilustrados numéricamente.

PALABRAS CLAVE: Sistema de facilidades de servicio a tiempo discreto, sistema de inventario discreto, Política de ordenes (s, S) , Proceso Markoviano de Decisión (MDP).

1. INTRODUCTION

Inventory control in service facility is a recent topic. Many researchers contributed to this domain with both deterministic and stochastic demand patterns. Berman, O., et al. [2] studied deterministic inventory control in service facility in 1993. Sapna, K.P. and Berman, O. [3], published their articles stochastic inventory control in service facility in 2000. Howard [12] used basic principles of Markov Chain theory and Dynamic programming to develop a policy-iteration algorithm for Markov decision processes problems: means sequential decision process with an infinite planning horizon. A theoretical foundations to Howard policy-iterations method has been developed by Blackwell [4], Denardo and Fox [6] and Veinott [25]. Markov decision model with Linear Programming method has been studied by De Ghellinck [5] and Manne [18], Derman [7] and Hordijk and Kallenberg ([10], [11]). Another method for solving MDP problem is value-iteration algorithm which was developed by Odani [19] to find lower and upper bounds for the minimal average cost.

The Markov decision models finds applications in a wide variety of fields. Some important applications are in the fields of inventory, queueing, communication networks. Machine maintenance is a major area of research with MDP in the last eighties(Golabi et al. [9], Kawai [13], Stengos and Thomas [22] and Tijims and Van der Duyn Schouten [24]. A wide survey of real application of MDP models has been given in White [28].

Discrete queues is object of study in the recent past. A discrete time queue with two identically independent streams of customers and service time of slot with multiple servers is studied in Lin and Silvester [16]. They found the stationary plot of loss of customers from each class under complete sharing. Takahashi and Hashida [23] analyze a discrete time queue with infinite buffer and identical and independent arrivals from multiple priority from head of the queue. Markov modulated arrival process has been studied in Garcia and Casals [8]. Le Boudec [14] discuss numerical example method for computing the loss of difficult classes in a discrete queue with Markov modulated Bernoulli process. Liao [15] obtained the probability of loss of different classes in a discrete queue with Markov modulated Poisson arrivals. Rubin and Tsai [20] and Wanger [27] are other articles dealing with discrete queues. Specifically Rubin and Tsai [20] considered the problem in generating functions approach, by deriving waiting time distributions for the infinite buffer queue and general arrival process.

Wagner [27] studied queueing systems with Markov modulated arrivals and infinite buffer. He obtained system performance measures like waiting times for lower class (priority) customers. Sharma and Gangadhar [26] studied both delay moments and distributions of the two classes of customers (traffic) for a finite discrete queue when space and time priority is considered.

In this paper we considered a discrete time based service facility system maintaining inventory. Hence inventory is main level and used in the service station. For example pumper, tire used in the car service station. Unit blood used in a hospital for operations. In traditional inventory management systems, whenever a demand occurs an item from stock is issued immediately. But in service facility a specific item is issued to the customers after a fixed or random service time completion. So mean time accumulations of demands form a queue emerges. This category of problem in deterministic demand was first considered by Berman, O. et al [2], Sapna, K. and Berman, O., [3] studied both Poisson arrival and renewal type arrivals and inventory with Zero and positive lead times. Arivarignan et al. [29] studied the service facility system with negative demands. All the
above articles deal with only continuous time Markov processes with specific policy in inventory management.

We considered the problem with an analyzing Markov Decision Processes involved in controlling inventory reordering (replenishment) processes with optimal policy. The objective of the model is to determined an optimal ordering policy for the SFS (Service Facility System) with inventory by using average cost criteria. The rest of the paper is organized as follows. Section 1 is of introducing nature and motivation for this research and literature review. In section 2 formulation of the problem as a MDP is proposed. In section 3 analysis of the system and transition probability matrix are introduced to fix the problem within MDP frame. In the section 5, system performance measures are computed. A numerical example is provided in section 9 to substantiate the results obtained is previous sections.

2. MODEL FORMULATION

Consider a service facility in which inventory is maintained to serve the customers. The time horizon considered is of infinite which is divide into intervals of equal length say η (discrete time). We assume that the system activities like arrival, service and departure occurs only of the end points of the time intervals (slots). That is the events are assumed to occur at some points but one after another. The demanded items from the system is issued to the customer only after a random service time. In general retail stores customers, will not or cannot be allowed to wait for long period of time. In our problem customers to wait until the spare part (inventory) is taken from stock to complete the service. Also we assume for convenience of our model that only one item is used per customer.

