Management Research and Practice
Volume 2, Issue 4 / December 2010

SUBSTANTIATING THE OPTIMAL DISTRIBUTION POLICY USING MARKOV DECISION PROCESSES

Gheorghe BĂŞANU[1], Victor TELEAŞĂ[2]
[1]*Academy of Economic Studies, Piata Romana 6, Bucharest, Romania, gheorghe.basanu@man.ase.ro*
[2]*Joint Logistic Command, Drumul Taberei, 7-9, Bucharest, Romania, victorteleasa@yahoo.com*

Abstract
The paper represents a means of substantiating the optimal politicy of transport and distribution of goods groupage using Markov's decisional processes, respectively through R. Howard's method of strategies' space. This kind of method is based on an iterative optimization algorithm, whose structure permits that, with the successive covering of each method, to limit the number of ulterior repetitions. In this way, the optimal politicy that result, constituted from a number of decisions, finite or not, aims at optimizing the following decisions in a close connection with the first decision's consequences, whatever that would be.

Concretely, a logistic centre's managers have the responsibility of identifying the optimal transport policy of the unit loads prepared for being shipped in three ways: palletized, containerized and packaged, according to the ways of transport which will be used: causeway, railway and airway, and the costs associated to each shipping means.

Keywords: distribution centre, optimal policy, decision, average cost, unit loads.

1. INTRODUCTION

The continuous increase of the amount of goods which circulate on the distribution chains between the producers, the wholesalers, the retailers and the final clients created the premises of identification and developing of a modern method of grouping them and of forming the unit loads, with the purpose of accelerating the flows of materials and products. These kinds of methods, continuously perfected in time, determined an increase in efficiency in handling the goods and in processing the orders, decrease of costs associated to these logistic operations, serving promptly the clients, and diminishing the material losses or their qualitative depreciation.

Utilizing on the large scale of the unit loads leads to:

1. The increasing of the materials' circulation speed, along their entire moving flow, irrespective of their origin and destination, through the mechanising and automation of the loading-unloading operations;
2. Warranting the quantitative and qualitative integrity of the goods and an increased degree of controlling them during the transport;

mrp.ase.ro
Management Research and Practice
Volume 2, Issue 4 / December 2010

3. Reducing the number of packages, an action that generates favourable consequences in the transport's efficient utilisation;
4. Increasing the work productivity inside the logistic centres, through diminishing the number of handling operations;
5. Reducing the expenses with moving the goods, creating several handling and loading operations in correlation with the beneficiary's wishes, having in mind the compatibility between the goods.
6. Simplifying the informational flow inside the logistic system, by placing on the unit loads, the radio frequency identification labels or the bar codes, with interesting information for the specialists in logistic, about the structure, contains and their route.

The dynamic programming is a mathematic technique of describing and solving of a great number of optimization issues. Among these, the problems of sequential decisions form an important category (Mihoc et al., 1978). With this purpose in mind, in order to identify the specific of the respective logistical problem, we'll introduce several main concepts.

In each moment in time, a system can be just in one of the *N* states, where *N* is finite. To each *i* state several *k* alternatives are associated *(k=1,2,…,K)*, through which the system transitions from the state *i* in the state *j*, in an aleatory way, with the known probabilities (Norris, 1997). The number of alternatives for *i* state can be different from the ones of the *j* state, and to each transition from a state to another it's associated an income or cost. The alternative established at a certain *n* moment, represents the decision d_i associated to the system's state *i*. A set of decisions $d_1, d_2, \ldots, d_N$, , which form a column vector, represent the *P* policy (Malița and Zidăroiu, 1971).

We consider a Markov decision process, which is developed in *N* stages, in which the decisions are taken in each stage and in which $p_{ij}(k)$ represents the transition matrix that corresponds to the decisions $d_i(P) = k$ inside the *P* politics. We suppose that in each stage a change of state and cost takes place, a change which is depend on the initial state, the final state and the decision taken, whose transition matrix is $Q_{ij}(k)$ (Bellman, 1957). Supposing that the system's initial state is known, the issue is choosing an optimal stationery politics, *P**, which, in our case, should ensure the average cost's minimization.

