Contributions to Game Theory and Management, XI, 103–112

Nash Equilibria in Mixed Stationary Strategies for m-Player Mean Payoff Games on Networks

Dmitrii Lozovanu[1] **and Stefan Pickl**[2]

[1] *Institute of Mathematics and Computer Science of Moldova Academy of Sciences, Academiei 5, Chisinau, MD-2028, Moldova,*
E-mail: lozovanu@math.md

[2] *Institute for Theoretical Computer Science, Mathematics and Operations Research, Universität der Bundeswehr München, 85577 Neubiberg-München, Germany,*
E-mail: stefan.pickl@unibw.de

Abstract We consider a class of non-zero-sum mean payoff games on networks that extends the two-player zero-sum mean payoff game introduced by Ehrenfeucht and Mycielski. We show that for the considered class of games there exist Nash equilibria in mixed stationary strategies and propose an approach for determining the optimal strategies of the players.

Keywords: mean payoff game, pure stationary strategy, mixed stationary strategy, Nash equilibria

1. Introduction

In this paper we consider a class of m-player mean payoff games on networks that generalizes the following two-player zero-sum mean payoff game introduced by Ehrenfeucht and Mycielski, 1979, and considered by Gurvich et al., 1988.

Let $G = (X, E)$ be a finite directed graph in which every vertex $x \in X$ has at least one outgoing directed edge $e = (x, y) \in E$. On the edge set E it is given a function $c : E \to R$ which assigns a cost $c(e)$ to each edge $e \in E$. In addition the vertex set X is divided into two disjoint subsets X_1 and X_2 $(X = X_1 \cup X_2,\ X_1 \cap X_2 = \emptyset)$ which are regarded as position sets of the two players. The game starts at a given position $x_0 \in X$. If $x_0 \in X_1$ then the move is done by the first player, otherwise it is done dy second one. Move means the passage from position x_0 to a neighbor position x_1 through the directed edge $e_0 = (x_0, x_1) \in E$. After that if $x_1 \in X_1$ then the move is done by the first player, otherwise it is done by the second one and so on indefinitely. The first player has the aim to maximize $\displaystyle\lim_{t\to\infty} \inf \frac{1}{t}\sum_{\tau=0}^{t-1} c(e_\tau)$ while the second player has the aim to minimize $\displaystyle\lim_{t\to\infty} \sup \frac{1}{t}\sum_{\tau=0}^{t-1} c(e_\tau)$. Ehrenfeucht and Mycielski, 1979, proved that for this game there exists a value $v(x_0)$ such that the first player has a strategy of moves that insures $\displaystyle\lim_{t\to\infty} \inf \frac{1}{t}\sum_{\tau=0}^{t-1} c(e_\tau) \geq v(x_0)$ and the second player has a strategy of moves that insure $\displaystyle\lim_{t\to\infty} \sup \frac{1}{t}\sum_{\tau=0}^{t-1} c(e_\tau) \leq v(x_0)$. Furthermore Ehrenfeucht and Mycielski, 1979 and Gurvich et al., 1988 showed that the players can achieve the value $v(x_0)$ applying the strategies of moves which do not depend on t but depend only on the vertex from which the player is able to move.

Ehrenfeucht and Mycielski, 1979 and Zwick and Paterson,1996, called such strategies *positional strategies*. Gurvich et al., 1988 and Lozovanu and Pickl, 2006, called these strategies stationary strategies.

In this paper we will regard such strategies as *pure stationary strategies* because each move in a position of the game is chosen from the set of feasible strategies of moves by the corresponding player with the probability equal to 1 and in each position such a strategy does not change in time.

A generalization of the zero-sum mean payoff game to m-player, where $m \geq 2$, is the following. Consider a finite directed graph $G = (X, E)$ in which every vertex has at least one outgoing directed edge. Assume that the vertex set X is divided into m disjoint subsets $X_1, X_2, \dots, X_m$ ($X = X_1 \cup X_2 \cup \dots \cup X_m$; $X_i \cap Xj = \emptyset, i \neq j$) which we regard as position sets of the m players. Additionally, we assume that on the edge set m functions $c^i : F \to R,\ i = 1, 2, \dots, m$ are defined that assign to each directed edge $e = (x, y) \in E$ the values $c_e^1, c_e^2, \dots, c_e^m$ that are regarded as the rewards for the corresponding players $1, 2, \dots, m$.

