STUDIA UNIVERSITATIS MOLDAVIAE, 2014, nr.2(72)
Seria "Științe exacte și economice" *ISSN 1857-2073* *ISSN online 2345-1033* *p.8-14*

ALGORITHMS FOR SOLVING STOCHASTIC DISCRETE CONTROL PROBLEMS ON NETWORKS WITH VARYING TIME OF STATES' TRANSITIONS OF THE DYNAMICAL SYSTEM

Maria CAPCELEA
Universitatea de Stat din Moldova

The stochastic versions of discrete optimal control problem on networks with varying time of state transitions of the dinamical system are studied. Polynomial time algorithms for determining the optimal stationary strategies in this problems are proposed.
Keywords: *discrete processes, stochastic optimal control problem, stationary strategies, linear programming approach, polynomial time algorithm.*

ALGORITMI PENTRU REZOLVAREA PROBLEMELOR DE CONTROL OPTIMAL DISCRET PE REȚELE CU TIMP DE TRECERE VARIABIL ÎNTRE STĂRILE SISTEMULUI DINAMIC
În lucrare se examinează variantele stocastice ale problemei de control optimal discret pe rețele cu timp variabil de tranziție între stările sistemului dinamic. Sunt propuși algoritmi polinomiali pentru determinarea strategiilor optime staționare.
Cuvinte-cheie: *procese discrete, problemă de control optimal stocastic, strategii staționare, metoda programării liniare, algoritm polinomial.*

1. Introduction and Problem Formulation

In this paper we consider the stationary case of the stochastic discrete optimal control problem on network, with an average cost optimization criterion, when the time of systems' transitions from one state to another may vary in the control process. The main results we describe are based on the control model with a fixed unit time of state's transitions, on linear programming approach and the concept of Markov decision processes. The problem will be reduced to the corresponding case of the problem with unit time of states' transitions of the system. The statement of the problem is the following.

Let a discrete dynamical system L with finite set of states X be given. At every discrete moment of time $t = t_0, t_1, t_2, \ldots$ 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 $x(t)$ to the state $x(t + \tau_e)$, where τ_e is the time of the system's transition from the state x to the state y through the edge $e = (x, y)$. So, on the edge set E it is defined the function $\tau : E \to \mathbb{N}$, which associates to each edge a natural number τ_e, which means that if the system L at the moment of time t is in the state $x = x(t)$, then the system can reach the state y at the moment $t + \tau_e$ if it passes through the edge $e = (x, y)$, i.e., $y = x(t + \tau_e)$. We assume that graph G does not contain deadlock vertices, i.e., for each x there exists at least one leaving directed edge $e = (x, y) \in E$. In addition, on the edge set E it is defined 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 $x = x(t)$ to the state $y = x(t + \tau_e)$ for an arbitrary discrete moment of time t. So, finally we have that to each edge $e = (x, y) \in E$ the cost c_e and the transition time τ_e from x to y are associated.

We assume that the set of states X of dynamical system may 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. So, the set of states X is divided into two subsets X_C and X_N ($X = X_C \bigcup X_N, X_C \bigcap X_N = \varnothing$), where X_C represents the set of controllable states $x \in X$ in which the transitions of the system to a next state y can be controlled by the decision maker at every discrete moment of time t and X_N represents the set of uncontrollable states $x \in X$, in which the decision maker is not able to control the transition because the system passes to a next state y randomly. The probability distribution function $p: E_N \to [0,1]$ on the set $E_N = \{e = (x,y) \in E \,|\, x \in X_N\}$ is defined such that $\sum_{y \in X^+(x)} p_{x,y} = 1$, $X^+(x) = \{y \in X \,|\, e = (x,y) \in E\}$. Here $p_{x,y}$ expresses the probability of system's transition from the state x to the state y for every discrete moment of time t. Note, that the condition $p_{x,y} = 0$ for a directed edge $e = (x,y)$ is equivalent with the condition that G does not contain this edge.