In the first stage, the *absolute and relative values g(P), respectively* $v_i(P)$*, will be established* (Howard, 1960).

As a result of any decision taken inside a *P* policy, it will result an immediate cost $C_i(k)$ which results from the formula:

$$C_i(k) = \sum_{j=1}^{N} p_{ij}(k) \cdot Q_{ij}(k) \qquad (1)$$

If we represent with $v_i^n(P)$ the total expected cost of a process in N stages, beginning with the *i* state and applying a given policy *P*, the following recurrence equation results:

$$v_i^n(P) = \sum_{j=1}^{N} p_{ij}(k) \cdot \left[Q_{ij} + v_j^{n-1}(P)\right],\ i = 1,2,\ldots,N \qquad (2)$$

Replacing in a proper way the equation (1) in the equation (2), it results that:

$$v_i^n(P) = C_i(k) + \sum_{j=1}^{N} p_{ij}(k) \cdot v_j^{n-1}(P),\ i = 1,2,\ldots,N \qquad (3)$$

For *n* grows large, the $v_i^n(P)$ behaves asymptomatically and can be expressed as the sum of two components (Mihoc et al., 1978):

$$v_i^n(P) \approx n \cdot g(P) + v_i(P),\ i = 1,2,\ldots,N \qquad (4)$$

In this expression it's noticed the fact that just the second term depends of the initial *i* state of the system.

Similarly, for $v_j^{n-1}(P)$ we deduct the expression:

$$v_j^{n-1}(P) \approx (n-1) \cdot g(P) + v_j(P),\ i = 1,2,\ldots,N \qquad (5)$$

Introducing the relations (4) and (5) in the recurrence equation (2) it's obtained:

$$g(P) + v_i(P) = C_i(k) + \sum_{j=1}^{N} p_{ij}(k) \cdot v_j(P),\ i = 1,2,\ldots,N \qquad (6)$$

The equation (6) represents a system of *N* linear equations with *N+1* unknowns. In order to be able to solve the system, it's chosen by convention $v_N(P) = 0$, and the solution of the system's equations accentuates the average cost on a stage, *g(P),* obtained through implementing the policy *P*. We obtain also the relative values of the initial states $v_i(P)$, for the policy *P*.

In order to obtain a cost's decrease, in the second stage there is raised the issue of *improving the policy applied* (Howard, 1960).

Thinking that, initially, a P_n policy was applied and there were obtained the relative values $v_i(P_n)$, an alternative policy P_{n+1} must be identified in the next step, so that for each state *i* of the system, $d_i(P_{n+1}) = k$, must represent a decision, through which to obtain:

$$\min_{k=1,2,\ldots,K}\left[C_i(k) + \sum_{j=1}^{N} p_{ij}(k) \cdot v_j(P_n)\right] \qquad (7)$$

The policy applied, $P_{n+1,}$ is an optimal one, if this is identical with the previous one, so it must result:

$$P_{n+1} \equiv P_n \qquad (8)$$

$$g(P_{n+1}) \leq g(P_n) \qquad (9)$$

If the result doesn't observe the relationships (8) and (9), then another policy must be searched.

In order to perceive the practical abilities of applying the model described above, we present a concrete example of distribution centers of the goods from the urbane environment. At the logistic center's level, the preparation of the delivering orders involves the developing of a complex of operations, like the ones related to the selection of the goods, the packaging, the wrapping, the matching, the forming of the unit loads and their shipping under a palletized, containerized and packaged form, toward the unloading points, which the logistic center serve. The respective center has a direct access to the causeways and the railway, being very close to an airport endowed with a cargo terminal.

In developing this example, the following elements were considered:

- the distribution centers, where the unit loads are assigned to arrive, are located at relatively far distances and equal towards the logistic center;
- a shipping (a means of transport) can group variously type of unit loads;
- the distribution centers can be supply all, through the same transport means (according to the established politics);
- due to the physic-chemical characteristics, the materials and the packaged products aren't accepted in the air transport;
- as a structure, the unit loads differ from one another, according to the materials' nature and the nature of the products from which they are constituted;
- the costs associate to the shipping also include the expenses with the handling operations executed in the logistic center's platform, respectively between the transportation means, where it's necessary;
- containerizing the materials and the products is realized in box pallets, mini containers or in TEU 20' containers, according to the transportation means chosen.