On G we consider the following m-person dynamic game. The game starts at given position $x_0 \in X$ at the moment of time $t = 0$ where the player $i \in \{1, 2, \dots, m\}$ who is owner of the starting position x_0 makes a move from x_0 to a neighbor position $x_1 \in X$ through the directed edge $e_0 = (x_0, x_1) \in E$. After that players $1, 2, \dots, m$ receive the corresponding rewards $c_{e_0}^1, c_{e_0}^2, \dots, c_{e_0}^m$. Then at the moment of time $t = 1$ the player $k \in \{1, 2, \dots, m\}$ who is owner of position x_1 makes a move from x_1 to a position $x_2 \in V$ through the directed edge $e_1 = (x_1, x_2) \in E$, players $1, 2, \dots, m$ receive the corresponding rewards $c_{e_1}^1, c_{e_1}^2, \dots, c_{e_1}^m$, and so on, indefinitely. Such a play of the game on G produces the sequence of positions $x_0, x_1, x_2, \dots, x_t \dots$ where each x_t is the position at the moment of time t. In this game the players make moves through the directed edges in their positions in order to maximize their average rewards per move

$$\omega_{x_o}^i = \lim_{t \to \infty} \inf \frac{1}{t} \sum_{\tau=0}^{t-1} c_{e_\tau}^i, \quad i = 1, 2, \dots, m.$$

The game formulated above in the case $m = 2$ and $c_e^1 = -c_e^2 = c_e,\ \forall e \in E$ is transformed into a two-player mean payoff game for which Nash equilibria in pure stationary strategies exist. In general a non-zero-sum mean payoff game may have no Nash equilibrium in pure stationary strategies. This fact has been shown by Gurvich et al., 1988, where an example of two-player non-zero-sum mean payoff game that has no Nash equilibrium in pure strategies is constructed. Alpern, 1991 and Lozovanu and Pickl, 2015, have shown that Nash equilibria for non-zero m-player mean payoff games may exist only for some special cases. A class of m-player mean payoff games for which Nash equilibria in pure stationary strategies exist is presented in Lozovanu and Pickl, 2006.

In this paper we consider the non-zero-sum mean payoff games in mixed stationary strategies. We define a mixed stationary strategy of moves in a position $x \in X_i$ for the player $i \in \{1, 2, \dots, m\}$ as a probability distribution over the set of feasible moves from x. We show that an arbitrary m-player mean payoff game possesses a Nash equilibrium in mixed stationary strategies. Based on a constructive proof of this result we propose an approach for determining the optimal mixed stationary strategies of the players.

The paper is organized as follows. In Section 2 a class of average stochastic positional games that generalizes non-zero-sum mean payoff games is considered. Then in Sections 3 is shown how an average stochastic positional game can be formulated in terms of pure and mixed stationary strategies. In Section 4 some results concerned with the existence of Nash equilibria in mixed stationary strategies for average stochastic positional games are presented. Additionally an approach for determining the optimal strategies of players is proposed. In Sections 5,6, based on results from the Sections 3,4, is proven the existence of Nash equilibria in mixed stationary strategies for non-zero-sum mean payoff games and an approach for determining the optimal strategies of the players is proposed.

2. A generalization of mean payoff games to average stochastic positional games

The problem of determining Nash equilibria in mixed stationary strategies for mean payoff games leads to a special class of average stochastic games that Lozovanu and Pickl, 2015 called average stochastic positional games. Lozovanu, 2018 shown that this class of games possesses of Nash equilibria in mixed stationary strategies. Therefore in the paper we shall use the average stochastic positional games for studying the non-zero-sum mean payoff games.

A stochastic positional game with m players consists of the following elements:

- a state space X (which we assume to be finite);
- a partition $X = X_1 \cup X_2 \cup \cdots \cup X_m$ where X_i represents the position set of player $i \in \{1, 2, \ldots, m\}$;
- a finite set $A(x)$ of actions in each state $x \in X$;
- a step reward $f^i(x, a)$ with respect to each player $i \in \{1, 2, \ldots, m\}$ in each state $x \in X$ and for an arbitrary action $a \in A(x)$;
- a transition probability function $p : X \times \prod_{x \in X} A(x) \times X \rightarrow [0, 1]$ that gives the probability transitions $p^a_{x,y}$ from an arbitrary $x \in X$ to an arbitrary $y \in X$ for a fixed action $a \in A(x)$, where $\sum_{y \in X} p^a_{x,y} = 1, \ \ \forall x \in X, \ a \in A(x)$;
- a starting state $x_0 \in X$.

