”Vasile Alecsandri” University of Bacău
Faculty of Sciences
Scientific Studies and Research
Series Mathematics and Informatics
Vol. 25(2015), No. 1, 11-22

ALGORITHMS FOR SOLVING STOCHASTIC DISCRETE OPTIMAL CONTROL PROBLEMS ON NETWORKS

MARIA CAPCELEA, TITU CAPCELEA

Abstract. In this paper we consider the stationary stochastic discrete optimal control problem with average cost criterion. We formulate this problem on networks and propose polynomial time algorithms for determining the optimal control by using a linear programming approach.

Work presented at CAIM 2014, September 19-22, 2014, ”Vasile Alecsandri” University of Bacău

1. Introduction

The aim of this paper is to develop algorithms for determining the optimal solutions of a new class of a discrete control problems with an infinite time horizon. The set of states of the system in the considered problems is finite and the starting state is fixed. We study a class of stochastic discrete control problems that emphasis Markov decision problems and deterministic optimal control problems with an infinite time horizon.

Keywords and phrases: time-discrete system, optimal control, stochastic network, Markov process, optimal strategy, linear programming.
(2010)Mathematics Subject Classification: 93E20.

We obtain the stochastic versions of classical discrete control problems assuming that the dynamical system in the control process may admit dynamical states in which the vector of control parameters is changing in a random way according to given distribution functions of the probabilities on given feasible sets.

So, in the considered control problems we assume that the dynamics of the system may contain controllable states as well as uncontrollable states. These problems are formulated on networks and polynomial time algorithms for determining their optimal solutions are proposed.

2. Problem Formulation

Let a discrete dynamical system $\mathbb{L}$ with finite set of states X be given, where $|X| = n$. At every discrete moment of time $t = 0, 1, 2, \ldots$ the state of $\mathbb{L}$ is $x(t) \in X$. The dynamics of the system is described by a directed graph of states' transitions $G = (X, E)$ where the set of vertices X corresponds to the set of states of the dynamical system and an arbitrary directed edge $e = (x, y) \in E$ expresses the possibility of the system $\mathbb{L}$ to pass from the state $x = x(t)$ to the state $y = x(t+1)$ at every discrete moment of time t. So, a directed edge $e = (x, y)$ in G corresponds to a stationary control of the system in the state $x \in X$. We assume that graph G does not contain deadlock vertices, i.e., for each vertex x there exists at least one leaving directed edge $e = (x, y) \in E$. In addition, we assume that to each edge $e = (x, y) \in E$ a quantity $c_e \in \mathbb{R}$ is associated, which expresses the cost of the system $\mathbb{L}$ to pass from the state $x = x(t)$ to the state $y = x(t)$ for every $t = 0, 1, 2, \ldots$.

A sequence of directed edges $E' = \{e_0, e_1, e_2, \ldots, e_t, \ldots\}$, where $e_t = (x(t), x(t+1))$, $t = 0, 1, 2, \ldots$, determines in G a control of the dynamical system with a fixed starting state $x_0 = x(0)$. An arbitrary control in G generates a trajectory $x_0 = x(0), x(1), x(2), \ldots$ for which the average cost per transition can be defined in the following way

$$f(E') = \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} c_{e_\tau}.$$

In [1] it is shown that this value exists and that if G is strongly connected, then for an arbitrary fixed starting state $x_0 = x(0)$ there exists the optimal control $E^* = \{e_0^*, e_1^*, e_2^* \dots\}$ for which

$$f(E^*) = \min_{E'} \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} c_{e_\tau}$$

and this optimal control does not depend either on the starting state or on time. Therefore, the optimal control for this deterministic problem can be found in the set of stationary strategies $\mathbb{S}$.

