

Strategic Support of the Shapley Value in Stochastic Games*

Elena M. Parilina

Saint Petersburg State University,

7/9 Universitetskaya nab., St. Petersburg, 199034, Russia.

E-mail: e.parilina@spbu.ru

WWW home page:

<http://www.apmath.spbu.ru/en/staff/parilina/index.html>

Abstract We consider the cooperative behavior in stochastic games. We assume that players cooperate in the game and agree on realizing the Shapley value as an imputation of their total payoff. The problem of subgame (time) consistency of the Shapley value is examined. The imputation distribution procedure is constructed to make the Shapley value subgame consistent. We redefine the payoffs in stochastic game applying the imputation distribution procedure. The problem of strategic support of the Shapley value is examined. We prove that the cooperative strategy profile is the Nash equilibrium in the stochastic game with re-defined payoff functions when some conditions are satisfied. The theoretical results are demonstrated on the example of a data transmission game for a wireless network of a specific topology.

Keywords: cooperative stochastic game, time consistency, subgame consistency, imputation distribution procedure, strategic support

1. Introduction

We consider the class of stochastic games with discounted payoffs when players use stationary strategies. This class of games for two players was introduced by (Shapley, 1953a). Most papers devoted to stochastic games examine the non-cooperative behavior of the players, e. g. see (Herings and Peeters, 2004), (Jáskiewicz and Nowak, 2015), (Rosenberg et al., 2003). The cooperative model of a stochastic game was initially proposed by (Petrosjan, 2006). He investigated the problem of subgame consistency of cooperative solutions in a stochastic game played over a finite tree. The same problem was examined for discounted stochastic games when the set of states is finite and players use stationary strategies in (Petrosjan and Baranova, 2006). In this paper the method of finding a cooperative solution and verifying if the solution satisfies the principle of subgame (time) consistency. The problem of time consistency of the cooperative solution was proposed by Petrosyan in (Petrosyan, 1977). He proposed to modify the payment mechanism along cooperative trajectory of the initial game and introduced the IDP (imputation distribution procedure) to make the cooperative solution time-consistent (Petrosyan and Danilov, 1979). This idea was realised for the class of differential games but it is actual for stochastic games as well.

Two other principles of stable cooperation in dynamic games were formulated in (Petrosyan and Zenkevich, 2009) including the principle of strategic support of a cooperative solution. If the cooperative solution is strategically supported, then

* This work was supported by Saint Petersburg State University (research project 9.38.245.2014) and Russian Foundation for Basic Research (project 16-01-00713).

there exists the Nash equilibrium in trigger strategies with the players' payoffs equal to their payoffs in cooperation. The trigger strategy punishes the deviating player by allowing him to obtain the maxmin value in a subgame starting from the stage following the stage when the deviation has been observed (Petrosyan, 2008). The problem of strategic support is considered in (Parilina and Zaccour, 2015b) where the subgame perfect ε -equilibrium is constructed for the dynamic games played over event trees¹. Another principle of stable cooperation is irrational-behaviour-proof (Yeung, 2006). It allows players to guarantee that their payoffs in cooperation will be not less than the payoffs when the cooperation breaks down at some stage and then players proceed playing the game as singletons. These three principles of stable cooperation were adopted to the discounted stochastic games (Parilina, 2015). Recently, the existence of cooperative solutions (Harsanyi, Shapley, Nash solutions) for discounted stochastic games was proved in (Kohlberg and Neyman, 2015).

In our paper we focus on two problems of cooperation in stochastic games: subgame consistency and strategic support. We prove the theorem which allows to construct the imputation distribution procedure to make the Shapley value subgame consistent. Then we define the behavior strategy profile in trigger strategies to support cooperation in case when some player deviates from the cooperative strategy profile. We need to mention here that initially we find the cooperative solution assuming that players use stationary strategies, but the construction of the trigger strategies requires considering the class of behavior strategies. Behavior strategies allow to observe the player's deviation and switch to a trigger mode of the trigger strategy.

As an example of a stochastic game we examine the problem of data transmission in the simple wireless network. The simple network of data transmission consisting of three nodes is taken as a basis of network topology. Two of the nodes generate data packages in each time slot with the corresponding probabilities. The third node is the destination one. The first two nodes are connected by a channel, the connection is one-way, i.e. the first node (first player) can transmit a package directly to node 3 or to node 2. For the transmission of a package to node 2 node 1 receives a nonnegative reward. The system of rewards and costs makes it possible to support cooperation between nodes 1 and 2 which are players 1 and 2 in the game, respectively. The described situation can be solved as a cooperative stochastic game.

Modeling data transmission as a stochastic game was introduced in (Altman et al., 2003, Parilina, 2010, Sagduyu and Ephremides, 2006). The game theory models of the behaviour in ad hoc wireless networks with emphasis on the development of cooperation mechanisms to stimulate package forwarding are considered in (Michiardi and Molva, 2003). Game theoretical models are useful for modeling the data transmission not only in ad hoc but also in CSMA networks (Benslama et al., 2013). The problem of constructing and analyzing the simple mechanism to stimulate the nodes for package forwarding is investigated in (Buttayan and Hubaux, 2003).

The rest of the paper is organized as follows: Section 2 describes the model of non-cooperative stochastic game, and Section 3 deals with the construction of the cooperative version of the stochastic game. Section 4 contains the description of the subgame consistency problem and the method to make the cooperative solu-

¹ The details of the specification of a game played over event trees may be found in (Haurie et al., 2012)

tion subgame-consistent. The idea of strategic support of a cooperative solution is investigated in Section 5. We provide an illustrative example of data transmission in wireless network in Section 6, and briefly conclude in Section 7.