- Inventory is maintained upto maximum capacity of S and the waiting hall is used to be finite (say N).
- The customer arrives to the service facility system following Bernoulli process with probability p and $\bar{p}$ denote the non-arrival of customers.
- If the server is free the customer immediately enter the server, otherwise join the waiting space with limited capacity N.
- A customer to find the waiting space full (N customers) is forced to balk the system.
- The server from inventory provides uses one item to serve customer. After service completion customer leaves the system with one item.
- Customer's serve times are independent and geometrically distributed with probability q, where $\bar{q}$ = 1- q is the probability of non - completion of serve.

- Order quantity $Q = S\text{-}s$ is placed whenever the inventory levels is $\leq s$ with $Q > s$. The order is delivered after a lead time having geometric distributed with parameter $r > 0$, where $\bar{r} = 1 - r$ is the probability that ordered items not received.

Discrete time systems are concerned the events happens within the period [t, t + 1], that is between time epochs t and t + 1, for t = 0, 1, 2, ... We assume that demand, event occurs first then service completion event occurred of customers and finally the replenishment event occurred.

Notation

$\bar{p} = 1 - p$

$\bar{q} = 1 - q$

$\bar{r} = 1 - r$

3. ANALYSIS

Let the random variable, X_t denote the number of customers waiting in the space provided with maximum capacity N and L_t denote inventory (on hand) level at time epoch t. Clearly the stochastic process $\{(X_t, L_t), t = 0, 1, 2, \ldots\}$ is a discrete time Markov chain with state space $E = \{(i, k): i = 0, 1, 2, \ldots, N, j = 0, 1, 2, \ldots, S\}$, a two dimensional space.

This discrete Markov Chain is given has a probability transition function.

$P((i, k)(j, l)) = Pr\{X_{t+1} = j, L_{t+1} = l \mid X_t = i, L_t = k\}$,

$(i, k), (j, l) \in E$. Various transition probabilities at the occurred of events are defined as

(i) When the number of customers i = 0, and order placed

(a) Pr{A demand arrives} = p, i = 0, k = s + 1, ..., S.

(b) Pr{No demand arrives} = $\bar{p}$, i = 0, k = s + 1, ..., S.

(c) Pr{An arrival of demands occur ordered quantity not replenished} = $p\bar{r}$, i = 0, k = 0, 1, 2 . . . , s.

(d) Pr{No demand arrives and ordered items not replenished} = $\bar{p}\bar{r}$, i = 0, k = 0, 1, 2, . . . , s.

(e) Pr {No demand arrives and quantity ordered replenished} = $\bar{p}$r, i = 0, k = 0, 1, 2, . . . , s.

(f) Pr {An arrival of demand occurs and ordered items replenished} = pr, i = 0, k = 0, 1, 2, . . . , s.

(ii) The number of customers, i ≥1 and k > s (order not placed)

(a) Pr{A demand arrives but service not completed} = $\bar{p}$q, i = 1, 2, , . . .N, k = s + 1, s + 2, . . . , S.

(b) Pr{No demand arrived and service also not completed} = $\bar{p}\bar{q}$,
i = 1, 2, 3, . . . ,N, k = s + 1, s + 2, . . . , S.

(c) Pr{A demand arrives and ordered items not received} = p $\bar{r}$.

(d) Pr{A demand occurs but service is not completed and ordered items are not received} = p $\bar{q}$ $\bar{r}$,
i = 1, 2, 3, . . .N, k = 1, 2, 3, . . . , s.

We can find all possible combinatorics of events and the probabilities. For example same of the one step transitions probabilities of the Markov chain are obtained as follows,

- Transfer from state i = 0, k = 0, 1, 2 . . . , s to the state j = i and l = k with probability $\bar{p}\bar{r}$.
- Transfer from state i = 1, 2 . . .N, k = 0, to the state j = i, l = j+Q with probability $\bar{p}$r.
- Transfer from state i = 1, 2, . . .N, k = 1, 2, . . . s to states j = i, l = j + Q -1 with probability pqr.

From the above set of arguments we get the transition probability matrix (tpm), P = (p(i, k), (j, l)). Now, we define the states of the system in dictionary order as follows. < i >= (i, 0)(i, 1) . . . (i, S), where i = 0, 1, 2 . . .N. Now we an order the above states that (< 0 > < 1 > < 2 > … < N >) a finite set of vectors.

