STOCHASTIC GAMES WITH REAL TRANSITION COSTS AND FINAL SEQUENCE OF STATES

Alexandru LAZARI

Moldova State University, Republic of Moldova (lazarialexandru@mail.md)

DOI: 10.19062/1842-9238.2017.15.2.7

Abstract: *The stochastic games with unit transition time and final sequence of states were previously investigated and an efficient algorithm for determining the game duration was developed. The purpose of this paper is to generalize this problem for the case when the transition costs are real and to elaborate a method for determining the expectation of the game outcome. The elaborated approach is based on extended network method, that allows us to reduce the real transition costs case to unit transition time case.*

Keywords: *stochastic game, final sequence of states, game outcome, extended network.*

1. PROBLEM FORMULATION

Let L_P be a stochastic system with finite set of states V_P, $|V_P| = \omega_P$. The system L_P starts its evolution from the state $v \in V_P$ with probability $p_P^*(v)$, $\forall v \in V_P$. Also, at every step t, the system L_P passes from the state $u \in V_P$ to the state $v \in V_P$ with probability $p_P(u,v)$ and real transition cost $c_P(u,v)$, that do not depend on t.

We assume that a sequence of states $X_P = (x_1, x_2, \ldots, x_m)$ is given and the system L_P stops transitions as soon as the states $x_1, x_2, \ldots, x_m$ are reached consecutively in given order. We denote by C_P the evolution outcome of the system L_P, i.e. $C_P = \sum_{(u,v)\in E_P} c_P(u,v)$, where E_P is the set of transitions that compose the evolution of the given system L_P.

The system L_P represents a Markov process with set of states V_P, initial distribution of the states $p_P^* = (p_P^*(v))_{v\in V_P}$, probability transition matrix $p_P = (p_P(u,v))_{u,v\in V_P}$ and real transition cost matrix $c_P = (c_P(u,v))_{u,v\in V_P}$. One more thing that is specific to the system L_P is the property to have a stopping rule, represented by final sequence of states X_P.

Several interpretations of these extended Markov processes were analyzed in [1], [6] and [7]. Also, polynomial algorithms for determining the main probabilistic characteristics of the evolution outcome of the given stochastic system L_P were proposed.

Next, the following game Γ_P, that represents a generalization of the games defined in [2] and [3], is considered. Initially, each player Π_l, $l = \overline{0, r-1}$, defines his stationary strategy, represented by one transition matrix $p_P^{(l)} = (p_P^{(l)}(u,v))_{u,v\in V_P}$, $l = \overline{0, r-1}$. The initial distribution of the states is $p_P^* = (p_P^*(v))_{v\in V_P}$. The game is started by first player Π_0.

At every moment of time, the system L_{P} passes consecutively to the next state according to the strategy of the current player. After the last player Π_{r-1}, the first player Π_0 acts on the system evolution and the game continues in this way until the given final sequence of states X_{P} is achieved. The winner is the player who acts the last on the system evolution.

Our goal is to study the game outcome C_{P}, knowing the initial distribution of the states $p^*_{\mathrm{P}} = (p^*_{\mathrm{P}}(v))_{v\in V_{\mathrm{P}}}$, the real cost transition matrix $c_{\mathrm{P}} = (c_{\mathrm{P}}(u,v))_{u,v\in V_{\mathrm{P}}}$, the stationary strategy $p^{(l)}_{\mathrm{P}} = (p^{(l)}_{\mathrm{P}}(u,v))_{u,v\in V_{\mathrm{P}}}$ of each player Π_l, $l=\overline{0,r-1}$ and the final sequence of states X_{P} of the system L_{P}. We will show how this problem can be reduced to the unit transition costs case using extended network method and how to find the win probabilities of the players.

2. TRANSITION COSTS DISCRETIZATION

2.1 Reduction to the positive real transition costs case. First of all, our goal is to reduce the game network to the positive real transition costs case, for avoiding the problems that can be encountered when we deals with negative and zero numbers.