2. Non-cooperative stochastic game

We consider a stochastic game with finite number of states when player's payoff is a discounted sum of the stage payoffs which players obtain along the realized trajectory (sequence of the realized action profiles). The game begins with the chance turn, i.e. with the choice of the initial state of the game which the game process begins with. The state of the stochastic game is determined as a normal form game of n players. One of the finite number of states is realized at each stage of the stochastic game. In the state some action profile is realized depending on the transition probabilities.

Definition 1. Let stochastic game G be determined by the set

$$\left(N, \{ \Gamma^j \}_{j=1}^t, \delta, \pi^0, \{ p(j, k; x^j) \}_{j=\overline{1, t}, k=\overline{1, t}, x^j \in \prod_{i=1}^n X_i^j} \right), \quad (1)$$

where

- $N = \{1, \dots, n\}$ is the set of players.
- $\Gamma^j = \langle N, X_1^j, \dots, X_n^j, K_1^j, \dots, K_n^j \rangle$ is a non-cooperative normal form game which defines the state j , $j = 1, \dots, t$, X_i^j is the finite set of pure actions of player i in Γ^j , $K_i^j(x_1^j, \dots, x_n^j) = K_i^j(x^j)$ is a payoff function of player i in state Γ^j , $j = 1, \dots, t$.
- $p(j, k; x^j)$ is the probability that state Γ^k is realized if at the previous stage (in state Γ^j) action profile $x^j = (x_1^j, \dots, x_n^j)$ has been realized, $p(j, k; x^j) \geq 0$ and $\sum_{k=1}^t p(j, k; x^j) = 1$ for each $x^j \in X^j = \prod_{i \in N} X_i^j$ and for any $j, k = 1, \dots, t$.
- $\delta \in (0, 1)$ is the discount factor.
- $\pi^0 = (\pi_1^0, \dots, \pi_t^0)$ is the vector of the initial distribution on states $\Gamma^1, \dots, \Gamma^t$ where π_j^0 is the probability that state Γ^j is realized at the first stage of the game, $\sum_{j=1}^t \pi_j^0 = 1$.

For constructing the cooperative model of a stochastic game we need to define its subgame and the class of strategies which players use in the game.

Definition 2. Stochastic game (1) with vector $\pi^0 = (0, \dots, 0, 1, 0, \dots, 0)$ (the j -th component is equal to 1), i.e. game beginning from state Γ^j , is called a stochastic subgame G^j , $j = 1, \dots, t$.

We assume that players realise stationary strategies in the game. Let $\Xi_i = \{\eta_i\}$ be the set of stationary strategies of player $i \in N$ in game G . Using stationary strategies a player chooses an action in each state from $\{\Gamma^1, \dots, \Gamma^t\}$ depending only on which state is realized at this stage, i.e. $\eta_i : \Gamma^j \mapsto x_i^j \in X_i^j, j = 1, \dots, t$. Considering stochastic game in stationary strategies, and taking into account that the set of states is finite and the game has an infinite horizon, there are a finite number of subgames of game G . The number of subgames equals the number of the states.

Remark 1. Obviously, the stationary strategy of player i in game G is also a stationary strategy of the player in subgames G^1, \dots, G^t .

The payoff to the player in game G is a random variable. Consider the mathematical expectation of the player's payoff as his payoff in game G . Let $\bar{E}_i(\eta)$ be the expected payoff of player i in game G and $E_i^j(\eta)$ be the expected payoff of player i in subgame G^j when strategy profile η is realised in game G (subgame G^j). Let $E_i(\eta)$ be the vector $(E_i^1(\eta), \dots, E_i^t(\eta))'$.

The expected payoff to player i in subgame G^j satisfies the following recurrent equation:

$$E_i^j(\eta) = K_i^j(x^j) + \delta \sum_{k=1}^t p(j, k; x^j) E_i^k(\eta) \tag{2}$$

s. t. $\eta(\Gamma^j) = x^j$, i. e. $\eta(\cdot) = (\eta_1(\cdot), \dots, \eta_n(\cdot))$ where $\eta_i(\Gamma^j) = x_i^j \in X_i^j$, $x^j = (x_1^j, \dots, x_n^j)$ for any $j = 1, \dots, t$, $i \in N$. Hereinafter, let $\eta(\cdot) = (\eta_1(\cdot), \dots, \eta_n(\cdot))$ be the stationary strategy profile such as $\eta_i(\Gamma^j) = x_i^j \in X_i^j$ where $j = 1, \dots, t$, $i \in N$.

The transition matrix of stochastic game G when stationary strategy profile $\eta(\cdot)$ is realised is:

$$\Pi(\eta) = \begin{pmatrix} p(1, 1; x^1) & \dots & p(1, t; x^1) \\ p(2, 1; x^2) & \dots & p(2, t; x^2) \\ \dots & \dots & \dots \\ p(t, 1; x^t) & \dots & p(t, t; x^t) \end{pmatrix}. \tag{3}$$

We can rewrite equation (2) in a matrix form using matrix (3) in the following way:

$$E_i(\eta) = K_i(x) + \delta \Pi(\eta) E_i(\eta), \tag{4}$$

where $K_i(x) = (K_i^1(x^1), \dots, K_i^t(x^t))$, and $K_i^j(x^j)$ is the payoff to player i in state Γ^j when action profile $x^j \in X^j$ is realized in this state.

Equation (4) is equivalent to the equation²:

$$E_i(\eta) = (\mathbb{I}_t - \delta \Pi(\eta))^{-1} K_i(x), \tag{5}$$

where \mathbb{I}_t is an identity $t \times t$ matrix.