A directed edge $e = (x,y)$ in G corresponds to a stationary control of the system in the state $x \in X$ which provides a transition from $x = x(t)$ to $y = x(t + \tau_e)$ for every discrete moment of time t. A sequence of directed edges $\tilde{E} = \{e_0, e_1, \ldots, e_t, \ldots\}$, where $e_j = \left(x(t_j), x\left(t_j + \tau_{e_j}\right)\right)$, $j = 0,1,2,\ldots$, determines in G a control of the system with fixed starting state $x(0)$. An arbitrary control in G generates a trajectory $x(t_0), x(t_1), x(t_2), \ldots$, for which the mean integral-time cost by a trajectory can be defined by the formula $f\left(\tilde{E}\right) = \lim_{t\to\infty}\left(\frac{1}{\sigma}\sum_{j=0}^{t-1} c_{e_j}\right)$, where $\sigma = \sum_{j=0}^{t-1} \tau_{e_j}$.

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 on network as a map:

$$s: x \to y \in X^+(x) \text{ for } x \in X_C.$$

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

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

In the considered Markov process, for an arbitrary state $x \in X_C$, the transition $(x, s(x))$ from the state $x \in X_C$ to the state $y = s(x) \in X$ is made with the probability $p_{x,s(x)} = 1$ if the strategy s is applied.

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

We describe a general scheme how to reduce the stochastic control problems with varying time of states' transitions to the case with unit time of states' transition of the system. We show that our problem can be reduced to the problem with unit time of states' transitions on an auxiliary graph $G' = (X', E')$, which is obtained from $G = (X, E)$, using a special construction. This means that after such a reduction we can apply the linear programming approach described in [1,4]. Graph $G' = (X', E')$ with unit transitions on directed edges $e' \in E'$ is obtained from G, where each directed edge $e = (x,y) \in E$ with corresponding transition

STUDIA UNIVERSITATIS MOLDAVIAE, 2014, nr.2(72)
Seria "Științe exacte și economice" ISSN 1857-2073 ISSN online 2345-1033 p.8-14

time τ_e is changed by a sequence of directed edges $e'_1 = (x, x_1^e)$, $e'_2 = (x_1^e, x_2^e), \ldots, e'_{\tau_e} = (x_{\tau_e - 1}^e, y)$. This means that a transition from a state $x = x(t)$ at the moment of time t to the state $y = x(t+\tau_e)$ at the moment $t+\tau_e$ in G we represent in G' as the transition of a dynamical system from the state $x = x(t)$ at the time-moment t to $y = x(t+\tau_e)$ when the system makes transitions through a new fictive intermediate set of states $x_1^e, x_2^e, \ldots, x_{\tau-1}^e$ at the corresponding discrete moments $t+1, t+2, \ldots, t+\tau_e - 1$.

The graphical interpretation of this construction is represented in Fig. 1 and Fig. 2. In Fig.1 it is represented an arbitrary directed edge $e = (x, y)$ with the corresponding transition time τ_e in G. In Fig.2 it is represented the sequence of directed edges e'_i and the intermediate states $x_1^e, x_2^e, \ldots, x_{\tau_e-1}^e$ in G' that correspond to a directed edge $e = (x, y)$ in G. So, the set of vertices X' of the graph G' consists of the set of states X and the set of intermediate states $XI = \{x_i^e \mid e \in E,\ i = 1, 2, \ldots, \tau_e - 1\}$, i.e., $X' = X \cup XI$. Also, we consider the sets X'_C and X'_N, so that $X' = X'_C \cup X'_N$, $X'_C = X_C$ and $X'_N = X' \setminus X_C$. The set of edges E' is defined as follows:

$$E' = \bigcup_{e \in E} \mathcal{E}^e, \quad \mathcal{E}^e = \left\{ (x, x_1^e),\ (x_1^e, x_2^e), \ldots,\ (x_{\tau_e-1}^e, y) \,\middle|\, (x, y) \in E \right\}.$$

We define the cost function $c' : E' \to \mathbb{R}$ in the following way:

$$c'_{x, x_1^e} = c_{x,y} \quad \text{if } e = (x, y) \in E, \quad c'_{x_1^e, x_2^e} = c_{x_2^e, x_3^e} = \ldots = c_{x_{\tau_e-1}^e, y} = 0.$$