Let us consider a new game Γ_{P^+}, defined on stochastic system L_{P^+} with the same final sequence of states $X_{\mathrm{P}^+} = X_{\mathrm{P}}$, finite set of states $V_{\mathrm{P}^+} = V_{\mathrm{P}}$, initial distribution of the states $p^*_{\mathrm{P}^+} = p^*_{\mathrm{P}}$ and stationary strategy $p^{(l)}_{\mathrm{P}^+} = p^{(l)}_{\mathrm{P}}$ of each player Π_l, $l=\overline{0,r-1}$. The difference between the games Γ_{P^+} and Γ_{P} consists in the way how the transition costs are defined.

Let $c^{\min}_{\mathrm{P}} = \min\limits_{u,v\in V_{\mathrm{P}}} c_{\mathrm{P}}(u,v)$ be the minimal transition cost in the game Γ_{P}. We define the transition costs in the new game Γ_{P^+} in the following way:

$$c_{\mathrm{P}^+}(u,v) = c_{\mathrm{P}}(u,v) - c^{\min}_{\mathrm{P}} + 1 \ge 1,\ \forall u,v \in V_{\mathrm{P}^+}\ . \tag{1}$$

Next, we consider a new game Γ_1, defined on stochastic system L_1 with the same final sequence of states $X_1 = X_{\mathrm{P}}$, finite set of states $V_1 = V_{\mathrm{P}}$, initial distribution of the states $p^*_1 = p^*_{\mathrm{P}}$ and stationary strategy $p^{(l)}_1 = p^{(l)}_{\mathrm{P}}$ of each player Π_l, $l=\overline{0,r-1}$. The transition costs in the new game Γ_1 are all equal to 1.

It is easy to observe that the following relation holds:

$$C_{\mathrm{P}} = C_{\mathrm{P}^+} + (c^{\min}_{\mathrm{P}} - 1)\cdot C_1, \tag{2}$$

where C_1 is the outcome of the game Γ_1, C_{P^+} is the outcome of the game Γ_{P^+} and C_{P} is the outcome of the game Γ_{P}. Indeed, we have

$$\begin{aligned} C_{\mathrm{P}^+} &= \sum_{(u,v)\in E_{\mathrm{P}^+}} c_{\mathrm{P}^+}(u,v) = \sum_{(u,v)\in E_{\mathrm{P}}} \left(c_{\mathrm{P}}(u,v) - c^{\min}_{\mathrm{P}} + 1\right) = \sum_{(u,v)\in E_{\mathrm{P}}} c_{\mathrm{P}}(u,v) - \sum_{(u,v)\in E_{\mathrm{P}}} \left(c^{\min}_{\mathrm{P}} - 1\right) = \\ &= \sum_{(u,v)\in E_{\mathrm{P}}} c_{\mathrm{P}}(u,v) - (c^{\min}_{\mathrm{P}} - 1) \sum_{(u,v)\in E_{\mathrm{P}}} 1 = C_{\mathrm{P}} - (c^{\min}_{\mathrm{P}} - 1)\cdot C_1. \end{aligned}$$

2.2 Reduction to the positive rational transition costs case. At this step, our goal is to reduce the game network to the case when the transition costs of related stochastic system are positive rational values.

Similar to the approach elaborated in [4], we will approximate the real cost of each transition with a rational number, such that the accuracy for the expectation of the game outcome to be less than a given value $\varepsilon > 0$.