The expected payoff to player i in game G is calculated by formula:

$$\bar{E}_i(\eta) = \pi^0 E_i(\eta). \tag{6}$$

3. Cooperative stochastic game

Suppose that the players from the grand coalition N decide to cooperate and receive the maximal total payoff. Denote the strategy profile maximizing the sum of the expected players' payoffs in game G as $\bar{\eta}(\cdot) = (\bar{\eta}_1(\cdot), \dots, \bar{\eta}_n(\cdot))$:

$$\max_{\eta \in \prod_{i \in N} \Xi_i} \sum_{i \in N} \bar{E}_i(\eta) = \sum_{i \in N} \bar{E}_i(\bar{\eta}). \tag{7}$$

² Matrix $(\mathbb{I}_t - \delta \Pi(\eta))^{-1}$ always exists for $\delta \in (0, 1)$. The proof follows. It is known that all the eigenvalues of stochastic matrix $\Pi(\eta)$ are in the interval $[-1, 1]$. For the existence of matrix $(\mathbb{I}_t - \delta \Pi(\eta))^{-1}$ it is necessary and sufficient that the determinant of matrix $(\Pi(\eta) - \frac{1}{\delta} \mathbb{I}_t)$ be not equal to zero. Thus matrix $(\Pi(\eta) - \frac{1}{\delta} \mathbb{I}_t)$ must not have the eigenvalue to be equal to $\frac{1}{\delta}$. The last condition takes place because $\frac{1}{\delta} > 1$, so this number cannot be the eigenvalue of stochastic matrix $\Pi(\eta)$.

Call the strategy profile $\bar{\eta}(\cdot)$ cooperative strategy profile.

The cooperative model of a non-cooperative game G is given by set (N, v) , where N is the set of players and v is a real-valued function, called the characteristic function of the game, defined on the set 2^N (the set of all subsets of N), and satisfying the property: $v(\emptyset) = 0$. The value $v(S)$ is a real number which is assigned to coalition $S \subset N$, and may be interpreted as the worth or power of coalition S . The members of coalition S play together as a unit.

Define the characteristic function $\bar{v}(S)$ in stochastic game G using characteristic functions $v^j(S)$ of stochastic subgames G^j , $j = 1, \dots, t$, as follows:

$$\bar{v}(S) = \pi^0 v(S) \quad (8)$$

for any coalition $S \subset N$ where $v(S) = (v^1(S), \dots, v^t(S))'$. And $v^j(S)$ is the value of the characteristic function for subgame G^j calculated for coalition S . Now the problem is to define the characteristic function $v^j(S)$ for any coalition S . We use α -approach to define the characteristic function. According to this approach the value of characteristic function for coalition S is equal to the maximal total payoff of coalition S which this coalition can guarantee when the left-out players cooperate and minimize total payoff of coalition S .

First, consider coalition $S = N$. Bellman equation for $v(N)$ is:

$$v(N) = \max_{\eta \in \prod_{i \in N} \Xi_i} \left[\sum_{i \in N} K_i(x) + \delta \Pi(\eta) v(N) \right] = \sum_{i \in N} K_i(\bar{x}) + \delta \Pi(\bar{\eta}) v(N),$$

where $\bar{\eta}(\cdot)$ is the cooperative strategy profile.

Therefore, the value $v(N)$ is:

$$v(N) = (\mathbb{I}_t - \delta \Pi(\bar{\eta}))^{-1} \sum_{i \in N} K_i(\bar{x}). \quad (9)$$

Second, consider coalition $S \subset N$, $S \neq \emptyset$. To define the value of characteristic function $v^j(S)$, $j = 1, \dots, t$, for each subgame G^j , we consider a zero-sum stochastic game G_S^j with two players (coalitions S and $N \setminus S$) where coalition $S \subset N$ plays as a maximizing player and coalition $N \setminus S$ plays as a minimizing player. Define the value $v^j(S)$ for subgame G^j as a maxmin of the payoff of coalition S in stochastic game G_S^j (in fact, the lower value of matrix game):

$$v^j(S) = \max_{\eta_S} \min_{\eta_{N \setminus S}} \sum_{i \in S} E_i^j(\eta_S, \eta_{N \setminus S}), \quad (10)$$

where $(\eta_S(\cdot), \eta_{N \setminus S}(\cdot))$ is a stationary strategy profile such that $\eta_S(\cdot) = (\eta_{i_1}(\cdot), \dots, \eta_{i_k}(\cdot))$ is a vector of stationary strategies of players $i_1, \dots, i_k \in S$, $i_1 \cup \dots \cup i_k = S$, $\eta_S(\cdot) \in \prod_{j=1}^k \Xi_{i_j}$, the set of stationary strategies of coalition $S \subset N$, and $\eta_{N \setminus S}(\cdot)$ is a vector of stationary strategies of players $i_{k+1}, \dots, i_n \in N \setminus S$, $i_{k+1} \cup \dots \cup i_n = N \setminus S$, $\eta_{N \setminus S}(\cdot) \in \prod_{j=k+1}^n \Xi_{i_j}$, the set of stationary strategies of coalition $N \setminus S$.

Third, consider $S = \emptyset$. Let the value of characteristic function be:

$$v^j(\emptyset) = 0. \quad (11)$$

Remark 2. Characteristic functions $\bar{v}(S)$ determined by (8) and $v^j(S)$ determined by (9)–(11) are superadditive.

Definition 3. Cooperative stochastic subgame G_{co}^j is a set (N, v^j) , where N is the set of players and $v^j : 2^N \rightarrow R$ is the characteristic function calculated by (9)–(11).

Definition 4. Cooperative stochastic game G_{co} is a set (N, \bar{v}) , where N is the set of players and $\bar{v} : 2^N \rightarrow R$ is the characteristic function calculated by (8).