In this paper we assume that the set of states X of the dynamical system admit states in which the system $\mathbb{L}$ makes transitions to the next state in a random way according to a given distribution function of probabilities on the set of possible transitions from these states [2]. So, the set of states X is divided into two subsets X_C and X_N ($X = X_C \cup X_N$, $X_C \cap X_N = \emptyset$), where X_C represents the set of states $x \in X$ in which the transitions of the system to the next state y can be controlled by the decision maker at every discrete moment of time t and X_N represents the set of states $x \in X$ in which the decision maker is not able to control the transition because the system passes to the next state y randomly. Thus, for each $x \in X_N$ a probability distribution function $p_{x,y}$ on the set of possible transitions (x, y) from x to $y \in X(x)$ is given, i.e.,

$$(1) \qquad \sum_{y\in X(x)} p_{x,y} = 1, \quad \forall x \in X_N; \quad p_{x,y} \geq 0, \quad \forall y \in X(x).$$

Here $p_{x,y}$ expresses the probability of the system's transition from the state x to the state y for every discrete moment of time t.

We call the graph G, with the properties mentioned above, decision network and denote it by (G, X_C, X_N, c, p, x_0). So, this network is determined by the directed graph G with a fixed starting state x_0, the subsets X_C, X_N, the cost function $c : E \to \mathbb{R}$ and the probability function $p : E_N \to [0, 1]$ on the subset of the edges $E_N = \{e = (x, y) \in E \,\big|\, x \in X_N,\ y \in X\}$, where p satisfies the condition (1).

We define a stationary strategy for the control problem on networks as a map:

$$s : x \to y \in X(x) \quad \text{for} \quad x \in X_C,$$

where $X(x) = \{y \in X | e = (x, y) \in E\}$.

Let s be an arbitrary stationary strategy. Then we can determine the graph $G_s = (X,\, E_s \cup E_N)$, where $E_s = \{e = (x, y) \in E \,|\, x \in X_C$, $y = s(x)\}$. This graph corresponds to a Markov process with the probability matrix $P^s = (p^s_{x,y})$, where

$$p^s_{x,y} = \begin{cases} p_{x,y}, & \text{if} \quad x \in X_N \quad \text{and} \quad y = X, \\ 1, & \text{if} \quad x \in X_C \quad \text{and} \quad y = s(x), \\ 0, & \text{if} \quad x \in X_C \quad \text{and} \quad y \neq s(x). \end{cases}$$

In the considered Markov process, for an arbitrary state $x \in X_C$, the transition $(x, s(x))$ from the states $x \in X_C$ to the states $y = s(x) \in X$ is made with the probability $p_{x,s(x)} = 1$ if the strategy s is applied. For this Markov process we can determine the average cost per transition for an arbitrary fixed starting state $x_i \in X$. Thus, we can determine the vector of average costs ω^s, which corresponds to the strategy s, according to the formula $\omega^s = Q^s \mu^s$, where Q^s is the limit matrix of the Markov process, generated by the stationary strategy s, and μ^s is the corresponding vector of the immediate costs, i.e., $\mu^s_x = \sum\limits_{y \in X(x)} p^s_{x,y} c_{x,y}$ [3]. A component ω^s_x of the vector ω^s represents the average cost per transition in our problem with a given starting state x and a fixed strategy s, i.e., $f_x(s) = \omega^s_x$.

In such a way we can define the value of the objective function $f_{x_0}(s)$ for the control problem on a network with a given starting state x_0, when the stationary strategy s is applied.

The control problem on the network (G, X_C, X_N, c, p, x_0) consists of finding a stationary strategy s^* for which

$$f_{x_0}(s^*) = \min_s f_{x_0}(s).$$

3. A Linear Programming Approach for Determining Optimal Stationary Strategies on Perfect Networks

We consider the stochastic control problem on the network (G, X_C, X_N, c, p, x_0) with $X_C \neq \emptyset$, $X_N \neq \emptyset$ and assume that G is a strongly connected directed graph. Additionally, we assume that in G for an arbitrary stationary strategy $s \in \mathbb{S}$ the subgraph $G_s = (X, E_s \cup E_N)$ is strongly connected. This means that the Markov chain induced by the probability transition matrix P^s is irreducible for an arbitrary strategy s. We call the decision network with such a condition a *perfect network*. At first we describe an algorithm for determining the optimal stationary strategies for the control problem on perfect networks.

