LINEAR PROGRAMMING APPROACH FOR SOLVING STOCHASTIC CONTROL PROBLEM ON NETWORKS WITH DISCOUNTED TRANSITION COSTS

D. D. Lozovanu [1], S. W. Pickl [2], M. A. Capcelea [3]

[1]*Institute of Mathematics and Computer Science, Academy of Sciences*
Chisinau, Moldova
E-mail: lozovanu@math.md
[2]*Institut für Theoretische Informatik, Mathematik und Operations Research,*
Fakultät fur Informatik, Universität der Bundeswehr
München, Germany
E-mail: stefan.pickl@unibw.de
[3]*Moldova State University, Faculty of Mathematics and Informatics*
Chisinau, Moldova
E-mail: ylan12@yahoo.com

The infinite horizon stochastic control problem on network with expected total discounted cost optimization criterion is studied. A linear programming approach for solving this problem on networks is developed. Moreover, a polynomial time

algorithm for determining the optimal stationary strategies in the considered problem is proposed and grounded.

Key words: Stochastic Control Problem, Discounted Costs, Optimal Stationary Strategies, Linear Programming Approach, Polynomial Time Algorithm.

INTRODUCTION AND PROBLEM FORMULATION

In this paper we consider the infinite horizon stochastic control problem with expected total discounted cost optimization criterion. We formulate and study this problem on networks.

Let L be a time-discrete system with finite set of states X. At every discrete moment of time $t = 0,1,2,\ldots$ the state of the dynamical system is $x(t) \in X$. Two states $x_0, x_f \in X$ are given where $x_0 = x(0)$ is the staring state of L and x_f is the state in which the dynamical system must be brought. The dynamics of system L is described by a directed graph $G = (X, E)$ in which the set of vertices X corresponds to the set of states of system L and a directed edge $e = (x, y) \in E$ signifies the possibility of system L to pass from the state $x = x(t) \in X$ to the state $y = x(t+1) \in X$ for arbitrary discrete moment of time $t = 0,1,2,\ldots$. Thus the set of edges $E(x) = \{e = (x, y) \in E \mid y \in X\}$ corresponds to the set of feasible controls in the state $x = x(t)$. On edge set E is defined a function $c : E - R$ which gives to each directed edge $e = (x, y)$ a cost $c_e = c_{x,y}$, i. e. $c_{x,y}$ expresses the cost of system's transition from the state $x = x(t)$ to the state $y = x(t)$ for arbitrary $t = 0,1,2,\ldots$. The transitions costs of the dynamical system we consider with given discount factor γ which means that if the system at the moment of time t passes from the state $x = x(t)$ to the state $y = x(t+1)$ then the the cost $c_{x,y}$ is multiplied by γ^t, i.e. the cost of system's transition from the state $x = x(t)$ at the moment of time t to the state $y = x(t+1)$ is equal to $\gamma^t c_{x,y}$.

For the control problem on network we have the following behavior of the dynamical system. If $x(0) \in X_1$ then the decision maker select a transition to the state $x(1) = x_1$ and we obtain the state x_1 at the moment of time $t = 1$; in the case $x(0) \in X_2$ the system passes from $x(0)$ to a state $x(1)$ in the random way. If at the moment of time $t = 1$ the state x_f is reached ($x_1 = x_f$), then the dynamical system stops transitions. In general case, if at the moment of time t the system is in the state $x(t) \in X_1$ and $x(t) \neq x_f$ then the decision maker select the transition to the next state $x(t+1)$; if $x(t) \in X_2$ then the system passes to to a state $x(t+1)$ in the random way. The dynamical system stops transitions if $x(t) = x_f$. It is evident that the final states x_f in this process can be reached with the probability that depends on probability distribution functions in uncontrollable states as well on control in controllable states. We assume that G possesses the property that the final state x_f can be reached with the probability equal to 1 for an arbitrary control and consider the problem of determining the control with with minimal expected total discounted cost. In general case we can estimate the optimal expected total discounted cost if the system have been reached the final state. Additionally we assume that the decision maker uses the stationary control (stationary strategy). A stationary strategy we define as a map

$$s : x \to y \in X_1(x) \ for \ x \in X_1 \setminus \{x_f\},$$

which uniquely determine the transitions from the states $x = x(t) \in X_1 \setminus \{x_f\}$ to the states $y = s(x) \in X$ for arbitrary discrete moment of time $t = 0,1,2,\dots$. In the terms of stationary strategies the discounted stochastic control problem on network can be formulated in the following way.

