International Journal of Engineering Science Invention (IJESI)
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org ||Volume 7 Issue 8 Ver II || Aug 2018 || PP 55-62

MDP in Supply Chain: Optimal Inventory Control System

P. K Santhi[1] and C. Elango[2]
[1]*Department of Mathematics, Jayaraj Annapackiam College for Women (Autonomous), Periyakulam, Theni(D.T),Tamilnadu, India.*
[2]*Department of Mathematical Sciences Cardamom Planters' Association College, Bodinayakanur, Tamilnadu, India*
Corresponding Author: P. K Santhi

Abstract: *In this article, we study a MDP based inventory control in a two stage Supply Chain. We consider a two stage Supply Chain having Distribution Center (DC) and Retail Vendor (RV) with their respective inventory systems and service facility. A two dimensional MDP is formulated and the optimal decision policy is obtained by Linear Programming technique. Some instances of a numerical is produced to study the behavior of the system.*

Keywords: *Markov Decision Processes, Supply Chain Management, Optimal inventory policy, Linear programming procedure, Two-stage Inventory control system*

Date of Submission: 29-07-2018 Date of acceptance: 22-08-2018

I. Introduction

Supply chain can be defined as the management of flow of products and services, which begins from the products point and ends at the consumption point(retailer). This process comprises of movement and storage of raw materials that are involved in work progress, inventory and fully furnished goods. Supply Chain exists in both service and manufacturing organizations, but the complexity of the chain may vary greatly from industry to industry.

Inventory decision is an important component of the supply chain management, because Inventories exist at each and every stage of the supply chain as raw material or semi-finished or finished goods. They can also be as Work-in-process between the stages or stations. Since holding of inventories can cost anywhere between 20% to 40% of their value, their efficient management is critical in Supply Chain operations

The usual objective for a multi-echelon inventory model is to coordinate the inventories at various echelons so as to minimize the total cost associated with the entire multi-echelon inventory system. It might also be a suitable objective when certain echelons are managed by either the suppliers or the retailers of the company. The reason is that a key concept of supply chain management is that a company should strive to develop an informal partnership relation with its suppliers and retailers that enable them jointly to maximize their total profit.

It would be appropriate to say that information technology is a vital organ of supply chain management. With the advancement of technologies, new products are being introduced within fraction of seconds increasing their demand in the market. Let us study the role of information technology in supply chain management briefly.

Multi-echelon inventory system has been studied by many researchers and its applications in supply chain management has proved worthy in recent literature. As supply chains integrates many operators in the network and optimize the total cost involved without compromising the customer service efficiency.

Continuous review models of multi-echelon inventory system in 1980's concentrated more on repairable items in a Depot-Base system than as consumable items(see Graves, Moinzadeh and Lee). All these models deal with repairable items with batch ordering. Sven Axsäter proposed an approximate model of inventory structure in SC. One of the oldest papers in the field of continuous review multi-echelon inventory system is a basic and seminal paper written by Sherbrooke in 1968. He assumed (S-1,S) polices in the Depot-Base systems for repairable items in the American Air Force and could approximate the average inventory and stock out level in bases. Seifbarghy,and Jokar, analyzed a two echelon inventory system with one warehouse and multiple retailers controlled by continuous review (r,Q) policy. A ccomplete review was provided by Benita M. Beamon (1998). The supply chain concept grow largely out of two-stage multi-echelon inventory models, and it is important to note that considerable research in this area is based on the classic work of Clark and Scarf(1960). In the case of continuous review perishable inventory models with random lifetimes for the items, most of the models assume instantaneous supply of order. The assumption of positive lead times further increases the complexity of the analysis of these models and hence there are only a limited number of models dealing with positive lead-times. A continuous review perishable inventory system at Service Facilities was studied by Elango(2001). A continuous review(s,S) policy with positive lead times in two-echelon Supply Chain was considered by K.Krishnan and

C.Elango. Service facilities in the inventory in supply chain management is a quiet new area

In this chapter we considered an inventory system model maintained in a service facilities at Retailer vendor in tandem supply chain having Retailer vendor(RV) and Distribution Centre(DC). One item from inventory at RV is used to serve the customer. (s, S) policy is adopted at Retail Vender node for inventory replenishment.