Definition 5. Vector $\alpha^j = (\alpha_1^j, \dots, \alpha_n^j)$ satisfying the two following conditions:

1. $\sum_{i \in N} \alpha_i^j = v^j(N)$,
2. $\alpha_i^j \geq v^j(\{i\})$ for any $i \in N$,

is called an imputation in subgame G_{co}^j ($j = 1, \dots, t$). Denote the imputation set in cooperative subgame G_{co}^j as A^j .

Definition 6. The vector $\bar{\alpha} = (\bar{\alpha}_1, \dots, \bar{\alpha}_n)$, where $\bar{\alpha}_i = \pi^0 \alpha_i$, $\alpha_i = (\alpha_i^1, \dots, \alpha_i^t)$, and $(\alpha_1^j, \dots, \alpha_n^j) = \alpha^j \in A^j$ is called an imputation in game G_{co} . Denote the imputation set in cooperative stochastic game G_{co} as \bar{A} .

Suppose that the imputation set in any subgame G_{co}^j , is nonempty, $j = 1, \dots, t$. Therefore, the imputation set in cooperative stochastic game G_{co} is also nonempty.

4. Subgame consistency of the Shapley value

Suppose that players decide to cooperate in stochastic game and for every subgame G_{co}^j they agree to choose an imputation $\alpha^j = (\alpha_1^j, \dots, \alpha_n^j) \in A^j$. The problem is to realize payments to the players at each stage of the stochastic game to guarantee the expected payoff α_i^j for player i in stochastic subgame G^j . If players receive stage payoffs according to their payoff functions they hardly ever obtain the components of the chosen imputation in mathematical expectation sense. To solve this problem we should suggest the method of redistribution of the total players' payoff in every state which may be realized during the game. Initially, the method was proposed by (Petrosyan and Danilov, 1979), for differential games.

There are two principles of constructing the payment scheme in a dynamic game which can be applied to the theory of stochastic games:

1. The sum of the payments to the players in every state is equal to the sum of the players' payoffs in action profile realized in this state according to the cooperative strategy profile $\bar{\eta}(\cdot)$.
2. The expected sum of the payments to player i in the game G is equal to the i th component $\bar{\alpha}_i$ of the imputation $\bar{\alpha}$.

Taking into account that in stochastic game (1) with stationary strategies the number of subgames is equal to the number of possible states, we need to define the vector $\beta_i = (\beta_i^1, \dots, \beta_i^t)$ for every $i \in N$, where β_i^j is a payment to player i in state Γ^j , $j = 1, \dots, t$. If these payments satisfy the two mentioned principles, they are called imputation distribution procedure (IDP) (see (Petrosyan and Danilov, 1979)). We are interested in constructing the subgame-consistent (time-consistent) IDP.

Definition 7. We call the IDP subgame-consistent in stochastic game G if for any subgame of game G the vector of the expected discounted sums of the payments to the players $1, \dots, n$ belong to the same cooperative solution³.

³ Let the cooperative solution be a singleton like the Shapley value. Then in any subgame we consider the Shapley value as a cooperative solution. The case where the cooperative solution is the set (e.g., the core) is considered in details by Parilina and Zaccour, 2015a for the games played over event trees.

In the paper we examine the Shapley value as a cooperative solution. Therefore, the subgame-consistent IDP guarantees any player to obtain the corresponding component of the Shapley value in any subgame.

Theorem 1. *Let the components of the IDP be calculated by equation:*

$$\beta_i = (\mathbb{I}_t - \delta\Pi(\bar{\eta}))\alpha_i, \tag{12}$$

where $\alpha_i = (\alpha_i^1, \dots, \alpha_i^t)$, and $(\alpha_1^j, \dots, \alpha_n^j) = \alpha^j$ is the Shapley value in the cooperative game G_{co}^j with characteristic function $v^j(S)$. Then the IDP is subgame-consistent.

Proof. First, prove that $\beta_i, i \in N$, calculated by equation (12) is the IDP. Taking into account equation (9) we obtain

$$\sum_{i \in N} \beta_i = (\mathbb{I}_t - \delta\Pi(\bar{\eta})) \sum_{i \in N} \alpha_i = (\mathbb{I}_t - \delta\Pi(\bar{\eta}))v(N) = \sum_{i \in N} K_i(\bar{x}).$$

Second, we calculate the expected sum of the payments to player i in the game G according to equation (12). Denote this sum for player i by B_i and it satisfies the equation:

$$\bar{B}_i = \pi^0 B_i = \pi^0(B_i^1, \dots, B_i^t)',$$

where B_i^j can be found from equation:

$$B_i^j = \beta_i^j + \delta \sum_{k=1}^t p(j, k; x^j) B_i^k,$$

or in a vector form:

$$B_i = \beta_i + \delta\Pi(\bar{\eta})B_i. \tag{13}$$

Equation (13) is equivalent to the following one:

$$B_i = (\mathbb{I}_t - \delta\Pi(\bar{\eta}))^{-1} \beta_i. \tag{14}$$

Taking into account equation (12), we prove that $B_i = \alpha_i$ and then $\bar{B}_i = \bar{\alpha}_i$. The equity $B_i = \alpha_i$ proves the subgame consistency of the IDP determined by equation (12).

Remark 3. Equation (12) is equivalent to the following one:

$$\alpha_i = \beta_i + \delta\Pi(\bar{\eta})\alpha_i. \tag{15}$$

The second summand at the right-hand side of the equation (15) is the expected value of the component of the Shapley value in subgame starting from the next stage. Therefore, any player will receive his component of the Shapley value in any subgame if the payments to the players are the IDP satisfying equation (12).