Now the transition probability matrix P of the discrete Markov chain can be expressed as block partitioned matrix of the form.

$$R = \begin{array}{c} \\ <0> \\ <1> \\ <2> \\ <3> \\ \vdots \\ <N> \end{array} \begin{array}{c} \begin{array}{ccccccc} <0> & <2> & <3> & \cdots & \cdots & \cdots & <N> \end{array} \\ \begin{pmatrix} B_0 & B_1 & 0 & 0 & 0 & \cdots & \cdots \\ A_2 & A_1 & A_0 & 0 & 0 & \cdots & \cdots \\ 0 & A_2 & A_1 & A_0 & 0 & 0 & \cdots \\ 0 & 0 & A_2 & A_1 & A_0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & A_0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & B_2 & B_3 \end{pmatrix} \end{array}$$

where

$$P_{ij} = \begin{cases} A_0 & for\ j = i+1,\ i = 1,2,\dots,N-1 \\ A_1 & for\ j = i,\ i = 1,2,\dots,N-1 \\ A_2 & for\ j = i-1,\ i = 1,2,\dots,N-1 \\ B_0 & for\ j = i,\ i = 0 \\ B_1 & for\ j = i+1,\ i = 0 \\ B_2 & for\ j = i-1,\ i = N \\ B_3 & for\ j = i,\ i = N \\ 0 & otherwise. \end{cases}$$

Now we have

$$[A_0]_{ij} = \begin{cases} p\bar{r} & k = i,\ i = 0 \\ p\bar{q}\bar{r} & k = i,\ i = 1,2,\dots,s \\ p\bar{q} & k = i,\ i = s+1, s+2, \dots, S \\ pq & k = i+Q,\ i = 0 \\ p\bar{q}r & k = i+Q,\ i = 1,2,\dots,s \\ 0 & otherwise \end{cases}$$

$$[A_1]_{ij} = \begin{cases} \bar{p}\bar{r} & k = i;\ i = 0 \\ \bar{p}\bar{q}\bar{r} & k = i;\ i = 1,2,\dots,s \\ \bar{p}\bar{q} & k = i;\ i = s+1, s+2, \dots, S \\ pq\bar{r} & k = i-1;\ i = 1,2,\dots,s \\ \bar{p}\bar{q}r & k = i+Q;\ i = 1,2,\dots,s \\ \bar{p}r & k = i+Q;\ i = 0 \\ pqr & k = i+Q-1;\ i = 1,2,\dots,s \\ pq & k = i-1;\ i = s+1, s+2, \dots, S \\ 0 & otherwise \end{cases}$$

$$[A_2]_{ij} = \begin{cases} \bar{p}qr & k = i+Q-1;\ i = 1,2,\dots,s \\ \bar{p}q\bar{r} & k = i-1;\ i = 1,2,\dots,s \\ \bar{p}q & k = i-1;\ i = s+1, s+2, \dots, S \\ 0 & otherwise \end{cases}$$

$$[B_0]_{ij} = \begin{cases} \bar{p}\bar{r} & k = i;\ i = 0,1,2,\dots,s \\ \bar{p} & k = i;\ i = s+1, s+2, \dots, S \\ \bar{p}r & k = i+Q;\ i = 0,1,2,\dots,s \\ 0 & otherwise \end{cases}$$

$$[B_1]_{ij} = \begin{cases} p\bar{r} & k = i;\ i = 0,1,2,\dots,s \\ p & k = i;\ i = s+1, s+2, \dots, S \\ pr & k = i+Q;\ i = 0,1,2,\dots,s \\ 0 & otherwise \end{cases}$$

$$[B_2]_{ij} = \begin{cases} qr & k = i+Q-1;\ i = s \\ q & k = i-1;\ i = s+1, s+2, \dots, S \\ q\bar{r} & k = i-1;\ i = s \\ 0 & otherwise \end{cases}$$

$$[B_3]_{ij} = \begin{cases} \bar{q}\bar{r} & k = i;\ i = s \\ \bar{q}r & k = i+Q;\ i = s \\ \bar{q} & k = i;\ i = s+1, s+2, \dots, S \\ r & k = Q;\ i = 0 \\ \bar{r} & k = 0;\ i = 0 \\ 0 & otherwise \end{cases}$$

Also sub matrices are the parts the systems transition with respective transition probabilities. The matrices A_0, A_1, A_2, B_0, B_1, B_2 and B_3 are square matrices of order S + 1.