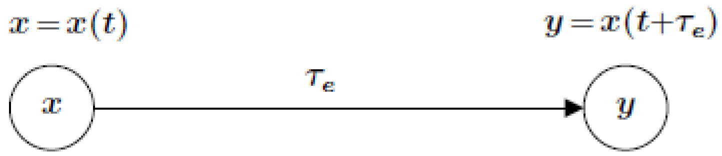

Fig. 1

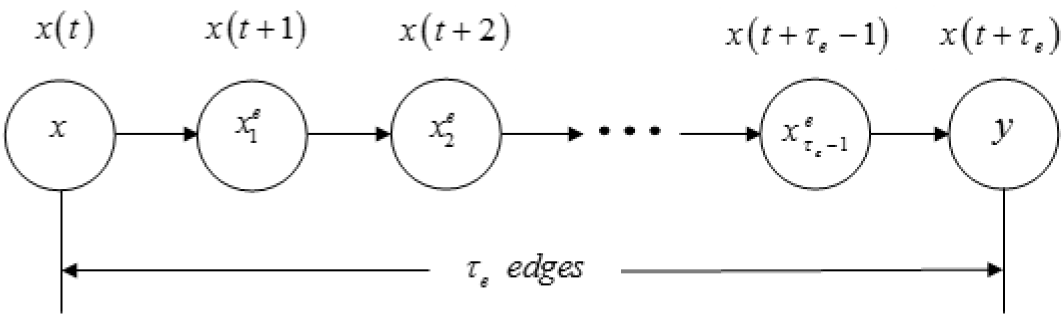

Fig. 2

The probability function $p' : E'_N \to [0,1]$ on the set $E'_N = \left\{ e' = (x', y') \in E' \,\middle|\, x' \in X'_N \right\}$ is defined as follows:

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

Between the set of stationary strategies $s : x \to y \in X^+(x)$ for $x \in X$ and the set of stationary strategies $s' : x' \to y' \in X'^+(x')$ for $x' \in X'$, there exists a bijective mapping, such that the corresponding average and discounted 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 $e=(x,y)$.

3. The main results

The linear programming algorithms for solving the control problem with a unit time of states' transitions have been developed in [1-4]. At first, we formulate an algorithm for determining the optimal stationary strategies for the control problem on perfect networks. Therefore, we consider the stochastic control problem on the network (G, X_C, X_N, c, p, x_0) with $X_C \neq \varnothing$, $X_N \neq \varnothing$, and assume that this network is perfect, i.e. the graph G and the subgraph $G_s = (X, E_s \cup E_N)$ (for an arbitrary stationary strategy $s \in S$) are strongly connected. In this case the network $(G', X'_C, X'_N, c', p', x_0)$ is perfect. The Markov chain induced by the probability transition matrix P^s is irreducible for an arbitrary strategy s.

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 $s'^*_{x',y'}$.

Based on the results from [1,4] we can prove the following

Theorem 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

$$\overline{\psi}(\alpha,q) = \sum_{x' \in X'_C} \sum_{y' \in X'^+(x')} c'_{x',y'} \alpha_{x',y'} + \sum_{z' \in X'_N} \mu_{z'} q_{z'}\,, \tag{1}$$

subject to

$$\begin{cases} \displaystyle\sum_{x' \in X'^-_C(y')} \alpha_{x',y'} + \sum_{z' \in X'_N} p_{z',y'} q_{z'} = q_{y'}, \quad \forall y' \in X', \\ \displaystyle\sum_{x' \in X'_C} q_{x'} + \sum_{z' \in X'_N} q_{z'} = 1, \\ \displaystyle\sum_{y' \in X'^+(x')} \alpha_{x',y'} = q_{x'}, \quad \forall x' \in X'_C, \\ \alpha_{x',y'} \geq 0,\ \forall x' \in X'_C,\ y' \in X';\ \ q_{x'} \geq 0,\ \forall x' \in X', \end{cases} \tag{2}$$

where $\mu_{z'} = \sum_{y' \in X'(z')} p'_{z',y'} c'_{z',y'}$, $\forall z' \in X'_N$, $X'^-_C(y') = \{x' \in X'_C \mid (x',y') \in E'\}$. *Then the optimal stationary strategy s'^* on a perfect network $(G', X'_C, X'_N, c', p', x_0)$ can be found as follows:*

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

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 $x' \in X'$.