Obviously, if players realise the cooperative strategy profile $\bar{\eta}(\cdot)$, the expected payoff of player i in stochastic game G with new payments in cooperative action profiles is equal to the expected value of the correspondent component of the Shapley value in cooperative stochastic game G_{co} .

Now for the imputation $\bar{\alpha} = (\bar{\alpha}_1, \dots, \bar{\alpha}_n)$, where $\bar{\alpha}_i = \pi^0 \alpha_i, \alpha_i = (\alpha_i^1, \dots, \alpha_i^t), (\alpha_1^j, \dots, \alpha_n^j) = \alpha^j \in A^j$, we determine the regularization of game G in the following way.

Definition 8. Noncooperative stochastic game G_α (subgame G_α^j , $j = 1, \dots, t$) is called α -regularization of stochastic game G (subgame G^j), if for any player $i \in N$ in state Γ^j payoff function $K_i^{\alpha,j}(x^j)$ is defined as follows:

$$K_i^{\alpha,j}(x^j) = \begin{cases} \beta_i^j, & \text{if } x^j = \bar{x}^j; \\ K_i^j(x^j), & \text{if } x^j \neq \bar{x}^j, \end{cases} \quad (16)$$

where IDP $\beta = (\beta_1, \dots, \beta_n)$ satisfies equation (12)⁴.

We suggest a method of construction of the new payoff function in the game G (subgame G^j) in every state when the action profile is cooperative. Here we may ask a question: “Do the players agree to redefine the payoff function in the game?” Our answer is “Yes”, if they want to make the payoff functions in the states subgame-consistent in the sense of Definition 7. Redistributing the payoffs using the IDP $\beta_1^1, \dots, \beta_n^t$ in the states $\Gamma^1, \dots, \Gamma^t$ respectively, player i receives the same sum (in terms of mathematical expectation) in game G_α (G_α^j) as he has planned to receive in the cooperative stochastic game G_{co} (G_{co}^j). Moreover, in any subgame his expected payoff will be the corresponding component of the Shapley value. In this case, we can state the *subgame consistency (time consistency)* of the chosen cooperative solution.

5. Strategic support of the Shapley value

In this section we need to consider the additional notations. Let $\Gamma(k) \in \{\Gamma^1, \dots, \Gamma^t\}$ be the state realized at stage k of game G_α . Let $x(k)$ be the action profile realized in state $\Gamma(k)$. Denote the subgame of game G_α starting from state $\Gamma(k)$ as $G_\alpha^{\Gamma(k)}$. Call the sequence $((\Gamma(1), x(1)), (\Gamma(2), x(2)), \dots, (\Gamma(k-1), x(k-1)))$ the history of stage k and denote it as $h(k)$. Let T be the set $\{(\Gamma^1, \bar{x}^1), (\Gamma^2, \bar{x}^2), \dots, (\Gamma^t, \bar{x}^t)\}$.

In this section we consider stochastic game G_α as the game with perfect information in the sense that at each stage k ($k = 1, 2, \dots$) all players know state $\Gamma(k)$ and the history of stage k . We would like to prove that the cooperative strategy profile in the game G_α is the Nash equilibrium in trigger strategies. To construct the Nash equilibrium we need to consider the sets of behavior strategies Φ_i , $i \in N$.

Definition 9. We call the behavior strategy profile $\varphi^* = (\varphi_1^*, \dots, \varphi_n^*)$ the Nash equilibrium in game G_α if for any player $i \in N$ the inequality

$$\bar{E}_i^\alpha(\varphi^*) \geq \bar{E}_i^\alpha(\varphi^* \parallel \varphi_i) \quad (17)$$

is true for any behavior strategy $\varphi_i \in \Phi_i$ of player i , and $\bar{E}_i^\alpha(\cdot)$ is the expected payoff of player i in α -regularization G_α .

The following theorem gives the condition when the cooperative strategy profile in the game G_α is the Nash equilibrium in the α -regularization G_α of game G .

Theorem 2. *If in the α -regularization G_α the following inequality is true for any coalition player $i \in N$:*

$$\beta_i \geq (\mathbb{I}_t - \delta \Pi(\bar{\eta}))F(\{i\}), \quad (18)$$

⁴ The IDP for stochastic games was initially proposed in (Petrosjan, 2006) when the game process is realised on a graph and in (Petrosjan and Baranova, 2006) when the number of states in stochastic game is finite and players use stationary strategies.

where $F(\{i\}) = (F^1(\{i\}), \dots, F^t(\{i\}))$,
 $F^j(\{i\}) = \max_{\substack{x_i^j \in X_i^j \\ x_i^j \neq \bar{x}_i^j}} \left\{ K_i^j(\bar{x}^j \parallel x_i^j) + \delta \sum_{l=1}^t p(j, l; \bar{x}^j \parallel x_i^j) v^l(\{i\}) \right\}$, then in the game G_α
 there exists the Nash equilibrium with payoffs $(\bar{\alpha}_1, \dots, \bar{\alpha}_n)$.

Proof. Consider the behavior strategy profile $\hat{\varphi}(\cdot) = (\hat{\varphi}_1(\cdot), \dots, \hat{\varphi}_n(\cdot))$ in game G_α :

$$\hat{\varphi}_i(h(k)) = \begin{cases} \bar{x}_i^j, & \text{if } \Gamma(k) = \Gamma^j, j = \overline{1, t}, h(k) \subset T; \\ \hat{x}_i^j(p), & \text{if } \Gamma(k) = \Gamma^j, j = \overline{1, t}, \exists l \in [1, k-1] \\ & \text{and } p \in N, p \neq i: h(l) \subset T, \\ & \text{and } (\Gamma(l), x(l)) \notin T, \\ & \text{but } (\Gamma(l), (x(l) \parallel \bar{x}_p(l))) \in T, \\ \text{anyone} & \text{in other cases,} \end{cases} \quad (19)$$

where $\hat{x}_i^j(p)$ is an action of player i in state Γ^j which with actions $x_k^j, k \neq i, k \neq p$ forms the strategy of coalition $\{N \setminus p\}$ in zero-sum game against player p in subgame G^{Γ^j} .