We consider the control problem for which the average cost per transition is the same for an arbitrary starting state, i.e., $f_x(s) = \omega^s, \quad \forall x \in X$.

Let $s \in \mathbb{S}$ be an arbitrary strategy. Taking into account that for every fixed $x \in X_C$ we have a unique $y = s(x) \in X(x)$, we can identify the map s with the set of boolean values $s_{x,y}$ for $x \in X_C$ and $y \in X(x)$, where

$$s_{x,y} = \begin{cases} 1, & \text{if} \quad y = s(x), \\ 0, & \text{if} \quad y \neq s(x). \end{cases}$$

For the optimal stationary strategy s^* we denote the corresponding boolean values by $s^*_{x,y}$.

Assume that the network (G, X_C, X_N, c, p, x_0) is perfect. Then the following lemma holds.

Lemma 3.1. *A stationary strategy s^* is optimal if and only if it corresponds to an optimal solution q^*, s^* of the following mixed integer bilinear programming problem:*
Minimize

$$\psi(s, q) = \sum_{x \in X_C} \sum_{y \in X(x)} c_{x,y} s_{x,y} q_x + \sum_{z \in X_N} \mu_z q_z \tag{2}$$

subject to

$$\begin{cases} \sum\limits_{x \in X_C} s_{x,y} q_x + \sum\limits_{z \in X_N} p_{z,y} q_z = q_y, \quad \forall y \in X, \\ \sum\limits_{x \in X_C} q_x + \sum\limits_{z \in X_N} q_z = 1, \\ \sum\limits_{y \in X(x)} s_{x,y} = 1, \quad \forall x \in X_C, \\ s_{x,y} \in \{0, 1\}, \ \ \forall x \in X_C,\, y \in X; \quad q_x \geq 0, \ \ \forall x \in X, \end{cases} \tag{3}$$

where

$$\mu_z = \sum_{y \in X(z)} p_{z,y} c_{z,y}, \quad \forall z \in X_N.$$

Proof. Denote $\mu_x = \sum\limits_{y \in X(x)} c_{x,y} s_{x,y}$ for $x \in X_C$. Then μ_x for $x \in X_C$ and μ_z for $z \in X_N$ represent, respectively, the immediate cost of the system in the states $x \in X_C$ and $z \in X_N$ when the strategy $s \in S$ is applied. Indeed, we can consider the values $s_{x,y}$ for $x \in X_C$ and $y \in X(x)$ as probability transitions from the state $x \in X_C$ to the state $y \in X(x)$.

Therefore, for fixed s the solution $q^s = (q^s_{x_{i_1}}, q^s_{x_{i_2}}, \dots, q^s_{x_{i_n}})$ of the system of linear equations

$$(4) \qquad \begin{cases} \sum\limits_{x\in X_C} s_{x,y}q_x + \sum\limits_{z\in X_N} p_{z,y}q_z = q_y, \quad \forall y \in X, \\ \sum\limits_{x\in X_C} q_x + \sum\limits_{z\in X_N} q_z = 1, \end{cases}$$

corresponds to the vector of limit probabilities in the ergodic Markov chain determined by the graph $G_s = (X, E_s \cup E_N)$ with the probabilities $p_{x,y}$ for $(x, y) \in E_N$ and $p_{x,y} = s_{x,y}$ for $(x, y) \in E_C$ $(E_C = E \backslash E_N)$. Therefore, for given s the value

$$\psi(s, q^s) = \sum_{x\in X_C} \mu_x q_x + \sum_{z\in X_N} \mu_z q_z,$$

expresses the average cost per transition for the dynamical system if the strategy s is applied, i.e.,

$$f_x(s) = \psi(s, q^s), \quad \forall x \in X.$$

So, if we solve the optimization problem (2), (3) on a perfect network then we find the optimal strategy s^*.□

In the following for an arbitrary vertex $y \in X$ we will denote by $X_C^-(y)$ the set of vertices from X_C which contain directed leaving edges $e = (x, y) \in E$ that end in y, i.e., $X_C^-(y) = \{x \in X_C \mid (x, y) \in E\}$; in an analogues way we define the set $X^-(y) = \{x \in X \mid (x, y) \in E\}$.

