International Journal of Computer Applications in Engineering Sciences
[VOL II, ISSUE I, MARCH 2012] [ISSN: 2231-4946]

MDP Application in Fixed-Quantity Inventory Control of HRMS

Bingjie He[1], Haiyan Zhao[2]

College of Advanced Vocational Technology, Shanghai University of Engineering Science,
Shanghai, 200437, P.R. China

[1]jessica_hbj@hotmail.com
[2]cemizhy@163.com

***Abstract*— We study a single item hybrid production system with remanufacturing and manufacturing (HRMS). It is assumed that remanufacturing is profitable, that on average, there are more demands than returns, and that the returns are repaired by fixed- quantity. The HRMS is modelled by finite-state discrete-time Markov chains. The inventory problem under consideration is formulated by a Markov Decision Problem (MDP) model. Using the dynamic programming algorithm, we get the optimal policy.**

***Keywords*— HRMS, inventory control, fixed-quantity, Markov Decision Process, dynamic programming algorithm**

I. INTRODUCTION

In the recent years, the growth of environmental concerns has given 'reuse' increasing attention. At the same time, researchers have paid attention to product recovery (reverse logistics) day by day. A large number of papers about it have been published (e.g. [1], [2], [3]-[9]).

Inventory control is one of the areas concerned. The return flow of used products needs appropriate integration into material planning. The situation can be characterized as follows (see Fig. 1). The main objective of inventory management is to increase profitability. A frequently used criterion for choosing the optimal policy is to minimize the total costs, which is equivalent to maximizing the net income in many cases. Scientific inventory management requires a sound mathematical model to describe the behavior of the underlying system and, quite often, an optimal policy with respect to the model.

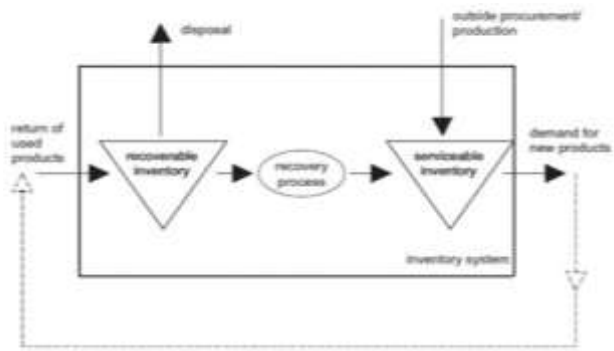

Fig. 1 Recoverable inventory system

A number of authors have studied such hybrid inventory systems. They either seek the structure of the optimal policy, or aim at finding an optimal policy in some prespecified class (e.g. [10], [11], [12]-[25]). The models based on the recoveries were used and would be remanufactured.

Similar to many other dynamic processes in the real world, demand variation encountered by retailers or manufacturers is both random and seasonal in nature. A random/stochastic process may be considered as an ensemble of random variables defined on a common probability space and evolving over time. It is desirable or sometimes necessary to quantify the dynamic relationships among these random events so as to better understand and effectively handle process uncertainties. Considering that the dynamics of such systems are often governed by Markov chains, we resort to Markovian models for solution. Markov chain, a well-known subject introduced by Markov in 1906, has been studied by a host of researchers for many years (e.g. [26],[27],[28],[29]). Markovian formulations (e.g. [30],[31],[32]) are useful in solving a number of real-world problems under uncertainties such as determining the inventory levels for retailers, maintenance scheduling for manufacturers, and scheduling and planning in production management. Markov chain approach has been applied in the design, optimization, and control of manufacturing processes, reliability studies and communication networks, where the underlying system is formulated as stochastic control problem driven by Markovian noise.

In this paper, we consider the inventory of a discrete time hybrid remanufacturing and manufacturing system (HRMS), in which the returns are assumed to be remanufactured by fixed- quantity. We consider two types of uncertainty, demand and recovery, which are assumed to be Poisson distribution. We assume that there are more demands than returns. In this paper, we adopt Markov decision process (MDP) models to solve the problem of inventory planning. The objective is to minimize the long-run average cost per time by considering the cost structure of fixed order costs, holding costs and remanufacturing costs. We provide the dynamic programming equation, and present computational algorithm that leads to an approximation of the optimal policy.