The proof of the theorem is based on the proof of any folk theorem (for example, see (Dutta, 1995)) using the structure of the trigger strategy (19). We prove that $\hat{\varphi}(\cdot) = (\hat{\varphi}_1(\cdot), \dots, \hat{\varphi}_n(\cdot))$ determined in (19) is a Nash equilibrium in stochastic game G_α .

If the player p does not deviate from the cooperative strategy profile $\bar{\eta}$, then taking into account the definition of the strategy (19), the expected payoff of player p in the subgame $G_\alpha^j, j = 1, \dots, t$, is

$$E_p^j(\hat{\varphi}(\cdot)) = E_p^j(\bar{\eta}(\cdot)).$$

Let $E_p(\hat{\varphi}(\cdot))$ be equal to the vector $(E_p^1(\hat{\varphi}(\cdot)), \dots, E_p^t(\hat{\varphi}(\cdot)))$, then for any player $p \in N$ the equation is true:

$$E_p(\hat{\varphi}) = (\mathbb{I}_t - \delta \Pi(\bar{\eta}))^{-1} \beta_p. \quad (20)$$

Consider the strategy profile $(\hat{\varphi}(\cdot) \parallel \varphi_p(\cdot)), p \in N$, when player $p \in N$ deviates from strategy $\hat{\varphi}_p(\cdot)$. Let stage k be such that there exists number $l \in [1, k-1]$ such that history $h(l) \subset T$ and state $(\Gamma(l), x(l)) \notin T$ but $(\Gamma(l), (x(l) \parallel \bar{x}_p(l))) \in T$. Without loss of generality, we suggest that $\Gamma(k) = \Gamma^j$. Calculate the payoff of player p in game G_α in strategy profile $(\hat{\varphi}(\cdot) \parallel \varphi_p(\cdot))$ as

$$\bar{E}_p^\alpha(\hat{\varphi} \parallel \varphi_p) = \pi^0 E_p^\alpha(\hat{\varphi} \parallel \varphi_p),$$

where

$$E_p^\alpha(\hat{\varphi} \parallel \varphi_p) = E_p^{\alpha, [1, k-1]}(\hat{\varphi} \parallel \varphi_p) + \delta^{k-1} \Pi^{k-1}(\hat{\varphi} \parallel \varphi_p) E_p^{\alpha, [k, \infty)}(\hat{\varphi} \parallel \varphi_p), \quad (21)$$

where $E_p^{\alpha, [1, k-1]}(\hat{\varphi} \parallel \varphi_p)$ is the expected payoff of player p at the first $k-1$ stages of game G_α , and $E_p^{\alpha, [k, \infty)}(\hat{\varphi} \parallel \varphi_p)$ is the expected payoff of player p in the subgame of game G_α starting from stage k . Since there were no deviations of any players

from the cooperative strategy profile $\bar{\eta}(\cdot)$ up to stage $k - 1$ inclusive as it was shown before, the following equalities holds for the elements of the right side of (21):

$$E_p^{\alpha, [1, k-1]}(\hat{\varphi} \parallel \varphi_p) = E_p^{\alpha, [1, k-1]}(\bar{\eta}),$$

$$\Pi^{k-1}(\hat{\varphi} \parallel \varphi_p) = \Pi^{k-1}(\bar{\eta}).$$

In the second term of the right side of (21) by $E_p^{\alpha, [k, \infty)}(\hat{\varphi} \parallel \varphi_p)$ we mean vector $(E_p^{\alpha, 1}(\hat{\varphi} \parallel \varphi_p), \dots, E_p^{\alpha, t}(\hat{\varphi} \parallel \varphi_p))$ where $E_p^{\alpha, j}(\hat{\varphi} \parallel \varphi_p)$ is the expected payoff of player p in regularized subgame G_α^j starting from the state Γ^j .

Now we calculate the expected payoff of player p in subgame G_α^j starting from stage k when the state $\Gamma(k)$ is Γ^j :

$$E_p^{\alpha, j}(\hat{\varphi} \parallel \varphi_p) = K_p^j(\bar{x}^j \parallel x_p^j) + \delta \sum_{l=1}^t p(j, l; \bar{x}^j \parallel x_p^j) v^l(\{p\}), \quad (22)$$

because players from coalition $N \setminus p$ will punish player p playing in zero-sum game against player p beginning from stage $k + 1$ according to the definition of strategy profile $\hat{\varphi}(\cdot)$.