II. Model Formulation

We consider a Supply Chain system consist of Distribution Centre(DC), Retail vendor(RV) with service facility and inventory is maintained at both DC and RV nodes. For every demand at the retailer node (RV) an item is supplied only after a exponential service time with parameter γ. The waiting space in the retailer node has maximum capacity N. An arriving customer seeing N customers in the system leaves.

Inventory policy adopted at RV node is (s, S) type in which order for Q=S-s>s items are placed when the inventory level reaches the prefixed level s, and lead time is exponentially distributed with parameter $\mu(>0)$. Demand at RV node follows a Poisson process with parameter $\lambda(>0)$. At DC, items are packed as Q items in one pocket with maximum inventory level nQ (n pockets). The ordering policy at DC is of (0,nQ) type where the inventory level reach 0, instantaneous replenishment nQ= M items is made. Deterministic Markov Decision policy is used solve the problem of MDP.

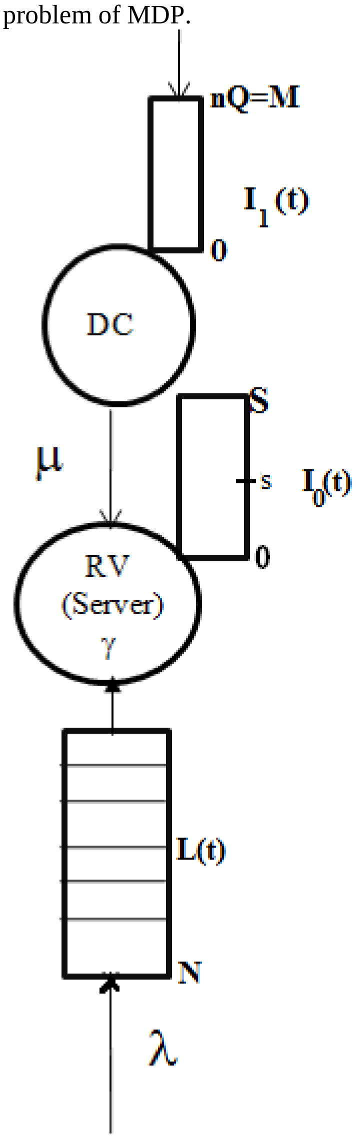

Fig (1)

Let $I_0(t)$ and L(t) denote the inventory level and the number of customers in the system at time t. Then $\{(I_0(t), L(t)) : t \geq 0\}$ is a finite two dimensional stochastic process with state space , $E_1 \times E_2$,

where $E_1 = \{0, 1, 2, \ldots ,S\}$ and $E_2 = \{0, 1, 2, \ldots,N\}$.

Decision Sets:

The reordering decisions taken at each state of the system $(i,q) \in E$, where, $I(t) = i$ and $X_0(t) = q$.
Let A_i (i =1, 2, 3) denotes the set of possible actions where, $A_1 = \{0\}$, $A_2 = \{0, 1\}$, $A_3 = \{2\}$, $A = A_1 \cup A_2 \cup A_3$, 0 represents 'no order', 1 means reorder for 'Q = S-i' items at level i and 2 means compulsory order for S items when inventory level is zero.
Suppose Đ denote the class of all stationary policies, then a policy *f* (sequence of decisions) can be defined as a function $f: E \to A$, given by

$$f(i,q) = \begin{cases} \{0,1\} & 1 \le i \le s, q \in E_2 \\ \{0\} & s+1 \le i \le S \quad , q \in E_2 \\ \{2\} & i = 0, q \in E_2 \end{cases}$$

Objective of the problem is to find the optimal reorder level s so that the long run expected total cost rate is minimum.

Notations and Assumptions:

1. $E_1 \times E_2 = E$ *is the state space of the Stochastic Process* $\{(I(t), L(t)) : t \ge 0\}$, *where* $E_1 = \{0,1,2,,...,S\}$ *and* $E_2 = \{0,1,2,,...,\mathrm{N}\}$
2. $A_{(i,q)}$ *– decision set corresponding to state* $(i,q) \in E$.
3. $C_{(i,q)}(a)$ – cost *occured when action a is taken at state* (i,q).
4. $p_{(i,q)}^{(j,r)}(a)$ *– the transition probability from state* (i,q) *to state* (j,r). *when action a is taken at state* $(i,q) \in E$.
5. *Inventory levels are reviwed at the time of service completion epochs.*
6. Reordering policy is (s, S): Q = S – s items ordered when the inventory level reaches s (prefixed level), where $0 \le s \le S$.
7. D- the class of stationary policies.