We consider the game Γ_{Θ^+}, defined on stochastic system L_{Θ^+} with the same final sequence of states $X_{\Theta^+} = X_{\mathrm{P}^+}$, finite set of states $V_{\Theta^+} = V_{\mathrm{P}^+}$, initial distribution of the states $p^*_{\Theta^+} = p^*_{\mathrm{P}^+}$ and stationary strategy $p^{(l)}_{\Theta^+} = p^{(l)}_{\mathrm{P}^+}$ of each player Π_l, $l = \overline{0, r-1}$. We define the transition costs in the new game Γ_{Θ^+} in the following way:

$$c_{\Theta^+}(u,v) = c_{\mathrm{P}^+}(u,v) + \varepsilon(u,v),\ \forall u,v \in V_{\Theta^+}, \tag{3}$$

choosing the nonnegative value $\varepsilon(u,v)$ such that $c_{\Theta^+}(u,v) \in \Theta^+$ for each $u,v \in V_{\Theta^+}$ and

$$\left| \mathrm{E}\left(C_{\Theta^+}\right) - \mathrm{E}\left(C_{\mathrm{P}^+}\right) \right| \le \varepsilon, \tag{4}$$

where the accuracy $\varepsilon > 0$ is given.

Consider that

$$0 \le \varepsilon(u,v) \le \varepsilon^* < 1,\ \forall u,v \in V_{\Theta^+}. \tag{5}$$

We obtain

$$\begin{aligned} C_{\Theta^+} &= \sum_{(u,v)\in E_{\Theta^+}} c_{\Theta^+}(u,v) = \sum_{(u,v)\in E_{\mathrm{P}^+}} \left(c_{\mathrm{P}^+}(u,v) + \varepsilon(u,v)\right) = \sum_{(u,v)\in E_{\mathrm{P}^+}} c_{\mathrm{P}^+}(u,v) + \sum_{(u,v)\in E_{\mathrm{P}^+}} \varepsilon(u,v) = \\ &= \sum_{(u,v)\in E_{\mathrm{P}^+}} c_{\mathrm{P}^+}(u,v) + \varepsilon(u,v) \cdot \sum_{(u,v)\in E_{\mathrm{P}^+}} 1 = C_{\mathrm{P}^+} + \varepsilon(u,v) \cdot C_1, \end{aligned}$$

that implies

$$\left| \mathrm{E}\left(C_{\Theta^+}\right) - \mathrm{E}\left(C_{\mathrm{P}^+}\right) \right| = \left| \mathrm{E}\left(C_{\Theta^+} - C_{\mathrm{P}^+}\right) \right| = \mathrm{E}(\varepsilon(u,v) \cdot C_1) = \varepsilon(u,v) \cdot \mathrm{E}(C_1) \le \varepsilon^* \mathrm{E}(C_1). \tag{6}$$

If we have $\varepsilon^* E(C_1) \le \varepsilon$, then also the inequality $\left| \mathrm{E}\left(C_{\Theta^+}\right) - \mathrm{E}\left(C_{\mathrm{P}^+}\right) \right| \le \varepsilon$ holds.

Consider an arbitrary value $\varepsilon^* \in \left(0, \min\{1, \varepsilon / \mathrm{E}(C_1)\}\right]$. If the natural number s satisfies the condition $\varepsilon^* > 10^{-s}$, then we can chose

$$\varepsilon(u,v) = 10^{-s}(1 - \{10^s c_{\mathrm{P}^+}(u,v)\}),\ \forall u,v \in V_{\Theta^+} \tag{7}$$

and these values verify the relations (3), (4) and (5), where by $\{a\}$ is denoted the fractional part of the arbitrary real number a. From relations (3) and (7) we obtain

$$c_{\Theta^+}(u,v) = 10^{-s}(1 + \lfloor 10^s c_{\mathrm{P}^+}(u,v) \rfloor),\ \forall u,v \in V_{\Theta^+}, \tag{8}$$

where by $\lfloor a \rfloor$ is denoted the integer part of the arbitrary real number a. So, in these conditions, we can consider $E\left(C_{\mathrm{P}^+}\right) \approx E\left(C_{\Theta^+}\right)$ with error of approximation ε.