The game starts at the moment of time $t = 0$ in the state x_0 where the player $i \in \{1, 2, \ldots, m\}$ who is the owner of the state position x_0 ($x_0 \in X_i$) chooses an action $a_0 \in A(x_0)$ and determines the rewards $f^1(x_0, a_0), f^2(x_0, a_0), \ldots, f^m(x_0, a_0)$ for the corresponding players $1, 2, \ldots, m$. After that the game passes to a state $y = x_1 \in X$ according to a probability distribution $\{p^{a_0}_{x_0,y}\}$. At the moment of time $t = 1$ the player $k \in \{1, 2, \ldots, m\}$ who is the owner of the state position x_1 ($x_1 \in X_k$) chooses an action $a_1 \in A(x_1)$ and players $1, 2, \ldots, m$ receive the corresponding rewards $f^1(x_1, a_1), f^2(x_1, a_1), \ldots, f^m(x_1, a_1)$. Then the game passes to a state $y = x_2 \in X$ according to a probability distribution $\{p^{a_1}_{x_1,y}\}$ and so on indefinitely. Such a play of the game produces a sequence of states and actions $x_0, a_0, x_1, a_1, \ldots, x_t, a_t, \ldots$ that defines a stream of stage rewards $f^1(x_t, a_t), f^2(x_t, a_t), \ldots, f^m(x_t, a_t), \quad t = 0, 1, 2, \ldots$. *The average stochastic positional game* is the game with payoffs of the players

$$\omega^i_{x_0} = \lim_{t\to\infty}\inf\ \frac{1}{t}\sum_{\tau=0}^{t-1}\mathsf{E}(f^i(x_\tau, a_\tau)), \quad i = 1, 2, \ldots, m$$

where E is the expectation operator with respect to the probability measure in the Markov process induced by actions chosen by players in their position sets and given a starting state x_0. Each player in this game has the aim to maximize his average reward per transition. In the case $m = 1$ this game becomes the average Markov decision problem with given action sets $A(x)$ for $x \in X$, a transition probability function $p : X \times \prod_{x\in X} A(x) \times X \rightarrow [0, 1]$ and step rewards $f(x, a) = f^1(x, a)$ for $x \in X$ and $a \in A(x)$. If in an average stochastic positional game the probabilities $p^a_{x,y}$ take only values 0 and 1, i. e. $p^a_{x,y} \in \{0, 1\}$, $\forall x, y \in X$ and $\forall a \in A(x)$, then such a game becomes a mean payoff game on the graph $G = (X, E)$, where $e = (x, y) \in E$ if and only if there exists $a \in A(x)$ such that $p^a_{x,y} = 1$. So, in this case the set of directed edges $E(x) = \{e = (x, y) \in E | y \in X\}$ with the common origin in x corresponds to the set of actions $A(x)$ in the position x of the game.

In the paper we will study the average stochastic positional game when the players use pure and mixed stationary strategies of choosing the actions in the states.

3. Average stochastic positional games in pure and mixed stationary strategies

A strategy of player $i \in \{1, 2, \ldots, m\}$ in a stochastic positional game is a mapping s^i that provides for every state $x_t \in X_i$ a probability distribution over the set of actions $A(x_t)$. If these probabilities take only values 0 and 1, then s^i is called *a pure strategy*, otherwise s^i is called *a mixed strategy*. If these probabilities depend only on the state $x_t = x \in X_i$ (i. e. s^i does not depend on t), then s^i is called *a stationary strategy*, otherwise s^i is called a *non-stationary strategy.*

Thus, we can identify the set of mixed stationary strategies $\mathbf{S}^i$ of player i with the set of solutions of the system

$$\begin{cases} \sum\limits_{a\in A(x)} s^i_{x,a} = 1, & \forall x \in X_i; \\ s^i_{x,a} \geq 0, & \forall x \in X_i, \quad \forall a \in A(x). \end{cases} \tag{1}$$

Each basic solution s^i of this system corresponds to a pure stationary strategy of player $i \in \{1, 2, \ldots, m\}$. So, the set of pure stationary strategies S^i of player i corresponds to the set of basic solutions of system (1).