The stability of the system in long-run is generated by the condition that the Markov chain (X, L) is irreducible and aperiodic, which gives the positive recurrence. The transition matrix $A = A_0 + A_1 + A_2$ gives the condition for stability as follow.

Let π denote the steady-state (equilibrium) probability vector for the system, then the vector π satisfies the condition π A = π and πe = 1, where $\pi = (\pi_0, \pi_1, ..., \pi_S)$ and
$A_{ij} = A_0 + A_1 + A_2$

$$= \begin{cases} \bar{r} & k = i;\ i = 0 \\ \bar{q}\bar{r} & k = i;\ i = 1, 2, \ldots, s \\ \bar{q} & k = i;\ i = s+1, s+2, \ldots, S \\ q\bar{r} & k = i-1;\ i = 1, 2, \ldots, s \\ \bar{q}r & k = i+Q;\ i = 1, 2, \ldots, s \\ r & k = i+Q;\ i = 0 \\ qr & k = i+Q-1;\ i = 1, 2, \ldots, s \\ 0 & otherwise. \end{cases}$$

Steady state analysis of the systems $\{X_t, L_t: t \geq 0\}$ with tpm P has state space E = {0, 1, 2, ...,N} x {0, ..., S}.
The limiting distribution of Markov Chain is defined by
$\psi(i, k) = \lim_{n\to\infty} \Pr\{X_t = i,\ L_t = k | X_0, L_0\}$P
where ψ (i, k), denote steady state probability with initial state (i, k). If $\psi = (\psi_0, \psi_1, \psi_2, ..., \psi_N)$ satisfies the equals $\psi P = \psi$ and $\psi e = 1$, then, we have for $0 \leq i \leq N$, $\psi_i = (\psi(i, 0), \psi(i, 1), ..., \psi(i, S))$.

4. WAITING TIME DISTRIBUTION

Let $\{x_i, i \geq 0\}$ denote the stationary probability vector, that a customers sees i customers ahead of him at arrival epochs. Then
$x_0 = \psi_0 + q\,\psi_i$

and $x_i = \bar{q}\psi_i + q\,\psi_{(i+1)}$, $1 \leq i \leq N$ (1)

Since $\psi_i = \psi_0 R^i$, $i \geq 1$, where $R = A_2(I - A_1 - A_2G)^{-1}G$ is a minimal matrix, which is a solution of the matrix polynomial
$G = A_2 + A_1G + A_0G^2$.
The matrix G is order S + 1, which can be obtained by solving the recurrence equations
$G_{(i+1)} = (I - A_1)^{-1}(A_0 + A_2G_i^2)$, $0 \leq i \leq N$
where I is the identity matrix and G_0 is a matrix with all entries are zero.
From (1) We have
$x_0 = \psi_0[I + qR]$

and $x_i = \psi_0 R^i[\bar{q}\, I + qR]$, $1 \leq i \leq N$,

where $I_{(S+1) \times (S+1)}$ is an identity matrix.
Let $W_q(r)$ denote the probability that a customer has to wait exactly r slots before getting service and let
$W_q^*(z) = \sum_{r=0}^{\infty} z^r W_q(r)$,
Then,

$W_q(0) = x_0$ and $W_q(r) = \sum_{i=1}^{r} y_i \varphi^i(r)$, $r \geq 1$ (2)

where

$$\begin{aligned} \phi^i(r) &= (\bar{q})^{r-1}q,\ r \geq 1 \\ \phi^i(i) &= q^i,\ i \geq 1 \\ and\ \phi^i(r) &= q\phi^{(i-1)}(r-1) + \bar{q}\phi^{(i)}(r-1)\ r \geq i+1,\ i \geq 2 \end{aligned} \quad (3)$$

From (2) and (3), we get
$W_q(r) = q\psi_0 R(\bar{q}I + qR)^r,\ r \geq 1.$

5. SYSTEM PERFORMANCE MEASURE

In this section we derive some system performance measures in the rub.