2.3 Reduction to the positive integer transition costs case. At this step, our goal is to reduce the game network to the case when the transition costs of related system are positive integer values. We will do this transformation by changing the measurement unit, i.e. we will multiply each positive rational transition cost $c_{\Theta^+}(u,v)$, $\forall u,v \in V_{\Theta^+}$, by μ, where μ is the least common multiple of the denominators of these rational transition costs.

Let us consider a new game Γ_{N^*}, defined on stochastic system L_{N^*} with the same final sequence of states $X_{N^*} = X_{\Theta^+}$, finite set of states $V_{N^*} = V_{\Theta^+}$, initial distribution of the states $p^*_{N^*} = p^*_{\Theta^+}$ and stationary strategy $p^{(l)}_{N^*} = p^{(l)}_{\Theta^+}$ of each player Π_l, $l = \overline{0, r-1}$. We define the transition costs in the new game Γ_{N^*} in the following way:

$$c_{N^*}(u,v) = \mu \cdot c_{\Theta^+}(u,v),\ \forall u,v \in V_{N^*}. \tag{9}$$

In this way, we will have $c_{N^*}(u,v) \in N^*$, $\forall u,v \in V_{N^*}$. We obtain

$$C_{N^*} = \sum_{(u,v)\in E_{N^*}} c_{N^*}(u,v) = \sum_{(u,v)\in E_{N^*}} \left(\mu \cdot c_{\Theta^+}(u,v)\right) = \mu \cdot \sum_{(u,v)\in E_{\Theta^+}} c_{\Theta^+}(u,v) = \mu \cdot C_{\Theta^+}, \tag{10}$$

that implies $C_{\Theta^+} = \frac{1}{\mu} \cdot C_{N^*}$.

2.4 Players' states synchronization. Next step prepares the transition costs for extended network method application. It is necessary that, the application of extended network method do not change the player who acts in the state v at each transition (u, v). From this reason, we will ensure that, after this transformation, each positive integer transition cost is congruent to 1 modulo r, where r is the number of the players.

Let us consider a new game Γ_f, defined on stochastic system L_f with the same final sequence of states $X_f = X_{N^*}$, finite set of states $V_f = V_{N^*}$, initial distribution of the states $p^*_f = p^*_{N^*}$ and stationary strategy $p^{(l)}_f = p^{(l)}_{N^*}$ of each player Π_l, $l = \overline{0, r-1}$. We define the transition costs in the new game Γ_f in the following way:

$$c_f(u,v) = r \cdot c_{N^*}(u,v) + 1,\ \forall u,v \in V_f, \tag{11}$$

i.e. we apply the function $f : N^* \to N^*$, $f(t) = rt + 1$, to each positive integer transition cost.

In this way, we will have $c_f(u,v) \in N^*$ and $c_f(u,v) \equiv 1 (\mathrm{mod}\ r)$, $\forall u,v \in V_f$. Also,

$$C_f = \sum_{(u,v)\in E_f} c_f(u,v) = \sum_{(u,v)\in E_f} \left(r \cdot c_{N^*}(u,v) + 1\right) = r \cdot \sum_{(u,v)\in E_{N^*}} c_{N^*}(u,v) + C_1 = r \cdot C_{N^*} + C_1, \tag{12}$$

that implies $C_{N^*} = \frac{1}{r} \cdot \left(C_f - C_1\right)$.

3. EXTENDED NETWORK METHOD

3.1 Reduction to the unit transition costs case. At this step, our goal is to reduce the game network to the case when the transition costs of related stochastic system are equal to 1. Similar with approach presented in [5], we will do this transformation by applying extended network method, i.e. each edge (u,v) with positive integer transition cost $c_f(u,v)$ will be replaced with $c_f(u,v)$ edges with unit transition cost, $\forall u,v \in V_f$. In this way, the total transition cost from arbitrary state $u \in V_f$ to another state $v \in V_f$ will not be changed, it is necessary only to ensure that also the transition probability from the state u to the state v will remain unchanged after extended network method application.