Let $\mathbf{s}= (s^1, s^2, \ldots, s^m) \in \mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \cdots \times \mathbf{S}^m$ be a profile of stationary strategies (pure or mixed strategies) of the players. Then the elements of probability transition matrix $P^{\mathbf{s}} = (p^{\mathbf{s}}_{x,y})$ in the Markov process induced by $\mathbf{s}$ can be calculated as follows:

$$p^{\mathbf{s}}_{x,y} = \sum_{a\in A(x)} s^i_{x,a} p^a_{x,y} \quad \text{for} \quad x \in X_i, \quad i = 1, 2, \ldots, m. \tag{2}$$

If we denote by $Q^{\mathbf{s}} = (q^{\mathbf{s}}_{x,y})$ the limiting probability matrix of matrix $P^{\mathbf{s}}$ then the average payoffs per transition $\omega^1_{x_0}(\mathbf{s}),\ \omega^2_{x_0}(\mathbf{s}), \ldots, \omega^m_{x_0}(\mathbf{s})$ for the players induced by profile $\mathbf{s}$ are determined as follows

$$\omega_{x_0}^i(\mathbf{s}) = \sum_{k=1}^{m} \sum_{y \in X_k} q_{x_0,y}^{\mathbf{s}} f^i(y, s^k), \quad i = 1, 2, \ldots, m, \tag{3}$$

where

$$f^i(y, s^k) = \sum_{a \in A(y)} s_{y,a}^k f^i(y, a), \ \text{ for } \ y \in X_k, \ k \in \{1, 2, \ldots, m\} \tag{4}$$

expresses the average reward (step reward) of player i in the state $y \in X_k$ when player k uses the strategy s^k.

The functions $\omega_{x_0}^1(\mathbf{s}),\ \omega_{x_0}^2(\mathbf{s}),\ \ldots,\ \omega_{x_0}^m(\mathbf{s})$ on $\mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \cdots \times \mathbf{S}^m$, defined according to (10),(11), determine a game in normal form that we denote by $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\omega_{x_0}^i(\mathbf{s})\}_{i=\overline{1,m}} \rangle$. This game corresponds to the *average stochastic positional game in mixed stationary strategies* that in extended form is determined by the tuple $(\{X_i\}_{i=\overline{1,m}},\ \{A(x)\}_{x\in X},\ \{f^i(x,a)\}_{i=\overline{1,m}},\ p,\ x_0)$. The functions $\omega_{x_0}^1(\mathbf{s}),\ \omega_{x_0}^2(\mathbf{s}), \ldots, \omega_{x_0}^m(\mathbf{s})$ on $S = S^1 \times S^2 \times \cdots \times S^m$, determine the game $\langle \{S^i\}_{i=\overline{1,m}}, \{\omega_{x_0}^i(\mathbf{s})\}_{i=\overline{1,m}} \rangle$ that corresponds to the *average stochastic positional game in pure strategies*. In the extended form this game is also determined by the tuple $(\{X_i\}_{i=\overline{1,m}},\ \{A(x)\}_{x\in X},\ \{f^i(x,a)\}_{i=\overline{1,m}},\ p,\ x_0)$.

A stochastic positional games can be considered also for the case when the starting state is chosen randomly according to a given distribution $\{\theta_x\}$ on X. So, for a given stochastic positional game we may assume that the play starts in the state $x \in X$ with probability $\theta_x > 0$ where $\sum_{x \in X} \theta_x = 1$. If the players use mixed stationary strategies then the payoff functions

$$\psi_\theta^i(\mathbf{s}) = \sum_{x \in X} \theta_x \omega_x^i(\mathbf{s}), \quad i = 1, 2, \ldots, m$$

on $\mathbf{S}$ define a game in normal form $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\psi_\theta^i(\mathbf{s})\}_{i=\overline{1,m}} \rangle$ that in extended form is determined by $(\{X_i\}_{i=\overline{1,m}}, \{A(x)\}_{x\in X}, \{f^i(x,a)\}_{i=\overline{1,m}}, p, \{\theta_x\}_{x\in X})$. In the case $\theta_x = 0, \forall x \in X \setminus \{x_0\},\ \ \theta_{x_o} = 1$ the considered game becomes a stochastic positional game with a fixed starting state x_0.

4. Nash equilibria for an average stochastic positional game and determining the optimal stationary strategies of the players

We present a Nash equilibria existence result and an approach for determining the optimal mixed stationary strategies of the players for the average stochastic positional game when the starting state of the game is chosen randomly according to a given distribution $\{\theta_x\}$ on the set of states X. In this case for the game in normal form $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\psi_\theta^i(\mathbf{s})\}_{i=\overline{1,m}} \rangle$, the set of strategies $\mathbf{S}^i$ and the payoff functions $\psi_\theta^i(\mathbf{s}),\ \ i = 1, 2, \ldots, m$ can be specified as follows:

Let $\mathbf{S}^i,\ i \in \{1, 2, \ldots m\}$ be the set of solutions of the system

$$\begin{cases} \sum_{a \in A(x)} s_{x,a}^i = 1, & \forall x \in X_i; \\ s_{x,a}^i \geq 0, & \forall x \in X_i, \ \ \forall a \in A(x). \end{cases} \tag{5}$$

On $\mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \cdots \times \mathbf{S}^m$ we define m payoff functions

$$\psi^i_\theta(s^1, s^2, \dots, s^m) = \sum_{k=1}^{m} \sum_{x \in X_k} \sum_{a \in A(x)} s^k_{x,a} f^i(x,a) q_x, \qquad i = 1, 2, \dots, m, \tag{6}$$

where q_x for $x \in X$ are determined uniquely from the following system of linear equations

$$\begin{cases} q_y - \sum\limits_{k=1}^{m} \sum\limits_{x \in X_k} \sum\limits_{a \in A(x)} s^k_{x,a}\ p^a_{x,y}\ q_x = 0, & \forall y \in X; \\ q_y + w_y - \sum\limits_{k=1}^{m} \sum\limits_{x \in X_k} \sum\limits_{a \in A(x)} s^k_{x,a}\ p^a_{x,y}\ w_x = \theta_y, & \forall y \in X \end{cases} \tag{7}$$

for an arbitrary fixed profile $\mathbf{s} = (s^1, s^2, \dots, s^m) \in \mathbf{S}$.

The functions $\psi^i_\theta(s^1, s^2, \dots, s^m)$, $i = 1, 2, \dots, m$, represent the payoff functions for the average stochastic game in normal form $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\psi^i_\theta(\mathbf{s})\}_{i=\overline{1,m}} \rangle$ determined by the tuple $(\{X_i\}_{i=\overline{1,m}},\ \{A(x)\}_{x \in X},\ \{f^i(x,a)\}_{i=\overline{1,m}},\ p,\ \{\theta_y\}_{y \in X})$ where θ_y for $y \in X$ are given nonnegative values such that $\sum_{y \in X} \theta_y = 1$.

If $\theta_y = 0,\ \forall y \in X \setminus \{x_0\}$ and $\theta_{x_0} = 1$, then we obtain an average stochastic game in normal form $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\omega^i_{x_0}(\mathbf{s})\}_{i=\overline{1,m}} \rangle$ when the starting state x_0 is fixed, i.e. $\psi^i_\theta(s^1, s^2, \dots,\ s^m) = \omega^i_{x_0}(s^1, s^2,\ \dots,\ s^m)$, $i = 1, 2, \dots, m$. So, in this case the game is determined by $(\{X_i\}_{i=\overline{1,m}},\ \{A(x)\}_{x \in X},\ \{f^i(x,a)\}_{i=\overline{1,m}},\ p,\ x_0)$.

Lozovanu, 2018, showed that each payoff function $\psi^i_\theta(s)$, $i \in \{1, 2, \dots, m\}$ in the game $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\psi^i_\theta(\mathbf{s})\}_{i=\overline{1,m}} \rangle$ is quasi-monotonic (quasi-convex and quasi-concave) with respect to $\mathbf{s}^i$ on a convex and compact set $\mathbf{S}^i$ for fixed $\mathbf{s}^1,\ \mathbf{s}^2,\ \dots,\ \mathbf{s}^{i-1},\ \mathbf{s}^{i+1},\ \dots, \mathbf{s}^m$. Moreover for the game $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\psi^i_\theta(\mathbf{s})\}_{i=\overline{1,m}} \rangle$ has been shown that each payoff function $\psi^i_\theta(s)$, $i \in \{1, 2, \dots, m\}$ is graph-continuous in the sense of Dasgupta and Maskin, 1986. Based on these properties Lozovanu, 2018, proved the following theorem.

Theorem 1. *The game $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}}, \{\psi^i_\theta(s)\}_{i=\overline{1,m}} \rangle$ possesses a Nash equilibrium $\mathbf{s}^* = (s^{1^*}, s^{2^*}, \dots, s^{m*}) \in \mathbf{S}$ which is a Nash equilibrium in mixed stationary strategies for the average stochastic positional game determined by $(\{X_i\}_{i=\overline{1,m}},\ \{A(x)\}_{x \in X},\ \{f^i(x,a)\}_{i=\overline{1,m}},\ p,\ \{\theta_y\}_{y \in X})$. If $\theta_y > 0,\ \forall y \in X$, then $\mathbf{s}^* = (s^{1^*}, s^{2^*}, \dots, s^{m*})$ is a Nash equilibrium in mixed stationary strategies for the average stochastic positional game $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}}, \{\omega^i_y(\mathbf{s})\}_{i=\overline{1,m}} \rangle$ with an arbitrary starting state $y \in X$.*