(a) mean inventory level
(b) mean reorder rate
(c) mean shortage rate and
(d) mean balking rate.
(e) mean number of customers in the system

(a) Mean Inventory Level

Let $\bar{I}$ denote the mean inventory level in the service facility system in the long-run. Suppose ψ(i,k) denotes the steady state probability that there are i customers waiting and k items in stock for service completion purpose. Then

$$\bar{I} = \sum_{i=0}^{N}\sum_{k=1}^{S} k\psi_{(i,k)}.$$

(b) Mean Reorder Rate

Let β_R denote the mean reorder rate, then

$$\beta_R = q\sum_{i=1}^{N}\psi_{(i,s+1)}$$

(c) Mean Shortage Rate

Let η_B denote the mean shortage rate for the system. Then

$$\eta_B = q\sum_{i=1}^{N}\psi_{(i,1)}$$

(d) Mean Balking Rate

Mean balking rate for the system when the waiting space is full is given by

$$\alpha_B = p\sum_{k=0}^{S}\psi_{(N,k)}$$

(e) Mean number of customers in the system

Mean number customers in the system is given by

$$\alpha_w = \sum_{i=1}^{N}\sum_{k=0}^{S}\psi_{(i,k)}$$

6. MARKOV DECISION PROCESS

MDP is a versatile tool developed by Howard [12] and Bellman [1]. Stochastic dynamic programming is the technical name for MDP. Discrete time MDP is the formal dynamical system which has application in machine maintenance, inventory control and communication systems. Any MDP formulation has 5 components (a) Time epochs (b) State space (c) Decision sets (actions) (d) Transition probabilities and (e) cost functions.

$$(T, E, A_e, p(\cdot \mid \cdot, a), c(\cdot \mid \cdot, a))$$

MDP FORMULATION

Time epochs.

The time epoch be 0, η, 2η, 3η, ... on the infinite time line [0,∞).

State Space.

The states of the MDP lie in the space

$$E = \{e = (i,k) : 0 \le i \le N;\ 0 \le k \le S\}$$

Action set.

The action set consist of three actions 0, 1 and 2, where 0 means no order, 1 means order for Q items Q = S -s > s and 2 means compulsory order for Q items at inventory level is zero.

Then

$$A_e = \begin{cases} \{0\} & if \quad s+1 \le k \le S,\ 0 \le i \le N \\ \{0,1\} & if \quad 1 \le k \le s,\ 0 \le i \le N \\ \{2\} & if \quad k = 0,\ 0 \le i \le N \end{cases}$$

The action set is

$$A = \bigcup_{e \in E} A_e = \{a :\ a = 0,1,2\}$$

Hence the state space E is partitioned as E = $E_1 \cup E_2 \cup E_3 \cup \ldots$

Transition Probability.

The Markov Chain $\{(X_t, L_t): t \geq 0\}$ with discrete time space $t \in \{0, \eta, 2\eta, 3\eta, \ldots\}$ has state transition probabilities $p_{((i,k),(j,l))}$ and tpm P. The detail of transition are described in section .

As a result of choosing action $a \in A_m$ in state m at decision epoch t,

(i) the decision maker incurs a cost $C_t(m, a)$ and

(ii) the system state at next decision epoch is determined by the transition probability $p_t(\cdot|m, a)$.

In our SFS the probability distribution is given by $p_t((j, l)|(i, k), a)$.

Thus $P_t = (p_t((j, l)|(i, k), a))$ is a square matrix of order (N + 1)(S + 1).

Reward/ Cost Function.

The real valued function $C_t((j, l), a)$ defined for $(j, l) \in E$ and $a \in A_{(i,l)}$, denote the cost incurred in period t.

$$C((j,l),a) = h\bar{I} + c_1\beta_R + c_2\eta_B + c_3\alpha_B + w\alpha_w$$

where h- Inventory holding cost/ unit item/ unit time, c_1- Cost/ order, c_2- Shortage cost in terms of inventory level, c_3- Cost/ customer lost, w- Waiting cost of the customer/ unit time. $\bar{I}$, β_R, η_B, α_B, and α_w are system performance measures of the service facility system considered.

The collection of objects $\{T, E, A_e, p_t(\cdot|s, a), c_t(s, a)\}$ is referred as Markov Decision Process.

7. MDP PROCEDURE

Fix a stationary policy R for the considered system. Under policy R each time action a = $R^*_{(i,k)}$ is taken whenever the system is in state (i, k) at decision epoch t.

The process $\{(X_t, L_t): t = 0, 1, \ldots\}$ describing the state of the system at the decision epochs is a Markov chain with one-step transition $p^{(j,l)}_{(i,k)}$p(j,l)(i,k)($R_{(i,k)}$), (i, k), (j, l) $\in$ E when policy R is used.