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.

1) Formulate the linear programming problem (1), (2) 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 $s^*_{x',y} = 0$.

We can show that the stochastic control problem on the non-perfect network, in which an arbitrary strategy s generates a Markov unichain [5], can be reduced to an auxiliary problem on perfect network. So, the proposed algorithm can be extended for the unichain control problem. In this case the graph G_s, induced by a stationary strategy, may not be strongly connected, but it contains a unique strongly connected component, that can be reached from any vertex $x \in X$. For this control problem the mean integral-time cost by a trajectory is the same for an arbitrary starting state.

A basic optimal solution α^*, q^* of the linear programming problem (1), (2) determines the strategy $s'^*_{x',y'}$ and a positive recurrent class $X'^* = \{x' \in X' \mid q^*_{x'} > 0\}$ in X'. For a unichain control problem Algorithm 1 determines the optimal stationary strategy of the problem only in the case if the system starts transitions in the state $x'_0 \in X'^*$. The remaining states $x' \in X' \setminus X'^*$ correspond to transient states and the optimal stationary strategies in this states can be chosen in order to reach X'^*. So, in G' we can find the optimal stationary strategy as follows:

Algorithm 2.

1) Find a basic optimal solution α^*, q^* of the linear programming problem (1),(2) and the subset of vertices $X'^* = \{x' \in X' \mid q^*_{x'} > 0\}$ which in G' corresponds to a strongly connected subgraph $G'^* = (X'^*, E'^*)$.

2) On G'^* we determine the optimal solution of the problem using the Algorithm 1.

3) 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:

a) 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 $X'^* \cup \{x'\}$;

b) if $x' \in X'_N \cap (X' \setminus X'^*)$ then change X'^* by $X'^* \cup \{x'\}$.

Now we consider the infinite horizon discounted stochastic control problem with varying time of states' transitions. The dynamics of the system is determined in the same way as for the problem with average cost, but the objective function which has to be minimized is defined by the sum $\sum_{j=0}^{\infty} \gamma^{t_j} c_{e_j}$, where $\gamma \in (0,1)$ is a given discount factor. An arbitrary control in the graph of states' transitions G generates a trajectory

$x(t_0), x(t_1), x(t_2), \dots,$ for which the discounted expected total cost is defined as follows $\sigma_{x_0}^{\gamma}(s) = \sum_{j=0}^{\infty} \gamma^{t_j} c_{e_j}$. We are seeking for a stationary strategy s^* such that $\sigma_{x_0}^{\gamma}(s^*) = \min_{s} \sigma_{x_0}^{\gamma}(s)$.

We develop a linear programming approach for the discounted stochastic control problem on the network (G, X_C, X_N, c, p, x_0) using the same scheme as above. We have the following result.

Theorem 2. *Let $\alpha^*_{x',y'}\ (x' \in X'_C, y' \in X')$, $\beta^*_{x'}\ (x' \in X')$ be an optimal solution of the following linear programming problem:*

Minimize

$$\phi_{x'_0}(\alpha, \beta) = \sum_{x' \in X'_C} \sum_{y' \in X'^+(x')} c_{x',y'} \alpha_{x',y'} + \sum_{z' \in X'_N} \mu_{z'} \beta_{z'}, \tag{3}$$