Thus, for an average stochastic positional game a Nash equilibrium in mixed stationary strategies can be found using the game model $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\psi^i_\theta(\mathbf{s})\}_{i=\overline{1,m}} \rangle$, where $\mathbf{S}^i$ and $\psi^i_\theta(\mathbf{s})$, $i = 1, 2, \dots, m$, are determined according to (5)-(7). This game model in the case $m = 2$, $f(x,a) = f^1(x,a) = -f^2(x,a)$, $\forall x \in X$, $\forall a \in A(x)$ corresponds to a zero-sum two-player average stochastic positional game and we can use it for determining the optimal stationary strategies of the players. Note that in this case the equilibrium may exists in pure stationary strategies and consequentely such a game model allows to determine the optimal pure stationary strategies of the players.

For antagonistic average stochastic positional games Lozovanu and Pickl, 2016, proposed an another approach for determining Nash equilibria in pure stationary strategies. However the approach from Lozovanu and Pickl, 2016, couldn't be extended for non-zero average stochastic positional games and for the non-zero mean payoff games. Nevertheless such an approach allows to ground finite efficient iterative procedures for determining the optimal pure stationary strategies of the players.

5. Formulation of mean payoff games in mixed stationary strategies

Consider a mean payoff game determined by the tuple $(G, \{X_i\}_{i=\overline{1,m}}, \{c^i\}_{i=\overline{1,m}}, x_0)$ where $G = (X, E)$ is a finite directed graph with vertex set X and edge set E, $X = X_1 \cup X_2 \cup \cdots \cup X_m$ $(X_i \cap Xj = \emptyset, i \neq j)$ is a partition of X that determine the corresponding position set of players, $c^i : E \to R^1$, $i = 1, 2, \ldots, m$ are the real functions that determine the rewards on edges of graph G and x_0 is the starting position of the game.

The pure and mixed stationary strategies in the mean payoff game on G can be defined in a similar way as for the average stochastic positional game. We identify the set of mixed stationary strategies S^i of player $i \in \{1, 2, \ldots, m\}$ with the set of solutions of the system

$$\begin{cases} \sum\limits_{y \in X(x)} s^i_{x,y} = 1, \quad \forall x \in X_i; \\ \qquad\quad s^i_{x,y} \geq 0, \quad \forall x \in X_i,\ y \in X(x) \end{cases} \tag{8}$$

where $X(x)$ represents the set of neighbor vertices for the vertex x, i.e. $X(x) = \{y \in X | e = (x, y) \in E\}$.

Let $\mathbf{s} = (s^1, s^2, \ldots, s^m)$ be a profile of stationary strategies (pure or mixed strategies) of the players. This means that the moves in the mean payoff game from an arbitrary $x \in X$ to $y \in X$ induced by $\mathbf{s}$ are made according to probabilities of the stochastic matrix $P^s = (s_{x,y})$, where

$$\mathbf{s}_{x,y} = \begin{cases} s^i_{x,y} & \text{if } \ e = (x, y) \in E, x \in X_i, y \in X;\ i = 1, 2, \ldots, m; \\ 0 & \text{if } \ e = (x, y) \notin E. \end{cases} \tag{9}$$

Thus, for a given profile $\mathbf{s}$ we obtain a Markov process with the probability transition matrix $P^{\mathbf{s}} = (\mathbf{s}_{x,y})$ and the corresponding rewards $c^i_{x,y}$, $i = 1, 2, \ldots, m$ on edges $(x, y) \in E$. Therefore, if $Q^{\mathbf{s}} = (q^{\mathbf{s}}_{x,y})$ is the limiting probability matrix of $P^{\mathbf{s}}$ then the average rewards per transition $\omega^1_{v_0}(\mathbf{s})$, $\omega^2_{x_0}(\mathbf{s})$, $\ldots, \omega^m_{x_0}(\mathbf{s})$ for the players can be determined as follows

$$\omega^i_{x_0}(\mathbf{s}) = \sum_{k=1}^{m} \sum_{y \in X_k} q^{\mathbf{s}}_{x_0,y} \mu^i(y, s^k), \quad i = 1, 2, \ldots, m, \tag{10}$$

where

$$\mu^i(y, s^k) = \sum_{x \in X(y)} s^k_{y,x} c^i(y, x), \quad \text{for } \ y \in X_k,\ k \in \{1, 2, \ldots, m\} \tag{11}$$