Since the expected payoffs of player p in strategy profiles $\hat{\varphi}(\cdot)$ and $(\hat{\varphi}(\cdot) \parallel \varphi_p(\cdot))$ equal until stage $k - 1$, then as a result of the deviation, player p can guarantee the increase of his payoff only at the sacrifice of the part of game G_α beginning with stage k , i.e. at the sacrifice of the expected payoff in subgame G_α^j , $j = 1, \dots, t$. Player p in strategy profile $(\hat{\varphi}(\cdot) \parallel \varphi_p(\cdot))$ can guarantee the following expected payoff from stage k :

$$\max_{\substack{x_p^j \in X_p^j \\ x_p^j \neq \bar{x}_p^j}} \left\{ K_p^j(\bar{x}^j \parallel x_p^j) + \delta \sum_{l=1}^t p(j, l; \bar{x}^j \parallel x_p^j) v^l(\{p\}) \right\}. \quad (23)$$

According to the definition of IDP, the expected payoff of player p in the subgame G_α^j in strategy profile $\hat{\varphi}(\cdot)$ can be calculated by the equation:

$$E_p^\alpha(\hat{\varphi}) = (\mathbb{I}_t - \delta \Pi(\bar{\eta}))^{-1} \beta_p, \quad (24)$$

where $E_p^\alpha(\hat{\varphi}) = (E_p^{\alpha, 1}(\hat{\varphi}(\cdot)), \dots, E_p^{\alpha, t}(\hat{\varphi}(\cdot)))$. Taking into account the inequality (18) and (23), (24) and we prove the inequality

$$E_p^\alpha(\hat{\varphi}(\cdot)) \geq E_p^\alpha(\hat{\varphi}(\cdot) \parallel \varphi_p(\cdot)).$$

Therefore, the behavior strategy profile (19) is the Nash equilibrium in α -regularization of game G . The expected payoff of player p in game G_α in strategy profile $\hat{\varphi}(\cdot)$ is equal to $\bar{\alpha}_p$ where $\bar{\alpha}_p = \pi^0 \alpha_p$ and vector $\alpha_p = (\alpha_p^1, \dots, \alpha_p^t)$ consists of p -th components of the Shapley value $\alpha^1, \dots, \alpha^t$ calculated for the cooperative subgames G^1, \dots, G^t accordingly. This completes the proof.

6. Data transmission game

6.1. Model

In this section we introduce an example of a stochastic game application in telecommunication systems. We consider a slotted synchronous system in which nodes 1

and 2 independently generate packages in each time slot with probabilities a_1 and a_2 , respectively, provided that their individual queues were empty at the end of the previous time slot. The graph of wireless network is depicted in Fig. 1. Some assumptions about this system are as follows:

1. Nodes 1 and 2 (players 1 and 2, respectively) are going to send their packages to a common destination (node 3).
2. The maximum buffer capacity of any node equals one. The destination node can accept only one transmitted package in one time slot. We do not assume multiple package transmissions or simultaneous transmissions and reception by any node in any time slot.
3. If players simultaneously transmit packages to the destination node, the last one rejects these packages and they return to their initial nodes, i.e. at the next time slot no new packages can be generated in nodes 1 and 2.
4. All transmitted packages have the same length, and it requires one time slot to transmit a package from one node to the other which has the direct channel with the first one.
5. Player 1 chooses between sending a package directly to node 3 or relying on node 2 to forward the package to the final destination (node 3).
6. If player 1 (node 1) transmits a package to player 2 (node 2) which has already had a package in its queue, player 2 rejects this package. Otherwise, player 2 decides on whether to accept or reject the package from player 1.

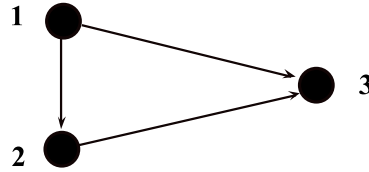


Fig. 1: Topology of a wireless network.

We suggest the following system of rewards and costs:

- $f \geq 0$ is a reward to player 1 or player 2 for each successful transmission to the destination node.
- Player 1 receives a reward $c \geq 0$ from player 2 for delivering a package to player 2 which can obtain the value f only after successful transmission of that particular package to the final destination in a subsequent time slot.
- Each time slot of package delay results in an additional cost $d \geq 0$ for the node that has that particular package in its queue (regardless of that package source).
- D_{ij} is an energy cost of one package transmission from node i to node j .

We suppose that the game ends in any time slot with the probability $0 < q < 1$. The probability $1 - q$ can be interpreted as a discount rate. The transmission problem in a wireless network can be solved as a stochastic game. Denote the pair (Q_1, Q_2) as the state of the stochastic game where Q_i is a queue content of node i , $i = 1, 2$. The queue content Q_i can be equal to 0 or 1 if no or one package is present at the queue of node i , respectively.

The set of states in stochastic game is

$$Q = \{(0, 0); (0, 1); (1, 0); (1, 1)\}.$$

Consider the game in a cooperative setting meaning that the players' actions are coordinated by one center to improve the work of the network. The coordination of the device actions are useful to increase the speed of data transmission. For solving the cooperative version of the stochastic game, we assume players have information not only on their own queues but also on the other player's queue.

Now we need to describe the states, i. e., the games in normal form corresponding to the states:

1. State $(0, 0)$: Player 1 has a unique action W (waiting), player 2 has the same action W (waiting). The payoffs to the players are $(0, 0)$.
2. State $(0, 1)$: Player 1 has a unique action W (waiting), player 2 also has a unique action $T3$ (transmission to node 3). The payoffs to the players are $(0, f - D_{23})$.
3. State $(1, 0)$: Player 1 has two actions: i) $T3$ (transmission to node 3), ii) $T2$ (transmission to node 2). Player 2 has two actions: i) Ac (accepting a package from node 1), ii) Rej (rejecting a package from node 1). The payoffs to the players are represented in the matrix:

$$\begin{pmatrix} (f - D_{13}, 0) & (f - D_{13}, 0) \\ (c - D_{12}, -c) & (-d - D_{12}, 0) \end{pmatrix}$$

4. State $\Gamma(1, 1)$: Player 1 has two actions: i) $T3$ (transmission to node 3), ii) W (waiting). Player 2 has two actions: i) $T3$ (transmission to node 3), ii) W (waiting). The payoffs to the players are as follows:

$$\begin{pmatrix} (-d - D_{13}, -d - D_{23}) & (f - D_{13}, -d) \\ (-d, f - D_{23}) & (-d, -d) \end{pmatrix}$$

6.2. Transition matrix

Assume the players use the stationary strategies. In the game defined in stationary strategies the players' choice of an action in the states depends neither on the history, nor on the time slot, in which the game is at present, but depends only on the state. In applications of stochastic games it is important to use a simple set of strategies for decreasing the number of calculations of players' expected payoffs.