According to these elements, the logistic managers must establish the optimal policy of transport of the unit loads, having in mind their way of constituting and the costs associated to them.

eISSN
2067- 2462

Analyzing the elements from above, through which the solving problem is defined, it results that the materials and the products that make the object of the distribution can be found in three states *(N=3)*: *palletized, containerized and packaged.* Furthermore, there are *three means of transport* of these unit loads, on which the logistic managers depend in their decisions – causeway, railway and airway.

-for states 1 and 2 there are 3 means of transport of these unit loads, ($n_1 = 3$; $n_2 = 3$);

-for state 3 there exist 2 transport possibilities: ($n_3 = 2$).

As a result, we have available 18 stationary policies.

In the table no. 1 there are presented the main elements of calculus, which are the ground of determining the optimal policy (the costs are expressed in thousands of monetary units).

As an example, from the data written in table no. 1, as a result of the analyses made by the specialists in logistics, about the availability of some transport capacities at the level of the carriers market, it results that, in state 1, decision 1, for the palletized materials and products there exists the probability that 70% of them to be shipped by causeway, with a palletized unit loads groupage, at a cost of 30.000 m.u. (monetary units), 20% of them to be shipped by causeway, with a containerized unit loads groupage, at a cost of 20.000 m.u. and 10% of them to be shipped by causeway, with a packaged unit loads groupage, at a cost of 40.000 u.m.

In a similar way, in the case of the containerized materials and products (state 2, decision 2), there exists the probability that 5% of them to be shipped by railway, with a palletized unit loads groupage, at a cost of 20.000 m.u., 90% of them to be shipped by railway, with a containerized unit loads groupage, at a cost of 15.000 m.u. and 5% of them to be shipped by railway, with a packaged unit loads groupage, at a cost of 20.000 m.u.

TABLE 1 - THE ELEMENTS THROUGH WHICH THE OPTIMAL POLICY IS CALCULATED

State	Decision	Transition probabilities from the state *i* in the state *j*			The associated partial cost			The immediate cost
i	*k*	$p_{ij}(k)$			$Q_{ij}(k)$			$C_i(k)$
1	1	0,700	0,200	0,100	30	20	40	29
	2	0,600	0,300	0,100	20	10	30	18
	3	0,850	0,100	0,050	60	80	90	63,500
2	1	0,100	0,800	0,100	25	20	30	21,500
	2	0,050	0,900	0,050	20	15	20	15,500
	3	0,100	0,850	0,050	70	85	95	84
3	1	0,095	0,005	0,900	60	70	75	73,550
	2	0,195	0,005	0,800	75	80	90	87,025

We initially apply a policy P_1, which regards the transport of the materials and products by causeway, irrespective of the way these products are grouped. This policy's vector will be:

$$P_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

We associate to the vector P_1, the transition probabilities matrix $p_{ij}(1)$ and the costs matrix $Q_{ij}(1)$:

$$pij(1) = \begin{pmatrix} 0{,}700 & 0{,}200 & 0{,}100 \\ 0{,}100 & 0{,}800 & 0{,}100 \\ 0{,}095 & 0{,}005 & 0{,}900 \end{pmatrix}$$

$$Qij(1) = \begin{pmatrix} 30 & 20 & 40 \\ 25 & 20 & 30 \\ 60 & 70 & 75 \end{pmatrix}$$

Applying the equation (1) it results the vector that corresponds to the immediate cost:

$$C_i(1) = \begin{pmatrix} 29 \\ 21{,}50 \\ 73{,}55 \end{pmatrix}$$

In a similar way, it's deduced the transition probabilities matrix $p_{ij}(2)$, the costs matrix $Q_{ij}(2)$ and the vector that corresponds to the immediate cost $C_i(2)$:

$$pij(2) = \begin{pmatrix} 0{,}600 & 0{,}300 & 0{,}100 \\ 0{,}050 & 0{,}900 & 0{,}050 \\ 0{,}195 & 0{,}005 & 0{,}800 \end{pmatrix}$$

$$Qij(2) = \begin{pmatrix} 20 & 10 & 30 \\ 20 & 15 & 20 \\ 75 & 80 & 90 \end{pmatrix}$$