Based on the lemma above we can prove the following result.

Theorem 3.1. *Let* $\alpha^*_{x,y}$ $(x \in X_C, y \in X)$, q^*_x $(x \in X)$ *be a basic optimal solution of the following linear programming problem:*
Minimize

$$(5) \qquad \overline{\psi}(\alpha, q) = \sum_{x\in X_C} \sum_{y\in X(x)} c_{x,y}\alpha_{x,y} + \sum_{z\in X_N} \mu_z q_z$$

subject to

$$(6) \qquad \begin{cases} \sum\limits_{x\in X_C^-(y)} \alpha_{x,y} + \sum\limits_{z\in X_N} p_{z,y}q_z = q_y, \quad \forall y \in X, \\ \sum\limits_{x\in X_C} q_x + \sum\limits_{z\in X_N} q_z = 1, \\ \sum\limits_{y\in X(x)} \alpha_{x,y} = q_x, \quad \forall x \in X_C, \\ \alpha_{x,y} \geq 0, \quad \forall x \in X_C,\ y \in X; \quad q_x \geq 0, \quad \forall x \in X. \end{cases}$$

Then the optimal stationary strategy s^* on a perfect network can be found as follows:

$$s^*_{x,y} = \begin{cases} 1, & if \quad \alpha^*_{x,y} > 0, \\ 0, & if \quad \alpha^*_{x,y} = 0, \end{cases}$$

where $x \in X_C$, $y \in X(x)$. Moreover, for every starting state $x \in X$ the optimal average cost per transition is equal to $\overline{\psi}(\alpha^*, q^*)$, i.e.,

$$f_x(s^*) = \sum_{x\in X_C} \sum_{y\in X(x)} c_{x,y}\alpha^*_{x,y} + \sum_{z\in X_N} \mu_z q^*_z$$

for every $x \in X$.

Proof. To prove the theorem it is sufficient to apply Lemma (3.1) and to show that the bilinear programming problem (2), (3) with boolean variables $s_{x,y}$ for $x \in X_C$, $y \in X$ can be reduced to the linear programming problem (5), (6). Indeed, we observe that the restrictions $s_{x,y} \in \{0, 1\}$ in the problems (2), (3) can be replaced by $s_{x,y} \geq 0$ because the optimal solutions after such a transformation of the problem are not changed. In addition, the restrictions

$$\sum_{y\in X(x)} s_{x,y} = 1, \quad \forall x \in X_C$$

can be changed by the restrictions

$$\sum_{y\in X(x)} s_{x,y}q_x = q_x, \quad \forall x \in X_C,$$

because for the perfect network it holds $q_x > 0$, $\forall x \in X_C$.

Based on the properties mentioned above in the problem (2), (3) we may replace the system (3) by the following system

$$(7) \qquad \begin{cases} \sum\limits_{x\in X^-_C(y)} s_{x,y}q_x + \sum\limits_{z\in X_N} p_{z,y}q_z = q_y, \quad \forall y \in X, \\ \sum\limits_{x\in X_C} q_x + \sum\limits_{z\in X_N} q_z = 1, \\ \sum\limits_{y\in X(x)} s_{x,y}q_x = q_x, \quad \forall x \in X_C, \\ s_{x,y} \geq 0, \ \forall x \in X_C, \ y \in X; \quad q_x \geq 0, \ \forall x \in X. \end{cases}$$