Let s be an arbitrary stationary strategy. We define the graph $G_s = (X, E_s \cup E_2)$, where $E_s = \{e = (x,y) \in E \mid x \in X_1, \ y = s(x)\}$, $E_2 = \{e = (x,y) \mid x \in X_2, \ y \in X\}$. This graph corresponds to a Markov process with the probability matrix $P^s = (p^s_{x,y})$, where

$$p^s_{x,y} = \begin{cases} p_{x,y}, & if \quad x \in X_2 \quad and \quad y = X; \\ 1, & if \quad x \in X_1 \quad and \quad y = s(x); \\ 0, & if \quad x \in X_1 \quad and \quad y \neq s(x). \end{cases}$$

For this Markov process with transition costs c_e, $e \in E$ we define the expected total discounted cost $\sigma^{\gamma}_{x_0}(s)$. We consider the problem of determining the strategy s^* for which

$$\sigma^{\gamma}_{x_0}(s^*) = \min_s \sigma^{\gamma}_{x_0}(s).$$

We show that this problem can be formulated as linear programming problem and the optimal stationary strategies can be found using a polynomial time algorithm.

THE MAIN RESULTS

To study and solve the considered stochastic control problem on networks we use the framework of discounted Markov decision process (X, A, c, p) with given discount factor $\gamma \, (0 < \gamma < 1)$, where X is the set of states of the system, A is the set of actions, $c : X \times X \to R$ is the transition cost function that gives the costs $c_{x,y}$ of system's transitions from the states $x \in X$ to $y \in X$ and $p : X \times X \times A \to [0,1]$ is the transition probability function that satisfy the condition $\sum_{y \in X} p^a_{x,y} = 1, \forall x \in X, a \in A(x)$ (see [3]). Here $A(x)$ is the set of actions in the state $x \in X$. A stationary strategy (a policy) in this Markov decision process we define as a map $s : X \to A$ that determines for each $x \in X$ an action $a \in A(x)$. We consider a Markov decision processes with given absorbing state x_f ,where $c_{x_f, x_f} = 0$. For such Markov processes we consider the problem of determining the stationary strategy s^* that provide optimal expected total discounted costs for an arbitrary starting state. We sow that optimal solution of the problem can be found by solving the following linear programming problem:

Minimize

$$\varphi(\alpha, \beta) = \sum_{x \in X} \sum_{a \in A(x)} \mu_{x,a} \alpha_{x,a} \tag{1}$$

subject to

$$\begin{cases} \beta_y - \gamma \sum_{x \in X} \sum_{a \in A(x)} p_{x,y}^a \alpha_{x,a} \ge 1, \ y \in X; \\ \sum_{a \in A(x)} \alpha_{x,a} = \beta_x, \ \forall x \in X \setminus \{x_f\}; \\ \beta_y \ge 0, \ \forall y \in X \setminus \{x_f\}; \ \alpha_{x,a} \ge 0, \ \forall x \in X \setminus \{x_f\}, a \in A(x). \end{cases} \tag{2}$$

The optimal stationary strategy s^* for the problem with absorbing state $z \in X$ is determined as follows: for $x \in X \setminus \{x_f\}$ fix $a = s^*(x)$ if $\alpha_{x,a}^* \ne 0$. Using the dual lineal programming model for problem (1), (2) we show that for determining the optimal stationary strategies in the discounted stochastic control problem can be used the following linear programming problem:

Maximize

$$\varphi(\sigma) = \sum_{x \in X} \sigma_x \tag{3}$$

subject to

$$\begin{cases} \sigma_x - \gamma\sigma_y \le c_{x,y}, \forall x \in X_1, y \in X(x); \\ \sigma_x - \gamma \sum_{y \in X} p_{x,y} \sigma_y \le \mu_x, \quad \forall x \in X_2 \setminus \{z\}. \end{cases} \tag{4}$$

The optimal stationary strategy for the problem on network can be determined by fixing $s^* : X \setminus \{z\} \to A$ such that $s^*(x) = y \in X^*(x) \, \forall x \in X_1 \setminus \{z\}$, where $X^*(x) = \{y \in X(x) \mid \sigma_x - \gamma\sigma_y = c_{x,y}\}$. Note that the considered model is valid also for $\gamma = 1$. It is easy to observe that the linear programming model (3), (4) in the case $X_2 = \emptyset, \gamma = 1$ is transformed in the linear programming model for minimum cost problem in a weighted directed graph with fixed sink vertex x_f. Based on results from [1, 2] we have elaborated an iterative algorithm for determining the optimal stationary strategy for control problem on network with positive costs.

REFERENCES

1. *Lozovanu, D.* Determining optimal stationary strategies for discounted stochastic optimal control problem on networks / D. Lozovanu, S. Pickl. 9th Cologne-Twente Workshop on Graphs and Combinatorial Optimization (Proceedings), Cologne, 2010, P. 115–118.
2. *Lozovanu, D.* Discounted Markov Decision Processes and Algorithms for Solving Stochastic Control Problem on Networks / D. Lozovanu, S. Pickl. 9th Cologne-Twente Workshop on Graphs and Combinatorial Optimization (Proceedings), Rome, 2011, P. 94–99.
3. *Puterman, M.* Markov Decision Processes:Stochastic Dynamic Programming / M. Puterman. John Wiley, New Jersey, 2005, P. 649.