$$C_i(2) = \begin{pmatrix} 18 \\ 15{,}500 \\ 87{,}025 \end{pmatrix}$$

respectively the transition probabilities matrix $p_{ij}(3)$, the costs matrix $Q_{ij}(3)$ and the vector which corresponds to the immediate cost $C_i(3)$ are deduced:

$$pij(3) = \begin{pmatrix} 0{,}850 & 0{,}100 & 0{,}050 \\ 0{,}100 & 0{,}850 & 0{,}050 \end{pmatrix}$$

$$Qij(3) = \begin{pmatrix} 60 & 80 & 90 \\ 70 & 85 & 95 \end{pmatrix}$$

$$C_i(3) = \begin{pmatrix} 63{,}50 \\ 84 \end{pmatrix}$$

Using the equation (6) it's determined the level of the necessary costs for implementing the P_1 policy. For simplifying the writing, in the equation system is used the simplified denotation $v_i(P_1) = v_i$. This way, it results the following equation system:

$$\begin{cases} g + v_1 = 29{,}00 + 0{,}700 \cdot v_1 + 0{,}200 \cdot v_2 + 0{,}100 \cdot v_3 \\ g + v_2 = 21{,}50 + 0{,}100 \cdot v_1 + 0{,}800 \cdot v_2 + 0{,}100 \cdot v_3 \\ g + v_3 = 73{,}55 + 0{,}095 \cdot v_1 + 0{,}005 \cdot v_2 + 0{,}900 \cdot v_3 \end{cases}$$

In order to solve this equation system, it's chosen the variant $v_3 = 0$. After making the necessary calculations, the following solutions are obtained:

$$v_1(P_1) = -241{,}031; \quad v_2(P_1) = -259{,}781; \quad v_3(P_1) = 0; \quad g(P_1) = 49{,}353$$

In conclusion, if materials and products are shipping with road carriers, irrespective of the way of grouping them, the average cost will be of 49.353 m.u. (monetary units).

In order to be obtained a cost's potential diminishing, there is appealed to identifying a new logistic policy P_2, improving the policy P_1.

Applying the equation (7) and considering that $v_3(P_1) = 0$, it's obtained:

-for state 1:

$$\min_{k=1,2,3}\left[C_1(k) + \sum_{j=1}^{3} p_{1j}(k) \cdot v_j(P_1)\right] = \min\{29 + 0{,}7 \cdot v_1 + 0{,}2 \cdot v_2; \;\; 18 + 0{,}6 \cdot v_1 + 0{,}3 \cdot v_2; \;\; 63{,}5 + $$
$$+ 0{,}85 \cdot v_1 + 0{,}1 \cdot v_2\} = \min\{-191{,}68; \;\; -204{,}55; \;\; -167{,}35\} = -204{,}55$$

It results the decision $d_1(P_2) = 2$.

- for state 2:

$$\min_{k=1,2,3}\left[C_2(k) + \sum_{j=1}^{3} p_{2j}(k) \cdot v_j(P_1)\right] = \min\{21{,}5 + 0{,}1 \cdot v_1 + 0{,}8 \cdot v_2; \;\; 15{,}5 + 0{,}05 \cdot v_1 + 0{,}9 \cdot v_2;$$
$$84 + 0{,}1 \cdot v_1 + 0{,}85 \cdot v_2\} = \min\{-210{,}43; \;\; -221{,}35; \;\; -118{,}17\} = -221{,}35$$

It results the decision $d_2(P_2) = 2$;

-for state 3:

$$\min_{k=1,2,3}\left[C_3(k) + \sum_{j=1}^{3} p_{3j}(k) \cdot v_j(P_1)\right] = \min\{73{,}55 + 0{,}095 \cdot v_1 + 0{,}05 \cdot v_2; \;\; 87{,}025 + 0{,}195 \cdot v_1 + $$
$$+ 0{,}05 \cdot v_2\} = \min\{49{,}35; \;\; 38{,}73\} = 38{,}73$$

From which it results the decision $d_3(P_2) = 2$.

Having in mind the three decisions, it results that the vector of policy P_2 will look like this:

$$P_2 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$

This new policy is applied to the materials and products that are shipped by railway, irrespective of the way they are grouped.