II. The Model

A. *Analysis of the Model*

Following the above motivation, we consider a single-item stochastic inventory model extended with a stochastic inbound item flow. We assume that the order is N time units (N∈N) during the period (see Fig.2). We assume M to be a remanufacture/ manufacture supplier, Mr to be a remanufacture supplier, Mn to be a production supplier. The manager of M checks the stock periodically. At the beginning of ordering period i, returns ξi from Mr will be added to M. At the same time, new products ui will be ordered and replenished from Mn to satisfy the demand Di, i =1, …, N. When the stock decreases to 0, we assume that it is replenished instantly, and that the returned items will be added to the serviceable stock immediately.

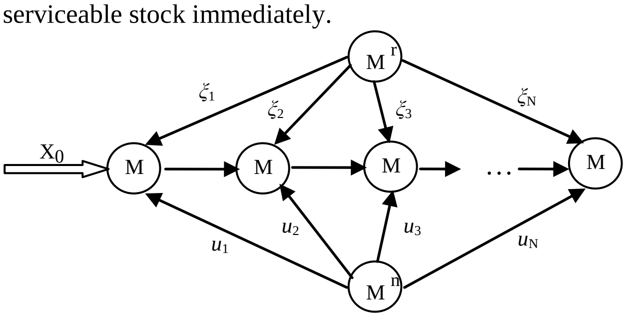

Fig. 2 Inventory control model in HRMS

We assume that the demand and returns are independent Poisson distribution, that there is no leadtimes in both demand and return, and that there is no backorder. Furthermore, we assume that the returns are repaired by fixed quantity, that the remanufacturing quantity is $\xi_{ij} (j = 1, \cdots, n)$, and that the returns are dealt with n times during the ordering period i. It means that if the remanufacturing period is T, the ordering period is nT.

Let

H= the capacity at ordering period i, i =1, …, N, N∈ ℕ;

λ_R= return intensity;

λ_D= demand intensity, $\lambda_R < \lambda_D$;

ξ_i= returns at ordering period i , i =1, …, N, $\xi_i = n\xi_{ij}$, $j = 1, \cdots, n$, N∈ ℕ;

u_i= replenishment at ordering period i , i =1, ……, N, N∈ ℕ, u_0=0;

V_i= net stock before ordering at the beginning of ordering period i , i=1, …, N, N∈ ℕ, $V_i \le H$.

Under these assumptions, the net stock level V_i before ordering forms a discrete Markov chain and the system dynamics are described as

$$V_{i+1} = \min\left[\left(X_i + u_i - \lambda_D \cdot \frac{\xi_i}{\lambda_R}\right)_+ + \xi_{i+1}, H\right]$$

$$= \min\left[\left(V_i + (\frac{1}{n} - 1)\xi_i\right)_+ + \xi_{i+1}, H\right],$$

$$V_1 = \min(V_0 + \xi_1, H),$$

$$(x)_+ = \max\{x, 0\}. \quad (2.1)$$

The states of the stochastic process $\{V_i\}$ take values in $S = \{0, 1, \ldots, H\}$.

We assume ξ_{ij} and ξ to be independently and identically distributed random variables by Poisson process. Let $\xi \sim P(\lambda_R)$,

$$p(\xi = k) = p_k,$$

$$\sum_{k=0}^{\infty} p_k = 1. \quad (2.2)$$

B. *The Long –Run Average Expected Cost*

Furthermore, let

K= fixed cost per replenishment order;

c= ordering cost per item per time;

c_r= remanufacturing cost per item per time;

h= holding cost per item per time;

$V_i(t)$ = inventory position at time t at state V_i;

$\xi_{ij}(T)$ = returns in the j th period T at state $V_i (j = 1, \cdots, n)$.