Let us consider a new game Γ_{ext}, defined on stochastic system L_{ext} in the following way. For each state $v \in V_f$ we choose the state $u \in V_f$ such that $c_f(u,v) = \max_{z \in V_f^-(v)} c_f(z,v)$, where $V_f^-(v) = \left\{ z \in V_f \;\middle|\; \sum_{l=0}^{r-1} p_f^{(l)}(z,v) > 0 \right\}$. We split the transition (u,v) in $c_f(u,v)$ unit transitions using the new states $z_0(u,v) = u;\ z_1(u,v);\ z_2(u,v);\ \ldots;\ z_{c_f(u,v)-1}(u,v);\ z_{c_f(u,v)}(u,v) = v$. If $c_f(u,v) > 1$, then the stationary strategies $p_f^{(l)}$, $l = \overline{0, r-1}$, are updated in the following way:

1. new rows and columns, filled with zeros, that correspond to intermediate states $z_j(u,v)$, $j = \overline{1, c_f(u,v)-1}$, are added;
2. the element $p_f^{(l)}(z_j(u,v), z_{j+1}(u,v))$ is set to 1, $j = \overline{1, c_f(u,v)-1}$;
3. the element $p_f^{(l)}(z, z_{c_f(u,v)-c_f(z,v)+1}(u,v))$ is set to $p_f^{(l)}(z,v)$ and, after that, the element $p_f^{(l)}(z,v)$ is set to 0, for each $z \in V_f^-(v)$ that satisfies the inequality $c_f(z,v) > 1$.

We consider $p_f^*(z_j(u,v)) = 0$, $j = \overline{1, c_f(u,v)-1}$, because the game cannot be started from one intermediate state. The final sequence of states becomes

$$\begin{aligned} X_{ext} \;&= (x_1,\ z_1(x_1,x_2),\ z_2(x_1,x_2), \ldots, z_{c_f(x_1,x_2)-1}(x_1,x_2), x_2, \\ &\quad z_1(x_2,x_3), z_2(x_2,x_3), \ldots, z_{c_f(x_2,x_3)-1}(x_2,x_3), x_3, \\ &\quad \ldots, \\ &\quad z_1(x_{m-1},x_m), z_2(x_{m-1},x_m), \ldots, z_{c_f(x_{m-1},x_m)-1}(x_{m-1},x_m), x_m). \end{aligned}$$

So, if we put all these things together, we obtain that the new game Γ_{ext} is defined on stochastic system L_{ext}, with a finite set of states V_{ext}, final sequence of states X_{ext}, an initial distribution of the states p_{ext}^* and stationary strategy $p_{ext}^{(l)}$ of each player Π_l, $l = \overline{0, r-1}$, defined in the way described above. Also, each transition cost in this stochastic system is equal to 1. This implies that the outcome C_{ext} of the game Γ_{ext} is

$$C_{ext} = \sum_{(u,v) \in E_{ext}} c_{ext}(u,v) = \sum_{(u,v) \in E_{ext}} 1 = |E_{ext}| = \sum_{(u,v) \in E_f} c_f(u,v) = C_f, \tag{13}$$

that implies $C_f = C_{ext}$.

3.2 Expectation of game outcome. Next, we will summarize the results obtained above. We have the following formula for calculating the outcome C_{P} of the game Γ_{P}:

$$\begin{aligned} C_{\mathrm{P}} \;&= C_{\mathrm{P}^+} + (c_{\mathrm{P}}^{\min} - 1) \cdot C_1 \approx C_{\Theta^+} + (c_{\mathrm{P}}^{\min} - 1) \cdot C_1 = \frac{1}{\mu} \cdot C_{\mathrm{N}^*} + (c_{\mathrm{P}}^{\min} - 1) \cdot C_1 = \\ &= \frac{C_f - C_1}{r\mu} + (c_{\mathrm{P}}^{\min} - 1) \cdot C_1 = \frac{C_{ext} - C_1}{r\mu} + (c_{\mathrm{P}}^{\min} - 1) \cdot C_1 = \frac{1}{r\mu} \cdot C_{ext} + \left(c_{\mathrm{P}}^{\min} - 1 - \frac{1}{r\mu} \right) \cdot C_1 . \end{aligned}$$