expresses the average step reward of player i in the state $y \in X_k$ when player k uses the mixed stationary strategy s^k. The functions $\omega^1_{x_0}(\mathbf{s}),\ \omega^2_{x_0}(\mathbf{s}),\ \dots,\ \omega^m_{x_0}(\mathbf{s})$ on $\mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \dots \times \mathbf{S}^m$, defined according to (10), (11), determine a game in normal form that we denote by $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\omega^i_{x_0}(\mathbf{s})\}_{i=\overline{1,m}} \rangle$. This game corresponds to the *mean payoff game in mixed stationary strategies* on G with a fixed starting position x_0. So this game is determined by the tuple $(G, \{X_i\}_{i=\overline{1,m}}, \{c^i\}_{i=\overline{1,m}}, x_0)$.

In a similar way as for an average stochastic game here we can consider the mean payoff game on G when the starting state is chosen randomly according to a given distribution $\{\theta_x\}$ on X. So, for such a game we will assume that the play starts in the states $x \in X$ with probabilities $\theta_x > 0$ where $\sum\limits_{x\in X} \theta_x = 1$. If the players in such a game use mixed stationary strategies of moves in their positions then the payoff functions

$$\psi^i_\theta(\mathbf{s}) = \sum_{x\in X} \theta_x \omega^i_x(\mathbf{s}), \quad i = 1, 2, \dots, m$$

on $\mathbf{S}$ define a game in normal form $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\psi^i_\theta(\mathbf{s})\}_{i=\overline{1,m}} \rangle$ that is determined by the following tuple $(G,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ \{\theta_x\}_{x\in X})$. In the case $\theta_x = 0$, $\forall x \in X \setminus \{v_0\}$, $\theta_{v_0} = 1$ this game becomes a mean payoff with fixed starting state x_0.

6. Nash equilibria in mixed stationary strategies for mean payoff games and determining the optimal strategies of the players

In this section we show how the results from the previous sections can be applied for determining Nash equilibria and the optimal mixed stationary strategies of the players for mean payoff games.

Let $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\psi^i_\theta(s)\}_{i=\overline{1,m}} \rangle$ be the game in normal form for the mean payoff game determined by $(G,\ \{X_i\}_{i=\overline{1,m}},\ \{c^i\}_{i=\overline{1,m}},\ \{\theta_x\}_{x\in X})$. We show that $\mathbf{S}^i$ and $\psi^i_\theta(s)$ for $i \in \{1, 2, \dots, m\}$ can be defined as follows:

$\mathbf{S}^i$ represents a set of the solutions of the system

$$\begin{cases} \sum\limits_{y\in X(x)} s^i_{x,y} = 1, \quad \forall x \in X_i; \\ \qquad\quad s^i_{x,y} \geq 0, \quad \forall x \in X_i,\ y \in X(x) \end{cases} \tag{12}$$

and

$$\psi^i_\theta(s^1, s^2, \dots, s^m) = \sum_{k=1}^{m} \sum_{y\in X_k} \sum_{y\in X(x)} s^k_{x,y} c^i(x,y) q_x, \tag{13}$$

where q_x for $x \in X$ are determined uniquely (via $s^k_{x,y}$) from the following system of equations

$$\begin{cases} q_y - \sum\limits_{k=1}^{m} \sum\limits_{x\in X_k} \sum\limits_{v\in X(x)} s^k_{x,y}\ q_x = 0, \qquad\qquad \forall y \in X(x); \\ \\ q_y + w_y - \sum\limits_{k=1}^{m} \sum\limits_{y\in X} \sum\limits_{x\in X(y)} s^k_{x,y}\ w_x = \theta_y, \quad \forall y \in X(x). \end{cases} \tag{14}$$

Here, θ_y for $y \in X$ represent arbitrary fixed positive values such that $\sum\limits_{y \in X} \theta_y = 1$.

Using Theorem 1 we can prove now the following result.