From above, we get

$$E[\xi_{ij}(T)] = \xi_{ij} \lambda_R T. \quad (2.3)$$

And we get the remanufacturing period $T = \frac{\xi_{ij}}{\lambda_R}$. The ordering period is

$$nT = n \cdot \frac{\xi_{ij}}{\lambda_R} = \frac{\xi_i}{\lambda_R}.$$

The following costs are incurred:

1) Holding cost at state V_i:

$$V_i(t) = u_i - \lambda_D \cdot T_{j-1} + (j-1) \cdot \xi_{ij} - \lambda_D \cdot (t - T_{j-1}), \quad (2.4)$$

$$T_{j-1} \le t \le T_j, \ \ j = 1, \cdots, n, T_0 = 0,$$

$$T = T_j - T_{j-1} = \frac{\xi_{ij}}{\lambda_R},$$

$$T_j = j \cdot \frac{\xi_{ij}}{\lambda_R}, T_n = n \cdot \frac{\xi_{ij}}{\lambda_R},$$

$$V_i(t) = u_i - \lambda_D \cdot (j-1) \cdot \frac{\xi_{ij}}{\lambda_R} + (j-1) \cdot \xi_{ij} - \lambda_D \cdot \left[t(j-1) \frac{\xi_{ij}}{\lambda_R} \right], \quad (2.5)$$

Since there is no backorder, and the stock is replenished instantly when it decreases to 0, letting $V_i(T_n) = 0, j = n, t = T_n,$ we get

$$V_i(T_n) = u_i + (n-1) \cdot \xi_{in} - n \cdot \lambda_D \cdot \frac{\xi_{in}}{\lambda_R} = 0,$$

$$u_i = (n \cdot \frac{\lambda_D}{\lambda_R} - n + 1)\xi_{ij} = \xi_i \cdot \left(\frac{\lambda_D}{\lambda_R} - 1 + \frac{1}{n} \right), \quad (2.6)$$

$$u_N = \xi_i \cdot \frac{\lambda_D}{\lambda_R}, \quad (2.7)$$

$$E[V_i(t)] = E\left[\xi_i \cdot \left(\frac{\lambda_D}{\lambda_R} - 1 + \frac{j}{n} \right) - \lambda_D \cdot t \right] = \frac{\xi_i \cdot \left(\lambda_D - \lambda_R + \frac{\lambda_R}{n} \right)}{2\lambda_R}, \quad (2.8)$$

$$E\,[\text{Holding cost in state } Vi\,] = h \cdot E(V_i) \frac{\xi_i}{\lambda_R}; \quad (2.9)$$

2) Ordering cost at state V_i:

$$c \cdot u_i + K = K + c \cdot \xi_i \cdot \left(\frac{\lambda_D}{\lambda_R} - 1 + \frac{1}{n} \right); \quad (2.10)$$

3) Remanufacturing cost in period T at state V_i:

$c_r \cdot n \cdot \xi_{ij}(T)$,

E (Remanufacturing cost in state V_i) $= c_r \cdot \xi_i$. (2.11)

From above, we conclude that the expected average cost per time in state V_i is

$$E(C_i) = h \cdot E(V_i) + c_r \cdot \lambda_R + K \cdot \frac{\lambda_R}{\xi_i} + c \cdot \left(\lambda_D - \lambda_R + \frac{\lambda_R}{n} \right). \quad (2.12)$$

The long-run average expected inventory cost is

$$C(V_i, \vec{\xi}, n) = \sum_{i=1}^{N} E(C_i), \vec{\xi} = \{\xi_1, \xi_2, \cdots, \xi_N\}. \quad (2.13)$$

The optimal policy is to find $(\vec{\xi}^*, n^*)$ satisfying

$$C(V_i, \vec{\xi}^*, n^*) = \min C(V_i, \vec{\xi}, n). \quad (2.14)$$

III. Analysis of the Cost Function

A. Markov Decision Process

Let V_i be the state of the HRMS mentioned above as in (2.1) at time i, $V_i \in S = \{0,1,\ldots,H\}$, ξ_i be the decision of remanufacturing, $\xi_i \in A(l) = \{0,1,\ldots,\lambda_D \cdot nT\}$. Under the fixed policy $\vec{\xi} = \{\xi_1, \xi_2, \cdots, \xi_N\}$ and the fixed time n, the pair $(V_i, \vec{\xi}, n)$ forms a three-dimensional Markov chain. At the same time, the transition probability from state V_i to V_j satisfies:

$$p_{ll'}(a) = P\{V_{i+1} = l' | V_1 = l_1, V_2 = l_2, \cdots, V_i = l, \xi_1, \xi_2, \cdots, \xi_N\} = P\{V_{i+1} = l' | V_i = l, \xi_i = a\}$$

$$= \begin{cases} p_{l' - \left(l + \frac{a}{n} - a\right)_+}, & l' < H \\ \sum_{k = l' - \left(l + \frac{a}{n} - a\right)_+}^{\infty} p_k, & l' > H \end{cases} \quad (3.1)$$

$$\sum_{l'=0}^{H} p_{ll'}(a) = 1.$$

The sequence of observed states and decisions made is so-called Markov Decision Process (MDP).