III. Analysis

Let R denote the stationary policy, which is time invariant and Markovian Policy (MR). From our assumptions it can be seen that $\{(I_0(t), L(t)) : t \ge 0\}$ is denoted as the controlled process $\{(I_0^R(t), L^R(t)) : t \ge 0\}$ when policy R is adopted. Since the process $\{(I_0^R(t), L^R(t)) : t \ge 0\}$ is a Markov Process with finite state space E. The process is completely Ergodic, if every stationary policy gives raise to an irreducible Markov chain. It can be seen that for every stationary policy $f \in F, \{I_0^f, L^f\}$ is completely Ergodic and also the optimal stationary policy R^* exists, because the state and action spaces are finite.

A Deterministic Markov decision rule from the class F is equivalent to the function $f: E \to A$ given by $P_{d_t} \in \wp(A_r), \mathrm{j} \in E_1,$ where d_t is the Markovian randomized decision rule for $t \in T$. We denote the set of decision rules at time t by D_t^{MR}.

If d_t is the Markovian randomized decision rule, the expected reward satisfies the transition probability relations.

$$p_t((j,r) \mid (i,q), d_t(i,q)) = \sum_{a \in A_s} p_t((j,\mathrm{r}) \mid (\mathrm{i,q}), \mathrm{a}) p_{\mathrm{d}_t(i,q)}(a).$$

$$r_t(i,q), d_t(i,\mathrm{q}) = \sum_{a \in A_s} r_t(\mathrm{i,q,a}) p_{\mathrm{d}_t(i,q)}(a).$$

For Markovian $f \in f^{MR}$, d_t depends on history analysis through the current state of the process $(i,q) \in E$ so that $p^f\{Y_t = a \mid Z_t = h_t\} = P_{d_t(h_t)}(a)$ where Y_t – denote the action at time t and the history process Z_t defined by $Z_1(w) = s_1$ and $Z_t(w) = \{s_1, s_2, s_3, \ldots, s_t\}$ for $1 \le t \le N, \mathrm{N} \le \infty$

Randomized Markovian Policy f

Order size	Q=S-s	Q+1=S-s+1	…	Q+s=S
Probability	p_s	p_{s-1}	…	p_0

f^{MR} is the randomized Markovian policy. Under this policy f an action $a \in A(j)$ is chosen with probability $f_a(j)$, whenever the process is in state $j \in E$.

Whenever $f_a(j) = 0 \, or \, 1$, the stationary randomized policy f reduces to a familiar stationary policy.

3.1 Steady State Analysis

Let $\left\{\left(I_0^R(t), L^R(t)\right): t \geq 0\right\}$ denote the process. $\left\{\left(I(t), L(t)\right): t \geq 0\right\}$ in which R is the policy adopted from our assumptions made in the previous section. The controlled process $\{I^R, L^R\}$ where R is the randomized Markovian policy in a Markov process. Under the randomized policy, f the expected long run total cost rate when policy f is adopted is given by

$$C^f = h\bar{I}^f + c_1 w^f + c_2 \eta_a^f + g\eta_b^f + \beta\eta_c^f \qquad (1)$$

h - holding cost / unit item / unit time
c_1 – waiting cost / customer / unit time
c_2 – reordering cost / order
g - balking cost / customer
β - service cost / customer
$\bar{I}^f$ - mean inventory level
$\bar{w}^f$ - mean waiting time in system
η_a^f - reordering rate
$\eta_b{}^f$ - balking rate
$\eta_c{}^f$ - service completion rate

Our objective is to find an optimal policy f^* for which $C^{f^*} \leq C^f$ for every MR policy in f^{MR}

For any fixed MR policy $f \in f^{MR}$ and $(i,q),(j,r) \in E,$ define

$$\Phi_{iq}^f(j,r,t) = Pr\left\{I_0{}^f(t) = j, L^f(t) = r \mid I_0{}^f(0) = i, L^f(0) = q\right\}, \; (i,q),(j,r) \in E. \qquad (2)$$

Now $\Phi_{iq}^f(j,r,t)$ satisfies the Kolmogorov forward differential equation $\Phi_i{}'(t) = \Phi(t)A,$ where A is an infinitesimal generator of the Markov process $\{(I_0{}^f(t), L^f(t)): t \geq 0\}.$

For each MR policy f, we get an irreducible Markov chain with the state space E and actions space A which are finite,

$$\Phi^f(j,r) = \lim_{t\to\infty} \Phi_{iq}^f(j,r;t)$$ exists and is independent of initial state conditions.