Then the n- step transition probability of the Markov chain is of the form

$$p^{(n)}_{(i,k),(j,l)}(R) = P\{(X_n, L_n) = (j,l)|(X_0, L_0) = (i,k)\}, \quad (i,k),(j,l) \in E$$

for n = 1, 2, 3, ...

By Chapman-Kolmogorov equation

$$p^{(n)}_{(i,k),(j,l)}(R) = \sum_{(u,v)\in E} p^{n-1}_{(i,k)(u,v)}(R)p_{(u,v)(j,l)}(R_{(u,v)})$$

where $p^{(0)}_{(i,k)(j,l)} = 1$ for $(j,l) = (i,k)$ and $p^{(R)}_{(i,k)(j,l)} = 0$ for $(j,l) \neq (i,k)$.

Define the expected value function

$V_n((i, k),R)$ = the total expected costs over the first n decision epoch when the initial state is (i,k) and policy R is used

$$= \sum_{m=0}^{n-1} \sum_{(j,l)\in E} p^{(m)}_{(i,k),(j,l)}(R)c_{(j,l)}(R_{(j,l)})$$

The average cost function $g_i(R)$ by

$$g_i(R) = \lim_{n\to\infty} \frac{1}{n} V_n((i,k),R), \; (i,k) \in E.$$

Next we discuss the Policy-iteration algorithm under unichain assumption.

For each stationary policy the associated Markov Chain $\{X_n, L_n\}$ has no two disjoint closed sets.

8. POLICY- ITERATION ALGORITHM

Step 0. (Initialization)

Choose a stationary policy R for the periodic review based inventory control in service facility system.

Step 1. (Value Determination Step)

For the current policy R, compute the unique solution $(q(R), v_e(R))$ to the following linear equations

$$v_e = c_e(R_e) - q + \sum_{e'\in E} p_t(e'|e)v_{e'}, \quad e = (i,k) \in E$$

$$v_e = 0, \quad \text{where } e = (i,k) \text{ is arbitarily chosen state in } E$$

$s \setminus s'$	(0,0)	(0,1)	(0,2)	(0,3)	(1,0)	(1,1)	(1,2)	(1,3)	(2,0)	(2,1)	(2,2)	(2,3)	(3,0)	(3,1)	(3,2)	(3,3)	(4,0)	(4,1)	(4,2)	(4,3)
(0,0)	0.3	0	0.3	0	0.2	0	0.2	0	0	0	0	0	0	0	0	0	0	0	0	0
(0,1)	0	0.3	0	0.3	0	0.2	0	0.2	0	0	0	0	0	0	0	0	0	0	0	0
(0,2)	0	0	0.6	0	0	0	0.4	0	0	0	0	0	0	0	0	0	0	0	0	0
(0,3)	0	0	0	0.6	0	0	0	0.4	0	0	0	0	0	0	0	0	0	0	0	0
(1,0)	0	0	0	0	0.3	0	0.3	0	0.2	0	0.2	0	0	0	0	0	0	0	0	0
(1,1)	0.18	0	0.18	0	0.12	0.12	0.12	0.12	0	0.08	0	0.08	0	0	0	0	0	0	0	0
(1,2)	0	0.36	0	0	0	0.24	0.24	0	0	0	0.16	0	0	0	0	0	0	0	0	0
(1,3)	0	0	0.36	0	0	0	0.24	0.24	0	0	0	0.16	0	0	0	0	0	0	0	0
(2,0)	0	0	0	0	0	0	0	0	0.3	0	0.3	0	0.2	0	0.2	0	0	0	0	0
(2,1)	0	0	0	0	0.18	0	0.18	0	0.12	0.12	0.12	0.12	0	0.08	0	0.08	0	0	0	0
(2,2)	0	0	0	0	0	0.36	0	0	0	0.24	0.24	0	0	0	0.16	0	0	0	0	0
(2,3)	0	0	0	0	0	0	0.36	0	0	0	0.24	0.24	0	0	0	0.16	0	0	0	0
(3,0)	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0	0.3	0	0.2	0	0.2	0
(3,1)	0	0	0	0	0	0	0	0	0.18	0	0.18	0	0.12	0.12	0.12	0.12	0	0.08	0	0.08
(3,2)	0	0	0	0	0	0	0	0	0	0.36	0	0	0	0.24	0.24	0	0	0	0.16	0
(3,3)	0	0	0	0	0	0	0	0	0	0	0.36	0	0	0	0.24	0.24	0	0	0	0.16
(4,0)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0.5	0
(4,1)	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0	0.3	0	0	0.2	0	0.2
(4,2)	0	0	0	0	0	0	0	0	0	0	0	0	0	0.6	0	0	0	0	0.4	0
(4,3)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.6	0	0	0	0	0.4