B. The Long-Run Expected Cost

From (2.13), we get the long-run expected average cost of the system starting in state V_i and evolving in a period of length N, that is $C(V_i, \vec{\xi}, n)$.

$$C(V_i, \vec{\xi}, n) = \sum_{i=1}^{N} E(C_i) = E\left[\sum_{i=1}^{N} \left(h \cdot V_i + \frac{K \cdot \lambda_R}{\xi_i} \right) \right] + N \cdot c_r \cdot \lambda_R + N \cdot c \cdot \left(\lambda_D - \lambda_R + \frac{\lambda_R}{n} \right) + h \cdot V_N. \quad (3.2)$$

Let

$$F(V_i, \vec{\xi}, n) = h \cdot V_i + \frac{K \cdot \lambda_R}{\xi_i}. \quad (3.3)$$

Then we get the new objective function：

$$C(V_i, \vec{\xi}, n) = E\left[\sum_{i=1}^{N} F(V_i, \vec{\xi}, n) | V_0 = l \right] + N \cdot c_r \cdot \lambda_R + N \cdot c \cdot \left(\lambda_D - \lambda_R + \frac{\lambda_R}{n} \right). \quad (3.4)$$

Let

$$G(l, \vec{\xi}, n) = E\left[\sum_{i=1}^{N} F(V_i, \vec{\xi}, n) | V_0 = l \right], \quad (3.5)$$

$$G_m(l, \vec{\xi}_m, n) = E\left[\sum_{i=m}^{N} F(V_i, \vec{\xi}_m, n) | V_{m-1} = l \right]. \quad (3.6)$$

The optimal policy is to find $\left(\vec{\xi}_m^{\ *}, n^*\right)$ satisfying

$$G_m\left(l, \vec{\xi}_m^{\ *}, n^*\right) = \min\left\{ E\left[\sum_{i=m}^{N} F(V_i, \vec{\xi}_m, n) | V_{m-1} = l \right] \right\}, \vec{\xi}_m^{\ *} = (\xi_m^{\ *}, \cdots, \xi_N^{\ *}). \quad (3.7)$$

At the same time, let

$$G_m(l) = G_m\left(l, \vec{\xi}_m^{\ *}, n^*\right).$$

C. The Optimal Policy Algorithm

Let $V_m = l, \xi_m^{\ *} = a^*, \xi_m = a, V_{m+1} = l'$.It can be shown that (e.g. [33])

$$\mathrm{F}(l,a^*,n^*)+\sum_{l'=0}^{\mathrm{H}}\mathrm{G}_{m+1}(l')\cdot\mathrm{P}_{ll'}(a^*)$$
$$=\min_{a\in\mathrm{A}(l)}\left\{\mathrm{F}(l,a,n)+\sum_{l'=0}^{\mathrm{H}}\mathrm{G}_{m+1}(l')\cdot\mathrm{P}_{ll'}(a)\right\},$$
$$m=1,\cdots,\mathrm{N}-1. \qquad (3.8)$$

To obtain the optimal policy, we use the following computational algorithm.