This implies the balance equations (5) – (16) given below. Transition in and out of a states give a system of equations.
Consider the typical state (j, r) that lies in the range $s+1 \leq j \leq S-1$; $1 \leq r \leq N-1$. When (j, r) lies in this range, there is no order pending and hence transition out of this state can be due to either by demand or a service completion. The corresponding balance equation is given by equation (7).
A service completion in state (j+1, r+1) will decrease both inventory level and number of customers by one unit, thus transition made to state (j, r).
When one customer arrives and enters the system (r < N) at state (j, r -1), the new state is (j, r). Considering two different ways of reaching state (j, r) and are reflected on the right hand side of Eq. (7).

Fig (2) represent the in-rate and out-rate flow diagram of the system states.

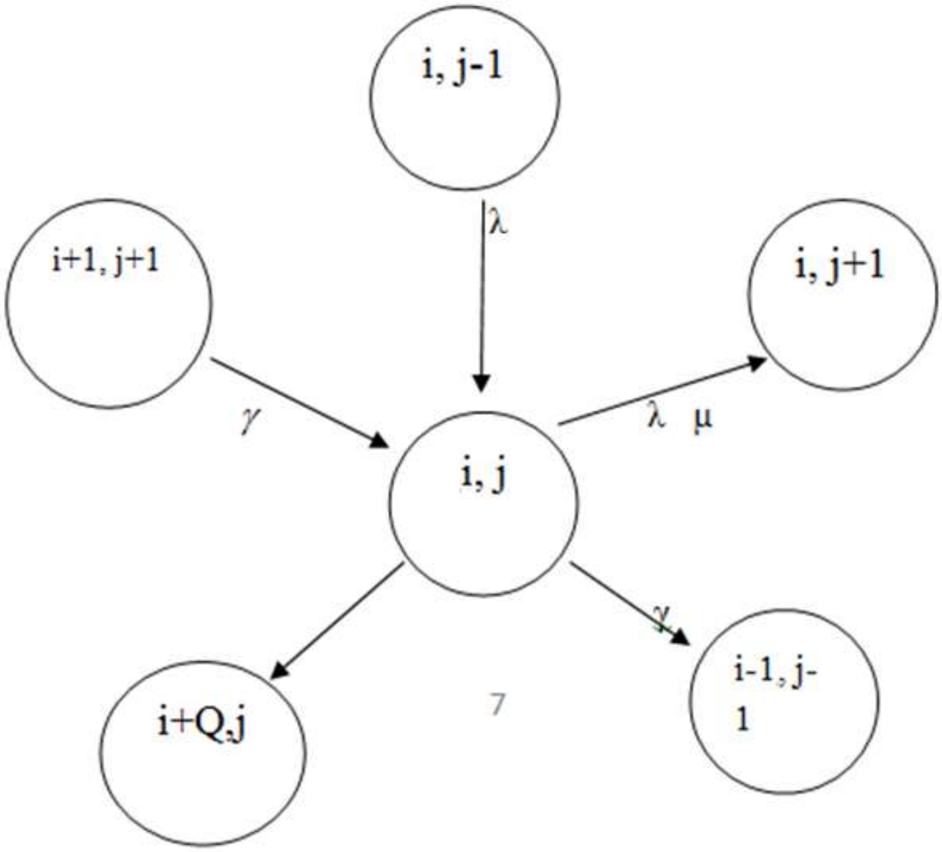

Fig (2)

Now the system of equations can be written in order as follows,

$$\lambda\Phi^{f}(\mathrm{S},0)=\mu\sum_{j=0}^{s} p_j\Phi^{f}(\mathrm{j},0) \tag{3}$$

$$(\lambda+\gamma)\Phi^{f}(\mathrm{S},\mathrm{r})=\mu\sum_{j=0}^{s} p_j\Phi^{f}(\mathrm{j},\mathrm{r})+\lambda\Phi^{f}(\mathrm{S},\mathrm{r}-1),\quad 1\le \mathrm{r}\le M-1 \tag{4}$$

$$\gamma\,\Phi^{f}(\mathrm{S},N)=\mu\sum_{j=0}^{s} p_j\Phi^{f}(j,N)+\lambda\Phi^{f}(\mathrm{S},N-1) \tag{5}$$