Thus, we may conclude that problem (2), (3) and problem (2), (7) have the same optimal solutions. Taking into account that for the perfect network $q_x > 0$, $\forall x \in X$ we can introduce in problem (2), (7) the notations $\alpha_{x,y} = s_{x,y}q_x$ for $x \in X_C$, $y \in X(x)$. This leads to the problem (5), (6). It is evident that $\alpha_{x,y} \neq 0$ if and only if $s_{x,y} = 1$. Therefore, the optimal stationary strategy s^* can be found according to the rule given in the theorem. □

So, if the network (G, X_C, X_N, c, p, x_0) is perfect then we can find the optimal stationary strategy s^* by using the following algorithm.

Algorithm 1. Determining the Optimal Stationary Strategy on Perfect Networks

1) Formulate the linear programming problem (5), (6) and find a basic optimal solution $\alpha^*_{x,y}$ $(x \in X_C,\ y \in X)$, q^*_x $(x \in X)$.
2) Fix a stationary strategy s^* where $s^*_{x,y} = 1$ for $x \in X_C$, $y \in X(x)$ if $\alpha^*_{x,y} > 0$; otherwise put $s^*_{x,y} = 0$.

4. Extension of the Algorithm 1 for Solving the Unichain Control Problem

We show that the algorithm 1 can be extended for the problem in which an arbitrary strategy s generates a Markov unichain. For this problem the graph G_s induced by a stationary strategy may not be strongly connected, but it contains a unique deadlock strongly connected component (which do not contain a leaving directed edge from the component to the exterior) that is reachable from every $x \in X$. A basic optimal solution α^*, q^* of the linear programming problem (5), (6) determines the strategy

$$s^*_{x,y} = \begin{cases} 1, & \text{if } \alpha^*_{x,y} > 0; \\ 0, & \text{if } \alpha^*_{x,y} = 0, \end{cases}$$

and a subset $X^* = \{x \in X \mid q^*_x > 0\}$, where s^* provides the optimal average cost per transition for the dynamical system $\mathbb{L}$ when it starts transitions in the states $x_0 \in X^*$. This means that for an arbitrary network algorithm 1 determines the optimal stationary strategy of the problem only in the case if the system starts transitions in the states $x \in X^*$.

For a unichain control problem algorithm 1 determines the strategy s^* and the recurrent class X^*. The remaining states $x \in X \setminus X^*$ in X correspond to transient states and the optimal stationary strategies in these states can be chosen in order to reach X^*.

We show how to use the linear programming model (5),(6) for determining the optimal stationary strategies of the control problem on the nonperfect network in which for an arbitrary stationary strategy s the matrix P^s corresponds to a recurrent Markov chain.

An arbitrary strategy s in G generates a graph $G_s = (X, E_s \cup E_N)$ with unique deadlock strongly connected component $G'_s = (X'_s, E'_s)$ that can be reached from any vertex $x \in X$. The optimal stationary strategy s^* in G can be found from a basic optimal solution by fixing $s^*_{x,y} = 1$ for the basic variables. This means that in G we can find the optimal stationary strategy by using the following algorithm:

Algorithm 2. Determining the Optimal Stationary Strategy for a Unichain Control Problem

1) Find a basic optimal solution α^*, q^* of the linear programming problem (5), (6) and the subset of vertices $X^* = \{x \in X | q^*_x > 0\}$ which in G corresponds to a strongly connected subgraph $G^* = (X^*, E^*)$.
2) If $x_0 \in X^*$ then we obtain the solution of the problem with fixed starting state x_0. To determine the solution of the problem for an arbitrary starting state we may select successively vertices $x \in X \setminus X^*$ which contain outgoing directed edges that end in X^* and will add them at each time to X^*, using the following rule:

- if $x \in X_C \cap (X \setminus X^*)$ then we fix an directed edge $e = (x, y)$, put $s^*_{x,y} = 1$ and change X^* by $X^* \cup \{x\}$;
- if $x \in X_N \cap (X \setminus X^*)$ then change X^* by $X^* \cup \{x\}$.