Step1. From *m*=N, *m*=*m*-1, until *m*=0, letting

$$n^*=\frac{\mathrm{c}\cdot\lambda_{\mathrm{R}}\cdot(1-\mathrm{K})}{\mathrm{c}\cdot\lambda_{\mathrm{D}}-\lambda_{\mathrm{R}}\cdot(\mathrm{c}+\mathrm{K})}, \qquad (3.9)$$

$$\xi_m{}^*(l)=\frac{2\cdot l\cdot\lambda_{\mathrm{R}}}{\lambda_{\mathrm{D}}-\lambda_{\mathrm{R}}+\frac{\lambda_{\mathrm{R}}}{n^*}}, \qquad (3.10)$$

we calculate

$$\mathrm{G}_m(l,\xi_m,n^*)=\left[\sum_{i=m}^{\mathrm{N}}\frac{\mathrm{h}\cdot\xi_i\cdot\left(\lambda_{\mathrm{D}}-\lambda_{\mathrm{R}}+\frac{\lambda_{\mathrm{R}}}{n^*}\right)}{2\lambda_{\mathrm{R}}}+\frac{\mathrm{K}\cdot\lambda_{\mathrm{R}}}{\xi_i{}^*}\right]$$
$$+\sum_{l'=0}^{\mathrm{H}}\mathrm{G}_{m+1}(l')\cdot\mathrm{P}_{ll'}. \qquad (3.11)$$

Step2. We calculate $\mathrm{C}(\mathrm{V}_0,\vec{\xi}^*,n^*)$ from (3.4).

TABLE I: OPTIMAL POLICY

	m=N=3			m=N-1=2			m=N-2=1			$\mathrm{C}(\mathrm{V}_0,\vec{\xi}^*,n^*)$
	$\xi_3(l)$	$\mathrm{G}_3(l)$	n^*	$\xi_2{}^*(l)$	$\mathrm{G}_2(l)$	n^*	$\xi_1{}^*(l)$	$\mathrm{G}_1(l)$	n^*	
l=1	1.63	10.58	4.5	1.63	21.16	4.5	1.63	31.74	4.5	91.48
l=2	3.28	8.28	4.5	3.28	20.7	4.5	3.28	35.19	4.5	92.17
l=3	4.9	8.85	4.5	4.9	17.7	4.5	4.9	26.55	4.5	81.1
l=4	6.55	10.14	4.5	6.55	25.35	4.5	6.55	43.10	4.5	106.59
l=5	8.18	11.7	4.5	8.18	23.4	4.5	8.18	35.1	4.5	98.2
l=6	9.81	13.4	4.5	9.81	26.8	4.5	9.81	40.2	4.5	108.4

IV. NUMERICAL EXAMPLE

We take $\lambda_R=2,\lambda_D=4$,K=7,c=3, $\mathrm{c_r}=1$,h=2, *p*(*ξ*=2)= *p*(*ξ*=4)=0.5, N=3,H=6. From the model (3.4) and the above computational algorithm, we calculate theoptimal policy (see Table Ⅰ).

From the table 1, we can find that the long-run average expected inventory cost is too more when *l* =4 and *l* =6, so we prefer to *l* =1,2,3,5...

V. CONCLUSION

In this paper, we discuss the inventory planning of a discrete time hybrid remanufacturing and manufacturing system (HRMS), in which the returns are assumed to be dealt with by fixed- quantity. The HRMS is modeled by finite-state discrete-time Markov chains. At the same time, we apply the Markov Decision Problem (MDP) model to the inventory problem under consideration. We get the optimal strategy by using the dynamic programming algorithm. We illustrate the conclusion in a numerical example.

REFERENCES

[1] Stock J R, Reverse Logistics, Oak Brook Illinois: Council of Logistics Management, 1992. 1-10

[2] Kopicky R J, Berg M J, Legg L, et al, Reuse and recycling: reverse logistics opportunities, Oak Brook Illinois: Council of Logistics Management, 1993: 1-8.

[3] M. Fleischmann, J.M. Bloemhof-Ruwaard, R. Dekker, E.Van derLaan, J.A.E.E. van Nunen, L.N. Van Wassenhove, Quantitative models for reverse logistics: A review, European Journal of Operational Research, 1997(103):1-17.

[4] M. Fleischmann, Roelof Kuik, On optimal inventory control with independent stochastic item returns, European Journal of Operational Research, 2003 (151):25-37.

[5] D. Rogers, R.S. Tibben- Lembke, Going Backwards: Reverse Logistics Trends and Practices, RLEC Press, Pittsburgh, 1999.

[6] V.D.R. Guide, V. Jayaraman, R. Srivastava, W.C. Benton, Supply-chain management for recoverable manufacturing systems,Interfaces, 30 (3) ,2000, 125-142.