Table 1: Prespecified transition probabilities of the system

Step 2. (Policy Improvement)
For each state e = (i, k) ∈ E determine the actions yielding

$$\arg\min_{a \in A_e} \{C_e(a) - q + \sum_{e' \in E} p_t(e'|e,a) v_{e'}(a)\}$$

The new stationary policy R^* is obtained by choosing $R_e^* = a_e$
Step 3. (Convergence Test)
If the new policy R^* = R, the old one. Then the process of searching stops with policy R. Otherwise go to Step 1 with R replaced by new R'.

9. NUMERICAL EXAMPLE

Consider a service facility system with inventory having limited waiting space, N = 4 and maximum inventory S = 3.
It is assumed that the arrival of customers to the systems follows a Bernoulli process with probability p = 0.4 and that of non-arrival $\bar{p}$ = 1-p. The service time of a customer are geometrically distributed with probability q = 0.6 and $\bar{q}$ = 1- q. We adopt (s, S) ordering policy which means a fixed quantity of Q items is ordered whenever the inventory level is to or less than a free fixed level s, where Q = S -s, (Q > s).

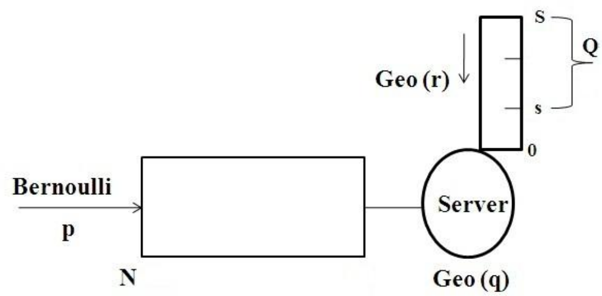

Figure 1: Discrete time Service Facility System with Inventory Management

The ordered items are delivered after a lead time having geometric distribution with parameter r = 0.5 and $\bar{r}$ = 1- r.
Assume the waiting cost w = 0.1 per customer/ period, the shortage cost c_2 = 0.15 in terms of inventory level, holding cost h = 0.1, balking cost c_3 = 0.15 per customer lost and ordering cost for Q items is c_1 = 0.2.

Initiate the policy iteration algorithm with the arbitrary policy:

$$R^{(0)} = (0, 0, 0, 2, 0, 0, 0, 2, 0, 0, 0, 2, 0, 0, 0, 2, 0, 0, 0, 2),$$

which prescribes the compulsory order when the inventory level is zero.

After two iterations we obtained the optimal policy.

$$R^{(2)} = (0, 0, 1, 2, 0, 0, 1, 2, 0, 0, 1, 2, 0, 0, 0, 2, 0, 0, 0, 2)$$

and this policy stabilizes for further iterations.

Thus the optimal policy is given by

$$R^{(*)} = (0, 0, 1, 2, 0, 0, 1, 2, 0, 0, 1, 2, 0, 0, 0, 2, 0, 0, 0, 2).$$

10. CONCLUSION

In this paper we investigated the problem of inventory control in a service facility system. For a discrete time service facility system we applied MDP tool to get an optimal reordering rule (policy) R^*, which yields lowest average cost during system implementation. A numerical example is also provided to substantiate the result obtained.

Ordering policy adopted in this study is (s, S) type with fixed order quantity Q = S-s >s. But other policies (i) varying order quantities at different inventory levels and (ii) modified ordering (up to S) policies also deserve further investigations in future.

REFERENCES

[1] BELLMAN, R. E. (1957): **Dynamic Programming**, Princeton University Press, Princeton .

[2] BERMAN, O., E.H. KAPLAN and D.G. SHIMSHAK (1993): Deterministic approximations for inventory management at service facilities, **IIE Transactions**, 25, 98-104,

[3] BERMAN O. and K. P. SAPNA (2000): Inventory management at service facilities for systems with arbitrarily distributed service times, **Stochastic Models**, 16, 343-360,

[4] BLACKWELL, D. , (1962): Discrete dynamic programming, **Ann. Math. Statist.**, 33, 719-726.