$$\lambda\Phi^{f}(\mathrm{j},0)=\gamma\Phi^{f}(\mathrm{j}+1,1),\quad \mathrm{s}+1\le j\le S-1 \tag{6}$$

$$(\lambda+\gamma)\Phi^{f}(\mathrm{j},\mathrm{r})=\gamma\,\Phi^{f}(\mathrm{j}+1,\mathrm{r}+1)+\lambda\Phi^{f}(\mathrm{j},\mathrm{r}-1),\ \mathrm{s}+1\le j\le S-1;1\le \mathrm{r}\le N-1 \tag{7}$$

$$\gamma\,\Phi^{f}(\mathrm{j},N)=\lambda\Phi^{f}(\mathrm{j},N-1),\quad \mathrm{s}+1\le j\le S-1 \tag{8}$$

$$(\lambda+\mu p_j)\Phi^{f}(\mathrm{j},0)=\gamma\,\Phi^{f}(\mathrm{j}+1,1)\ ,\ 1\le j\le s \tag{9}$$

$$(\lambda+\mu p_j+\gamma)\Phi^{f}(\mathrm{j},r)=\gamma\Phi^{f}(\mathrm{j}+1,r+1)+\lambda\Phi^{f}(\mathrm{j},r-1), \tag{10}$$
$1\le j\le s;1\le \mathrm{r}\le N-1$,

$$(\mu p_j+\gamma)\Phi^{f}(\mathrm{j},N)=\gamma\Phi^{f}(\mathrm{j},N-1),1\le j\le s\ , \tag{11}$$

$$(\lambda+\mu p_0)\Phi^{f}(0,0)=\gamma\,\Phi^{f}(1,1)\,, \tag{12}$$

$$(\lambda+\mu p_0)\Phi^f(0,\mathrm{r})=\gamma\ \Phi^f(1,\mathrm{r}+1)+\lambda\Phi^{f}(0,\mathrm{r}-1),\quad 1\le \mathrm{r}\le N-1\ , \tag{13}$$

$$\mu p_0\Phi^f(0,N)=\lambda\Phi^{f}(0,N-1) \tag{14}$$

The above set of equations together with the condition $\sum_{(j,r)\in E}\Phi^f(j,r)=1$ (15)

,determine the steady-state probabilities uniquely.

3.2 System Performance Measures

The probability ϕ^f(,j r) also gives the value that the long-run fraction of time the system is in the state (j, r),

1.The expected inventory level in the system is given by

$$\overline{I^f}=\sum_{j=1}^{S} j\sum_{r=0}^{N}\Phi^f(j,r) \tag{16}$$

2. The mean waiting time is given by

$$\overline{W}^{\ f}=\sum_{r=1}^{N}\frac{r}{\gamma}\sum_{j=0}^{S}\Phi^f(j,r)+\sum_{k=0}^{s}\frac{1}{\mu p_k}\sum_{m=1}^{[N/S]}\sum_{r=1}^{mS}m\Phi^f(j,r) \tag{17}$$

3. The reorder rate is given by

$$\eta_a{}^f=\mu\sum_{r=0}^{N}\ \sum_{j=0}^{s}p_j\Phi^f(j,r). \tag{18}$$

4. The balking rate is given by

$$\eta_b{}^f=\lambda\sum_{j=0}^{S}\Phi^f(j,N) \tag{19}$$

5.The service completion rate is given by

$$\eta_c{}^f=\gamma\sum_{r=1}^{N}\ \sum_{j=1}^{S}\Phi^f(j,r) \tag{20}$$

Hence the average cost rate of the system is given by

$$C^f=h\sum_{j=1}^{S}j\sum_{r=0}^{N}\Phi^f(j,r)+\frac{c_1}{\gamma}\ \sum_{r=1}^{N}r\sum_{j=0}^{s}\ \Phi^f(j,r)+\frac{c_1}{\mu}\sum_{m=1}^{[N/S]}\sum_{r=1}^{mS}\sum_{j=0}^{S}\frac{m}{p_j}\Phi^f(j,r)+c_2\mu\sum_{r=1}^{N}r\sum_{j=0}^{s}p_j\Phi^f(j,r)$$
$$+g\lambda\sum_{j=0}^{S}\Phi^f(j,M)+\beta\gamma\sum_{j=1}^{S}\sum_{r=1}^{N}\ \Phi^f(j,r) \tag{21}$$

IV. Linear programming problem

4.1 LPP Formulation

In this section we propose a LPP model within a MDP framework.