5. Algorithms for solving stochastic discrete control problems on networks with varying time of states'transitions of the dynamical system

We consider the stationary stochastic discrete optimal control problem on networks with an average cost optimization criterion, when the time of systems' transitions from one state to another may vary in the control process. The problem will be reduced to the case with unit time of states' transitions of the system.

Let a discrete dynamical system L with finite set of states X be given. At every discrete moment of time $t = t_0, t_1, t_2, ...$ the state of L is $x(t) \in X$ and at the starting moment of time $t_0 = 0$ the state of the dynamical system is $x_0 = x(0)$. Assume that the dynamics of the system is described by a directed graph of state's transitions $G = (X, E)$. An arbitrary vertex x of G corresponds to a state $x \in X$ and an arbitrary directed edge $e = (x, y) \in E$ expresses the possibility of the system L to pass from the state $x\left(t\right)$ to the state $x\left(t+\tau_{e}\right)$, where τ_e is the time of the system's transition from the state $x = x\left(t\right)$ to the state $y = x\left(t+\tau_{e}\right)$ through the edge $e = (x, y)$. So, on the edge set E it is defined the transition time function $\tau : E \to \mathbb{N}$ and, also, the cost function $c : E \to \mathbb{R}$, which associates to each edge the cost c_e of the system's transition from the state x to the state y.

We assume that the set X admit states in which system L makes transitions to a next state in the random way, according to given distribution function of probabilities on the set of possible transitions from these states. The set X is divided into two subsets X_C and X_N ($X = X_C \bigcup X_N, X_C \bigcap X_N = \emptyset$), where X_C represents the set of controllable states, and X_N represents the set of uncontrollable states.

The control problem on network (G, X_C, X_N, c, p, x_0) with an average cost optimization criterion consists in finding the stationary strategy s^* that provides the minimal mean integral-time cost by a trajectory.

We define a stationary strategy for the control problem as a map $s : x \to y \in X\left(x\right)$ for $x \in X_C$. For the arbitrary stationary strategy s the graph $G_s = (X,\, E_s \bigcup E_N)$, where $E_s = \left\{ e = (x, y) \in E \,|\; x \in X_C,\; y = s\left(x\right)\right\}$, corresponds to a Markov process with the probability matrix $P^s = \left(p^s_{x,y}\right)$, where

$$p^s_{x,y} = \begin{cases} p_{x,y}, & if\ x \in X_N\ and\ y \in X, \\ 1, & if\ x \in X_C\ and\ y = s(x), \\ 0, & if\ x \in X_C\ and\ y \neq s(x). \end{cases}$$

6. Reduction to the problem with unit time of states' transitions

Our problem can be reduced to the case with unit time of states' transitions on an auxiliary graph $G' = (X',\, E')$. Graph G' is obtained from G, where each directed edge $e = (x,\, y) \in E$ with corresponding transition time τ_e is changed by a sequence of directed edges $e'_1 = (x,\, x^e_1)\,,\;\; e'_2 = (x^e_1,\, x^e_2)\,,\, ...,\;\; e'_{\tau_e} = \left(x^e_{\tau_e - 1},\, y\right)$. So, the set of vertices X' of the graph G' consists of the set of states X and the set of intermediate states $XI = \{x^e_i |\; e \in E,\; i = 1, 2, ..., \tau_e - 1\}$, i.e., $X' = X \bigcup XI$. Also, we consider the sets X'_C and X'_N, so that $X' = X'_C \bigcup X'_N$, $X'_C = X_C$ and $X'_N = X' \backslash X_C$. The set of edges E' is defined as

$$E' = \bigcup_{e \in E} \mathcal{E}^e, \quad \mathcal{E}^e = \left\{ (x,\, x^e_1)\,,\;\; (x^e_1,\, x^e_2)\,,\, ...,\;\; \left(x^e_{\tau_e - 1},\, y\right) \Big|\;\; (x,\, y) \in E \right\}$$