[5] DE GHELLINCK, G. , (1960): Les problemes de decisions sequentielles, **Cahiers Centre Etudes Recherche Oper.,** 2, 161-179.

[6] DENARDO E.V., and B.L. FOX (1968): Multi-chain Markov renewal programs, **SIAM J. Appl. Math.** 16, 468-487.

[7] DERMAN, C. (1970): Finite State Markovian Decision Processes, Academic Press, New York

[8] GARCIA J. and O. CASALS, (1996): Priorities in ATM Networks, **NATO Advanced Research Workshop.**

[9] GOLABI, K., R.B. KULKARNI and C.B. WAY(1982): A statewide pavement management system, **Interfaces**, 12, 5-21,

[10] HORDIJK A: and L.C.M. KALLENBERG (1979): Linear programming and Markov decision chains, **Management Sci.,** 25, 352-362

[11] HORDIJK A: and L.C.M. KALLENBERG (1984): Constrained undiscounted stochastic dynamic programming, **Math. Operat. Res**., 9, 276-289.

[12] HOWARD R. A. (1960): **Dynamic Programming and Markov Processes**, Cambridge, Massachusetts: M. I. T. Press,

[13] KAWAI H. (1983): An optimal ordering and replacement policy of a Markovian degradation system under complete observation, part I, J**. Operat. Res. Soc. Japan**, 26, 279-290.

[14] Le Boudec .J. Y. (1991): An Efficient Solution Method for Markov models of ATM Links with Loss Priorities, **IEEE JSAC**, 9(3).

[15] LIAO K. Q. (1995): Queueing Analysis of Partial Buffer Sharing with Markov Modulated Poisson Inputs, In **ITC-14,** 55-64, Elsevier, Amsterdam.

[16] LIN A.Y.-M. and J. SILVESTER (1991): Priority Strategies and Buffer Allocation Protocols for Traffic Control at an ATM Integrated Broadband Switching System, **IEEE JSAC**, 9(9).

[17] MAHESWARI P. and C. ELANGO (2017): Optimal Admission and Service Control in a Discrete time Service Facility Systems: MDP Approach, **Annals of Pure and Applied Mathematics**, 15, 315-326.

[18] MANNE A. (1960): Linear programming and sequential decisions, **Management Sci**., 6, 259-267.

[19] Odoni A. (1969): On finding the maximal gain for Markov decision processes, **Operat. Res**., 17, 857-860.

[20] RUBIN I. and Z, TSAI (1991): Message Delay Analysis of Multiclass Priority TDMA, FDMA and Discrete Time Queueing Systems, **Trans. Info. Theory**, IT-35, 637-647.
[21] SELVAKUMAR, C., P. MAHESWARI and C. ELANGO (2017): Discrete MDP problem with Admission and Inventory Control in Service Facility Systems, **International Journal of Computational and Applied Mathematics,** 12, 478-493.
[22] STENGOS D. and L.C. THOMAS (1980): The blast furnaces problem, **Eur. J. Oper. Res**., 4, 330-336.
[23] TAKAHASHI Y. and O. HASHIDA (1991): Delay analysis of discrete time priority queue with structured inputs, **Queueing Systems**, 8, 149-164.
[24] TIJMS H. C. and F.A. VAN DER DUYN SCHOUTEN (1985): A Markov decision algorithm for optimal inspections and revisions in a maintenance system with partial information, **Eur. J. Oper. Res**., 21, 245-253.
[25] Veinott A.F. (1996): On finding optimal policies in discrete dynamic programming with no discounting, **Ann. Math. Statist**., 37, 1284-1294.
[26] VINOD S. and N. D. GANGADHAR (1997): Discrete time queues with priorities, **Queueing Systems**, 25, 281-305.
[27] Wagner D. (1995): Analysis of multi-sever model with non-preemptive priorities and non-renewal input, In J. Labetoulle and J.W. Roberts, editors, **ITC-14**, 757-767, Elsevier, Amsterdam.
[28] WHITE J. (1985): Real Applications of Markov Decision Processes, I**NFORMS**, 15, 73-83.
[29] YADAVALLI, V. S. S. , B. SIVAKUMAR, G. ARIVARIGNAN, and O. ADETUNJI (2011): A multi-server perishable inventory system with negative customer, **Computers and Industrial Engineering**, 61, 254-273.