First define the variables, D (j, r, k) as a conditional probability such that

D(j, r, k) = Pr {decision is k | state is (j, r)} --------(22)

Since $0 \le$ D (j, r, k) ≤ 1, this is compatible with the randomized time invariant Markovian policies.

Here, the Semi – Markovian decision problem can be formulated as a linear programming problem.

Hence

$0 \le$ D (j, r, k) ≤ 1 and $\sum_{k\in A=\{0,1,2\}} D(\mathrm{j,r,k})=1,$ **$0 \le$ r $\le$ N; $0 \le$ j $\le$ S.**

For the reformulation of the MDP as LPP, we define another variable
y (j, r, k) as follows.

$$y(\mathrm{j,r,k})=\mathrm{D(j,r,k)}\Phi^f(\mathrm{j,r}). \tag{23}$$

From the above definition of the transition probabilities

$$\Phi^{f}(\mathrm{j,r}) = \sum_{k\in A} y(\mathrm{j,r,k}),\ (\mathrm{j,r})\in \mathrm{E},\ \mathrm{k}\in A=\{0,\ 1,\ 2\} \qquad (24)$$

Expressing $P^{\pi}(\mathrm{j,r})$ in terms of $y(\mathrm{j,r,k})$, the expected total cost rate function (21) is given by

Minimize

$$C = h\sum_{k\in A}\sum_{j=1}^{S} j\sum_{r=1}^{N} \mathrm{y(j,r,k)} + h\sum_{j=1}^{S} j.y(\mathrm{j},0,0) + c_1\mu\sum_{k\in A}\sum_{r=1}^{N}\sum_{j=0}^{s} p_j.y(\mathrm{j,r,k})$$

$$+c_2\sum_{k\in A}\sum_{r=1}^{N}\frac{r}{\gamma}\sum_{j=0}^{S}\mathrm{y(j,r,k)} + \frac{c_2}{\mu}\sum_{k\in A}\sum_{m=1}^{\left[N/s\right]}\sum_{r=1}^{ms}\sum_{j=0}^{S}\left(\frac{m}{p_k}\right)\mathrm{y(j,r,k)}$$

$$+g\lambda\sum_{k\in A}\sum_{j=1}^{S}\mathrm{y(j,N,k)} + g\ \lambda\ y(0,\mathrm{N},0) + \sum_{k\in A}\sum_{j=1}^{S}\sum_{r=1}^{N}\ \beta_k y(\mathrm{j,r,k}) \qquad (25)$$

Subject to the constraints,

(1) $y(\mathrm{j,r,k})\geq 0,\ (\mathrm{j,r})\in \mathrm{E},\ \mathrm{k}\in \mathrm{A}_l,\ l=1,2$,

(2) $\sum_{l=1}^{2}\sum_{(\mathrm{j,r})\in \mathrm{E}_l}\sum_{k\in A_l} y(\mathrm{j,r,k}) = 1,$

and the balance equations (3) – (14) obtained by replacing $\Phi^{\pi}(\mathrm{j,r})$ by $\sum_{k\in A} y(\mathrm{j,r,k})$.

4.2 Lemma:

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

Proof:

From the equations

$$y(\mathrm{j,r,k}) = \mathrm{D(j,r,k)}\Phi^{\pi}(\mathrm{j,r}) \qquad (26)$$

and

$$\Phi^{\pi}(\mathrm{j,r}) = \sum_{k\in A} y(\mathrm{j,r,k}),\ (\mathrm{j,r})\in \mathrm{E}. \qquad (27)$$

Since the decision problem is completely ergodic every basic feasible solution to the above linear programming problem has the property that for each $(\mathrm{j,r})\in \mathrm{E},\ y(\mathrm{j,r,k})>0$ for every $\mathrm{k}\in\mathrm{A}$.

V. Numerical Illustration and Discussion

In this system we consider a problem to illustrate the method described in section 4, through numerical examples. We implemented TORA software to solve LPP by simplex algorithm.

We intuitively proposed a conjecture that the reordering rate($p_j\mu$) to be employed depends only on the inventory level.