So, the expectation of the outcome C_{P} of the game Γ_{P} is

$$\mathrm{E}(C_{\mathrm{P}}) \approx \frac{1}{r\mu} \cdot \mathrm{E}(C_{ext}) + \left(c_{\mathrm{P}}^{\min} - 1 - \frac{1}{r\mu} \right) \cdot \mathrm{E}(C_1). \quad (14)$$

The values $\mathrm{E}(C_{ext})$ and $\mathrm{E}(C_1)$ can be determined using the algorithm developed in [2], because the both games Γ_{ext} and Γ_1 are defined on stochastic systems with unit transition costs (which, in particular case, can be considered as unit transition time).

3.3 Win probabilities of the players. From the definition of the win probabilities of the players, it is easy to observe that they depend only on the network structure and the transition probabilities of the stochastic system on which the game is defined. Indeed, if we consider that each transition time in the game Γ_{P} is equal with 1, and T_{P} is the duration of the game Γ_{P}, then the win probability $w_{\mathrm{P}}^{(l)}$ of the player Π_l, $l = \overline{0, r-1}$, is given by the following relation:

$$w_{\mathrm{P}}^{(l)} = P(T_{\mathrm{P}} \equiv l \ (\mathrm{mod}\, r)) = \sum_{k=0}^{\infty} P(T_{\mathrm{P}} = k + l), \ l = \overline{0, r-1}. \quad (15)$$

So, the win probabilities of the players do not depend on the transition costs. The win probabilities of the players in the game Γ_{P} are the same as the win probabilities of the players in the game Γ_1, that can be found using the algorithm from [2].

CONCLUSIONS

The purpose of this paper was to generalize the problem of determining the expectation of the game outcome for the case when the transition costs are real. The elaborated approach is based on extended network method, that allows us to reduce the real transition costs case to unit transition time case.

This reduced case was investigated in [2] and algorithms for determining the expectation of the game duration (outcome) and win probabilities of the players were developed. So, solving the problem for the unit transition time case and performing the transformations presented in this paper, we obtain the solution of the problem in the general case, when the transition costs are real.

REFERENCES

[1] A. Lazari, D. Lozovanu and M. Capcelea, *Dynamical deterministic and stochastic systems: evolution, optimization and discrete optimal control (in romanian),* Chisinau, CEP USM, 2015;

[2] A. Lazari, Stochastic games on Markov processes with final sequence of states, *Bul. Acad. Ştiinţe Repub. Mold. Mat.*, no. 1 (83), pp. 77-94, 2017;

[3] A. Lazari, Determining the distribution of the duration of stationary games for zero-order Markov processes with final sequence of states, *Bul. Acad. Ştiinţe Repub. Mold. Mat.*, no. 3 (79), pp. 72-78, 2015;

[4] A. Lazari, Rationalization of discrete Markov processes with final sequence of states and real transition costs, *Proceedings of the international conference "Mathematical modelling, optimization and informational technologies"*, Chisinau, Evrica, pp. 120-126, 2014;

[5] A. Lazari, Extended network method for determining the main probabilistic characteristics of the evolution for stochastic discrete systems with positive rational transitions, *Scientific abstracts of International Conference of Young Researchers, VIII edition*, Chisinau, pp. 93, 2010;

[6] G. Zbaganu, Waiting for an ape to type a poem, *Bul. Acad. Ştiinţe Repub. Mold. Mat.*, no. 2 (8), pp. 66-74, 1992;

[7] L. Guibas and A. Odlyzko, Periods in strings, *J. Combin. Theory Ser. A*, no. 30, pp. 19-42, 1981.