[7] V.D.R. Guide Jr, Production planning and control for remanufacturing: industry practice and research needs, Journal of Operations Management , 18 (4) ,2000, 467-483.

[8] E. van der Laan, The e!ects of remanuf acturing on inventory control, Ph.D. Thesis, Erasmus University Rotterdam, the Netherlands, 1997.

[9] R. Lund, Remanufacturing, Technology Review, 87 (2), 1984, 18-23.

[10] Toktay L B, Wein M, Zenios S A, Inventory management of remanufacturable products, Managements Science, 2000, 46(11): 1412-1426.

[11] Minner S, Strategic safety stocks in reverse logistics supply chains, International Journal of Production Economics, 2001, 71(3): 417-428.
[12] Minner S. Kleber R, Optimal control of Production and remanufacturing in a simple recovery model with linear cost functions,O R Spekrtum, 2001, 23(1): 3-24.
[13] Kleber R, Minner S, Kiesmuller G, A continuous time inventory model for a product recovery system with multiple options, International Journal of Production Economics, 2002, 79: 121-141.
[14] Kiesmuller G P, Van der Lann E A, An inventory model with dependent product demands and returns, International Journal of Production Economics, 2001, 72(1): 73-87.
[15] D.P. Heyman, Optimal disposal policies for a single-item inventory system with returns, Naval Research Logistics Quarterly 24, 1977, 385-405.
[16] K. Inderfurth, Simple optimal replenishment and disposal policies for a product recovery system with leadtimes, OR Spektrum 19, 1997, 111-122.
[17] K. Inderfurth, T. Jensen, Analysis of MRP policies with recovery option, Preprint No. 1, Faculty of Economics and Management, Otto-von-Guericke University of Magdeburg, Magdeburg, Germany, 1998.
[18] K. Inderfurth, E. van der Laan, Leadtimes e!ects and policy improvement for stochastic inventory control with remanufacturing, Preprint No. 22, Faculty of Economics and Management, Otto-von-Guericke University of Magdeburg, Magdeburg, Germany, 1998.
[19] E. van der Laan, M. Salomon, Production planning and inventory control with remanufacturing and disposal, European Journal of Operational Research 102 (2) , 1997, 264-278.
[20] K. Richter, The EOQ and waste disposal model with variable setup numbers, European Journal of Operational Research 95 , 1996,313-324.
[21] V.P. Simpson, Optimum solution structure for a repairable inventory problem, Operations research 26 (2),1978, 270-281.
[22] R.H. Teunter, E. van der Laan, K. Inderfurth, How to set the holding cost rates in average cost inventory models with reverse logistics? OMEGA The International Journal of Management Science 28 , 2000, 409-415.
[23] P. Kelle, E.A. Silver, Purchasing policy of new containers considering the random returns ofpreviously issued containers, IIE Transactions 21 (4) ,1989,349-354.
[24] K. Richter, The extended EOQ repair and waste disposal model, International, Journal of Production Economics 45 (1-3) , 1996, 443-448.
[25] D.A. Schrady, A deterministic inventory model for repairable items, Naval Research Logistics Quarterly 14 , 1967, 391-398.
[26] Chung, K. L, Markov chain with stationary transition probabilities, Berlin: Springer, 1960
[27] Doob, J. L, Stochastic processes, New York: Wiley,1953.
[28] Feller, W, An introduction to probability theory and its applications, New York: Wiley, 1971.
[29] Kushner, H. J., & Yin, G. G, Stochastic approximation algorithms and applications, New York: Springer, 1997.
[30] Chiang, C. L., An introduction to stochastic processes and their applications, New York: Robert E. Krieger Publishing Co, 1980.
[31] Taylor, H. M., & Karlin, S, An introduction to stochastic modeling , Boston: Academic Press ,1998.
[32] Yang, H., Yin, G., Yin, K., & Zhang, Q, Control of singularly perturbed Markov chains: a numerical study, The ANZIAM Journal (formerly known as Journal of the Australian
[33] Mathematical Society, Series B: Applied Mathematics), 2002.
[34] Hillier, F.S. and Lieberman, G.J, Introductions to Operations Research, 6ed, Holden-Day, Inc. Oakland, C A, 1999.