This conjecture can be proved for zero lead time and reorder is made for fixed items at inventory level s. Sapna, K. P., & Berman, O already proved that the expected cost rate,

$$C(\gamma) = h\left(\frac{S+1}{2}\right) + \sum_{j=0}^{s}\frac{c_1}{s\left[m+\dfrac{\alpha(1)}{\lambda}\right]} + c_2\sum_{k\in\{0,1,2\}}\sum_{j=1}^{s} k\ m\ p(\mathrm{j,k}) + g\sum_{j=0}^{s}\mathrm{p(j,N)} + \beta(\mu)\sum_{j=1}^{s}\sum_{r=1}^{N}\mathrm{p(j,r)}$$

where $m=\dfrac{1}{\gamma},\quad \alpha(1)=\dfrac{1-\dfrac{\lambda}{\gamma}}{1-\left(\dfrac{\lambda}{\gamma}\right)^{N}}.$

For $0\leq \mathrm{j}\leq \mathrm{S},\ 1\leq \mathrm{r}\leq \mathrm{N}$

$$p\,(\mathrm{j},\mathrm{r}) = \left(\frac{\lambda}{\gamma}\right)^{r-1} P(\mathrm{j},0).$$

For $0 \le j \le S$, $$\mathrm{p}(\mathrm{j},0) = \left(\frac{1}{S}\right)\left(\frac{1-\frac{\lambda}{\gamma}}{1-\left(\frac{\lambda}{\gamma}\right)^{N}}\right).$$

Consider the MDP problem with the following parameters:

S = 3, s = 2, N = 4, λ = 2, μ = 3, γ = 4, h = 0.1, c_j = 3j; j = 0, 1, 2, g = 5, β(μ) = 2μ

Action(a)\prob. value	p_2	p_1	p_0
0	**0.5**	**0.2**	**0**
1	**0.5**	**0.8**	**0**
2	**0.0**	**0.0**	**1**

The optimum cost for the system is C = 13.96 and Optimal policy is R*(0, 1, 2, 3) is (2, 0, 1, 0).

VI. Conclusion and future research

In our problem, we use an (s, S) ordering policy at the vendor node. Policies such as one-to-one ordering, or (r,Q) systems or any other fixed ordering policy can be analyzed with same methodology. We used the tools of Semi-Markov decision processes to analyze the problem and linear programming technique is used to determine the optimal reorder level.

The main contribution of the chapter is the determination of the inventory control policy that smoothen the implement supply chain. The MDP considered in this model uses randomized Markov policy, which is first time introduced for MDP application in inventory control systems.

References:

[1]. S. Axsater, Simple solution for a class of two-echelon inventory problems, operations research, 38(1)(1990), 64-69.

[2]. Benita and M. Beamon, Supply Chain design and Analysis: Models and methods, International Journal of Production Economics. Vol. 55, No. 3(1998), 281 294

[3]. O. Berman, K. P Sapna , Optimal Control of service for facilities holding inventory , Computer & Operations Research 28(2001),429-441

[4]. E. Cinlar, Introduction to Stochastic Processes, Prentice Hall, Englewood Cliffs, NJ, 1975.

[5]. A. J. Clark and H. Scarf, Optimal policies for a multi-Echelon Inventory problem, Management science, 6(4)(1960),475-490.

[6]. C. Elango, A continuous review perishable inventory system at service facilities, Ph.D., Thesis, Madurai Kamaraj University, 2001.

[7]. S. C. Graves, The Applications of queueing theory to continuous perishable inventory systems, Management science, 28(1982), 400-406.

[8]. S.C. Graves, The multi-Echelon inventory model for repairable item with one-for-one replenishment, Management science, 31(10)(1985), 1247-1256.

[9]. K. Krishnan and C. Elango, A continuous review (s, S) policy with positive lead times in Supply Chain, Proceedings of National Conference on Intelligent Optimization Modelling (2005), 245-253.

[10]. K. Moinzadeh and H.L. Lee, Batch size and stocking levels in multi-echelon repairable systems, Management Science, 32(12)(1986), 1567-1581 .

[11]. Mine, H. and Osaki S., Markov Decision Processes, American Elsevier Publishing Company Inc, New York (1970).

[12]. Puterman, M.L., Markov Decision Processes: Discrete Stochastic Dynamic Programming, Wiley Interscience Publications Inc., (2005).

[13]. Tijms, H. C. A First Course in Stochastic Models, John Wiley & Sons Inc., (2003).

P. K Santhi, "MDP in Supply Chain: Optimal Inventory Control System" International Journal of Engineering Science Invention (IJESI), vol. 07, no. 8, 2018, pp 55-62.