Theorem 2. *For a mean payoff game on G the corresponding game in normal form $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}}, \{\psi^i_\theta(s)\}_{i=\overline{1,m}} \rangle$ possesses a Nash equilibrium $\mathbf{s}^* = (s^{1^*}, s^{2^*}, \dots, s^{m^*}) \in \mathbf{S}$ which is a Nash equilibrium in mixed stationary strategies for the mean payoff game on G with an arbitrary starting position $x_0 \in X$.*

Proof. To prove the theorem it is sufficient to show that the functions $\psi^i_\theta(s)$, $i \in \{1, 2, \dots, m\}$ defined according to (13), (14) represent the payoff functions for the mean payoff game determined by $(G, \{X_i\}_{i=\overline{1,m}}, \{c^i\}_{i=\overline{1,m}}, \{\theta_x\}_{x \in X})$. This is easy to verify because if we replace in (6) the rewards $f^i(x,a)$ for $x \in X$ and $a \in A(x)$ by rewards $c^i_{x,y}$ for $(x,y) \in E$ and in (6), (7) we replace the probabilities $p^a_{x,y}$, $x \in X_k$, $a \in A(x)$ for the corresponding players $k = 1, 2, \dots, m$ by $p^k_{x,y} \in \{0,1\}$ according to the structure of graph G then we obtain that (6), (7) are transformed into (13), (14). If after that we apply Theorem 1 then obtain the proof of the theorem.

So, the optimal mixed stationary strategies of the players in a mean payoff game can be found if we determine the optimal stationary strategies of the players for the game $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}}, \{\psi^i_\theta(s)\}_{i=\overline{1,m}} \rangle$ where $\mathbf{S}^i$ and $\psi^i_\theta(s)$ for $i \in \{1, 2, \dots, m\}$ are defined according to (12) - (14). If $m = 2$, $c_{x,y} = c^1_{x,y} = -c^2_{x,y}$, $\forall (x,y) \in E$ then we obtain a game model in normal form for the zero-sum two-player mean payoff on graph G. In this case the equilibrium exists in pure stationary strategies and the considered game model allows to determine the optimal pure stationary strategies of the players. For antagonistic mean payoff games on graphs the approach for antagonistic average stochastic positional games from Lozovanu and Pickl, 2016, can also be adapted if we take into account the mentioned transforms in the proof of Theorem 2, i.e. we should change the rewards $f^i(x,a)$ for $x \in X, a \in A(x)$ by rewards $c^i_{x,y}$ for $(x,y) \in E$ and replace the probabilities $p^a_{x,y}$, $x \in X_k$, $a \in A(x)$, $k = 1, 2, \dots, m$ by probabilities $p^k_{x,y} \in \{0,1\}$ according to the structure of the graph G.

7. Conclusion

The considered class of non-zero mean payoff games generalizes the zero-sum two-player mean payoff games on graphs studied by Ehrenfeucht and Mycielski, 1979, Gurvich et al., 1988, Lozovanu and Pickl, 2015 and Zwick and Paterson,1996. For zero-sum two-player mean payoff games on graphs there exist Nash equilibria in pure stationary strategies. For the case of non-zero-sum mean payoff games on networks Nash equilibria in pure stationary strategies may not exist. The results presented in the paper show that an arbitrary mean payoff game in mixed stationary strategies possesses a Nash equilibrium and such an equilibrium can be found using the game models proposed in Sections 5,6.

References

Alpern, S. (1991). *Cycles in extencsive form perfect information games.* Journal of Mathematical Analysis and Applications, **159**, 1–17.

Dasgupta, P. and Maskin, E. (1986). *The existence of an equilibrium in discontinuous economic games.* The Review of Economic Studies, **53**, 1–26.

Ehrenfeucht, A. and Mycielski, J. (1979). *Positional strategies for mean payoff games.* International Journal of Game Theory, **8**, 109–113.

Gurvich, V., Karzanov, A. and Khachiyan, L. (1988). *Cyclic games and an algorithm to find minimax cicle means in directed graphs.* USSR Computational Mathematics and Mathematical Physics, **28**, 85–91.

Lozovanu, D. (2018). *Stationary Nash equilibria for average stochastic positional games.* In: Petrosyan, L., Mazalov, V., Zenkevich, N. (eds) Static amd Dynamic Game Theory: Foundations and Applications, Springer (to appear).

Lozovanu, D. and Pickl, S. (2006). *Nash equilibria conditions for cyclic games with m players.* Electronic Notes in Discrete Mathematics, **28**, 85–91.

Lozovanu, D. and Pickl, S. (2015). *Optimization of Stochastic Discrete Systems and Control on Complex Networks.* Springer.

Lozovanu, D. and Pickl, S. (2016). *Determining the optimal strategies for zero-sum average stochastic positional games.* Electronic Notes in Discrete Mathematics, **55**, 155–159.

Zwick, U. and Paterson, M. (1996). *The complexity of mean payoffs on graphs.* Theoretical Computer Science, **158**, 343–359.