and the cost function $c' : E' \to \mathbb{R}$ by $c'_{x,x_1^e} = c_{x,y}$ if $e = (x, y) \in E$, $c'_{x_1^e,x_2^e} = c_{x_2^e,x_3^e} = ... = c_{x^e_{\tau_e-1},y} = 0$. The probability function $p' : E'_N \to [0, 1]$ is defined as follows:

$$p'_{x',y'} = \begin{cases} p_{x,y}, & if \;\; x' = x, \;\; x' \in X_N \subset X'_N \;\; and \;\; y' = x_1^e, \\ \qquad\qquad 1, & if \;\; x' \in X'_N \backslash X_N. \end{cases}$$

Between the set of stationary strategies $s : x \to y \in X(x)$ for $x \in X$ and $s' : x' \to y' \in X'(x')$ for $x' \in X'$, there exists a bijective mapping such that the corresponding average costs on G and on G' are the same. So, if s'^* is the optimal stationary strategy of the problem with unit transitions on G' , then the optimal stationary strategy s^* on G is determined by fixing $s^*(x) = y$ if $s'^*(x) = x_1^e$, where $e = (x, y)$.

We consider that the network (G, X_C, X_N, c, p, x_0) is perfect, i.e., the graphs G and G_s are strongly connected. In this case the network $(G', X'_C, X'_N, c', p', x_0)$ is perfect.

Let $s' \in S'$ be an arbitrary strategy in G'. Taking into account that for every fixed $x' \in X'_C$ we have a unique $y' = s'(x') \in X'(x')$, we can identify the map s' with the set of boolean values $s'_{x',y'}$ for $x' \in X'_C$ and $y' \in X'(x')$, where

$$s'_{x',y'} = \begin{cases} 1, & if \;\; y' = s'(x'), \\ 0, & if \;\; y' \neq s'(x'). \end{cases}$$

For the optimal stationary strategy s'^* we denote the corresponding boolean values by $s'^*_{x',y'}$.

So, if the network (G, X_C, X_N, c, p, x_0) is perfect then we can find the optimal stationary strategy s^* by using the following algorithm.

Algorithm 3. Determining the Optimal Stationary Strategy for a Problem with Varying Time of States Transitions

1) Formulate the linear programming problem analogous to the problem (5), (6) and find a basic optimal solution $\alpha^*_{x',y'}$ $(x' \in X'_C, \;\; y' \in X')$, $q^*_{x'}$ $(x' \in X')$;
2) Fix a stationary strategy s'^* in G': put $s'^*_{x',y'} = 1$ for $x' \in X'_C$, $y' \in X'(x)$ if $\alpha^*_{x',y'} > 0$; otherwise put $s'^*_{x',y'} = 0$;
3) Fix a stationary strategy s^* in G: for each $(x', y') \in E'$ so that $s'^*_{x',y'} = 1$ put $s^*_{x',y} = 1$ for $y \in X(x')$, so that (x', y') is edge of a directed path from x' to y; otherwise put $s^*_{x',y} = 0$.

The algorithm 3 can be extended for the unichain control problem.

References

[1] R. Bellman , *Functional equations in the theory of dynamic programming, XI-Limit theorems*, Rand. Circolo Math. Palermo, 8(3) (1959), 343–345.

[2] D. Lozovanu , S. Pickl, *Algorithmic solution of discrete control problems on stochastic networks*, Procceedings of CTW09 Workshop on Graphs and Combinatorial Optimization, Paris, (2009), 221–224.

[3] M. Puterman, *Markov Decision Processes: Stochastic Dynamic Programming*, John Wiley, New Jersey (2005).

Moldova State University ,
Str. Alexe Mateevici 60, Chişinău MD-2009,
Republic of MOLDOVA,
e-mail: mariacapcelea@yahoo.com, tcapcelea@yahoo.com