subject to

$$\begin{cases} \beta_{y'} - \gamma \sum\limits_{x' \in X'^-_C(y')} \alpha_{x',y'} - \gamma \sum\limits_{z' \in X'^-_N(y')} p_{z',y'} \beta_{z'} = 1, \quad y' = x'_0, \\ \beta_{y'} - \gamma \sum\limits_{x' \in X'^-_C(y')} \alpha_{x',y'} - \gamma \sum\limits_{z' \in X'^-_N(y')} p_{z',y'} \beta_{z'} = 0, \quad \forall y' \in X' \setminus \{x'_0\}, \\ \sum\limits_{y' \in X'^+(x')} \alpha_{x',y'} = \beta_{x'}, \quad \forall x' \in X'_C, \\ \beta_{x'} \ge 0,\ \forall x' \in X',\ \alpha_{x',y'} \ge 0, \quad \forall x' \in X'_C,\ y' \in X'^+(x'), \end{cases} \tag{4}$$

where $\mu_{z'} = \sum_{y' \in X'^+(z')} c_{z',y'} p_{z',y'}, \quad \forall z' \in X'_N$. *If in $G' = (X', E')$ an arbitrary vertex $x' \in X'$ is attainable from x'_0, then $\beta^*_{x'} > 0,\ \forall x' \in X'_C$ and*

$$\frac{\alpha^*_{x',y'}}{\beta^*_{x'}} \in \{0,1\},\ \forall x' \in X'_C,\ y' \in X'^+(x').$$

The optimal stationary strategy s'^* of the discounted stochastic control problem on the network can be found by fixing $s'^*_{x',y'} = \dfrac{\alpha^*_{x',y'}}{\beta^*_{x'}}$, for $x' \in X'_C$ and every $y' \in X'^+(x')$ if $\beta^*_{x'} \ne 0$; otherwise we put $s^*_{x',y'} = 0$.

Based on the Theorem 2 we can propose the following algorithm for determining the optimal solution of the discounted stochastic control problem on the network, with fixed starting state x'_0 and varying time of states' transitions.

Algorithm 3.

1. Formulate the linear programming problem (3), (4);
2. Determine an optimal solution $\alpha^*_{x',y'}\ (x' \in X'_C,\ y' \in X')$, $\beta^*_{x'}\ (x' \in X')$ of the problem (3), (4) and fix a stationary strategy s'^* in G': put $s'^*_{x',y'} = \alpha^*_{x',y'} / \beta^*_{x'}$ for $x' \in X'_C,\ y' \in X'^+(x')$ if $\beta^*_{x'} \ne 0$; 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 $s^*_{x',y} = 0$.

4. Conclusion

A linear programming approach for finding the optimal stationary strategies of stochastic discrete optimal control problems with infinite time horizon and varying time of states' transitions is proposed. Polinomial time algorithms based on such approach for solving the considered problems on networks are developed.

Bibliography:

1. CAPCELEA, M. Linear programming approach for stochastic discrete optimal control problems with infinite time horizon. In: *Proceedings of the International Student Conference on Pure and Applied Mathematics, ISCOPAM 2010.* Iasi: ed. Universității „Al.I. Cuza", 2011, p.41-54.
2. LOZOVANU, D., CAPCELEA, M. Algorithms for determining optimal stationary strategies of discounted stochastic optimal control problem on networks. In: *Materialele Conferinței Internaționale "Modelare Matematică, Optimizare şi Tehnologii Informaționale*", vol. I, 24-26 martie 2010, Ediția a II-a, Chişinău: Evrica, 2010, p.50-55.
3. LOZOVANU, D., PICKL, S. Discounted Markov decision processes and algorithms for solving stochastic control problem on networks. In: *Proceedings of the Conference „10th Cologne-Twente Workshop on Graphs and Combinatorial Optimization CTW 2011*", June 14-16, 2011, p.194-197.
4. LOZOVANU, D., PICKL, S. Optimal stationary control of discrete processes and a polynomial time algorithm for stochastic control problem on networks. In: *Proceedings of International Conference on Computational Science, ICCS 2010. Amsterdam, Elsevier, Procedia Computer Science (1)*, 2012, p.1417-1426.
5. PUTERMAN, M. *Markov decision processes*. Wiley, 1993.

Acknowledgments. The research reported in this paper has been supported in part by the project 14.819.02.14F "Numerical methods for solving stochastic optimization problems".

Prezentat la 09.06.2014