The transition probabilities matrix $p_{ij}(2)$, the costs matrix $Q_{ij}(2)$ and the vector that corresponds to the immediate cost $C_i(2)$, associated to this policy are the ones mentioned above.

Using the equation (6) the values of the costs that are necessary for the implementation of the P_2 policy are determined. For simplifying things, in the equation system it's used one more time the notation $v_i(P_2) = v_i$. This way, the following equation system results:

$$\begin{cases} g + v_1 = 18{,}000 + 0{,}600 \cdot v_1 + 0{,}300 \cdot v_2 + 0{,}100 \cdot v_3 \\ g + v_2 = 15{,}500 + 0{,}050 \cdot v_1 + 0{,}900 \cdot v_2 + 0{,}050 \cdot v_3 \\ g + v_3 = 87{,}025 + 0{,}195 \cdot v_1 + 0{,}005 \cdot v_2 + 0{,}800 \cdot v_3 \end{cases}$$

In order to solve it, it's chosen $v_3 = 0$. After making the calculations, the following solutions are obtained:

$$v_1(P_2) = \text{-}269{,}335;\ \ v_2(P_2) = \text{-}309{,}252;\ \ v_3(P_2) = 0;\ \ g(P_2) = 32{,}958$$

It can be noticed that, implementing that logistic policy *(P_2)*, it's obtained an average cost of 32.958 monetary units, a cost which is with 16.395 monetary units lower that in the case of the P_1 policy.

The process of optimization is continued in this sense, trying to identify a new logistic policy P_3, improving P_2 policy. The equation (7) is applied, considering that $v_3(P_2) = 0$

This way, it's obtained:

-for state 1:

$$\min_{k=1,2,3}\left[C_1(k) + \sum_{j=1}^{3} p_{1j}(k) \cdot v_j(P_2)\right] = \min\{-221{,}38; -236{,}38; -196{,}36\} = -236{,}38$$

from which it results the decision $d_1(P_3) = 2$;

-for state 2:

$$\min_{k=1,2,3}\left[C_2(k) + \sum_{j=1}^{3} p_{2j}(k) \cdot v_j(P_2)\right] = \min\{-252{,}84; -276{,}29; -205{,}80\} = -276{,}29$$

mrp.ase.ro
Management Research and Practice
Volume 2, Issue 4 / December 2010

from which it results the decision $d_2(P_3) = 2$;

-for state 3:

$$\min_{k=1,2}\left[C_3(k) + \sum_{j=1}^{3} p_{3j}(k)\cdot v_j(P_2)\right] = \min\{46,42;32,96\} = 32,96$$

from which it results the decision $d_3(P_3) = 2$

Based on these decisions, a new policy is defined:

$$P_3 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$

It can be noticed that this new policy is identical to the previous one. In the same time, we notice that the conditions specified in the relations (8) and (9) are fulfilled and, as a result, the optimal policy is $P^* = P_2$, respectively the transport by railway of the materials and products, irrespective of their way of grouping at an average cost per stage of 32.958 monetary units for each state.

In essence, the way of solving the example introduced hereby involved the decomposing of the logistic process in stages and the iterative optimization at the level of each stage, based on some recurrence situations. The optimal criterion is the average cost for each stage, considering that the planning horizon is infinite and the total cost is limitless (Puterman, 2005). Therefore, it results that the optimal policy identified is stationary and nonrandomized.

REFERENCES

Bellman, R. E. (1957). *Dynamic Programming*. Princeton: Princeton University Press;

Howard, R. D. (1960). *Dynamic Programming and Markov Processes*. New York: John Wiley & Sons;

Malița, M. and Zidăroiu, C. (1971). *The organizational mathematics.* Bucharest: Technical Press;

Mihoc, Gh., Bergthaller, C. and Urseanu, V. (1978). *Stochastic processes. Elements of theory and applications*. Bucharest: Scientific and Encyclopedic Publishing House;

Norris, J. R. (1997). *Markov Chains.* Cambridge: Cambridge University Press;

Puterman, M. L. (2005). *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* New York: John Wiley & Sons.

eISSN
2067- 2462