Denote the set of mixed stationary strategies of player i as $\bar{\Xi}_i$, $i = 1, 2$. According to the game structure the player 1's mixed stationary strategy assigns him to choose action W with probability one in the states $(0, 0)$, $(0, 1)$, action $T3$ with probability p_{11} in the state $(1, 0)$, and action $T3$ with probability p_{12} in the state $(1, 1)$. The player 2's mixed stationary strategy assigns him to choose action W with probability one in the state $(0, 0)$, action $T3$ in the state $(0, 1)$, action Ac with probability p_{21} in the state $(1, 0)$, and action $T3$ with probability p_{22} in the state $(1, 1)$. Denote a player i 's mixed stationary strategy as $\eta_i = (p_{i1}, p_{i2})$. A stationary strategy profile is $\eta = (\eta_1, \eta_2) = (p_{11}, p_{12}, p_{21}, p_{22})$.

The transition matrix when players realise stationary strategy profile η is

$$H(\eta) = \{p(k, l; x^k)\}_{k=1, \dots, t; l=1, \dots, t}, \quad (25)$$

where

$$p(1, 1; x^1) = (1 - a_1)(1 - a_2),$$

$$\begin{aligned}
 p(1, 2; x^1) &= (1 - a_1)a_2, \\
 p(1, 3; x^1) &= a_1(1 - a_2), \\
 p(1, 4; x^1) &= a_1a_2, \\
 p(2, 1; x^2) &= (1 - a_1)(1 - a_2), \\
 p(2, 2; x^2) &= (1 - a_1)a_2, \\
 p(2, 3; x^2) &= a_1(1 - a_2), \\
 p(2, 4; x^2) &= a_1a_2, \\
 p(3, 1; x^3) &= p_{11}(1 - a_1)(1 - a_2), \\
 p(3, 2; x^3) &= p_{11}(1 - a_1)a_2 + (1 - p_{11})p_{21}(1 - a_1), \\
 p(3, 3; x^3) &= p_{11}a_1(1 - a_2) + (1 - p_{11})(1 - p_{21})(1 - a_2), \\
 p(3, 4; x^3) &= p_{11}a_1a_2 + (1 - p_{11})p_{21}a_1 + (1 - p_{11})(1 - p_{21})a_2, \\
 p(4, 1; x^4) &= 0, \\
 p(4, 2; x^4) &= p_{12}(1 - p_{22})(1 - a_1), \\
 p(4, 3; x^4) &= (1 - p_{12})p_{22}(1 - a_2), \\
 p(4, 4; x^4) &= p_{12}p_{22} + (1 - p_{12})(1 - p_{22}) + p_{12}(1 - p_{22})a_1 + (1 - p_{12})p_{22}a_2.
 \end{aligned}$$

6.3. Payoff functions

If the stationary strategy profile η is realized, the payoff to player 1 in the stochastic game is

$$K_1(x) = (K_1^1(x^1), K_1^2(x^2), K_1^3(x^3), K_1^4(x^4))',$$

where

$$\begin{aligned}
 K_1^1(x^1) &= K_1^2(x^2) = 0, \\
 K_1^3(x^3) &= p_{11}(f - D_{13}) + (1 - p_{11})p_{21}(c - D_{12}) + (1 - p_{11})(1 - p_{21})(-d - D_{12}), \\
 K_1^4(x^4) &= p_{12}p_{22}(-d - D_{13}) + p_{12}(1 - p_{22})(f - D_{13}) + (1 - p_{12})(-d).
 \end{aligned}$$

If the stationary strategy profile η is realized, the payoff to player 2 in the stochastic game is

$$K_2(x) = (K_2^1(x^1), K_2^2(x^2), K_2^3(x^3), K_2^4(x^4))',$$

where

$$\begin{aligned}
 K_2^1(x^1) &= 0, \\
 K_2^2(x^2) &= f - D_{23}, \\
 K_2^3(x^3) &= (1 - p_{11})p_{21}(-c), \\
 K_2^4(x^4) &= p_{12}p_{22}(-d - D_{23}) + (1 - p_{12})p_{22}(f - D_{23}) + (1 - p_{22})(-d).
 \end{aligned}$$

We consider the set of pure stationary strategies which is denoted as $\Xi_i, i = 1, 2$. For example, player 1's pure stationary strategy $\eta_1 = (1, 0)$ assigns player 1 to choose action $T3$ in the state $(1, 0)$ and action W in the state $(1, 1)$. Each player has 4 pure stationary strategies in the stochastic game, therefore, there are 16 pure stationary strategy profiles. For each pure stationary strategy profile $\eta = (\eta_1, \eta_2)$ the transition matrix $\Pi(\eta)$ is determined by (25).

For example, for the pure stationary strategy profile $\eta^1 = (1, 1, 1, 1)$ the transition matrix is

$$\Pi(\eta^1) = \begin{pmatrix} (1 - a_1)(1 - a_2) & (1 - a_1)a_2 & a_1(1 - a_2) & a_1a_2 \\ (1 - a_1)(1 - a_2) & (1 - a_1)a_2 & a_1(1 - a_2) & a_1a_2 \\ (1 - a_1)(1 - a_2) & (1 - a_1)a_2 & a_1(1 - a_2) & a_1a_2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

For each strategy profile $\eta \in \Xi = \prod_{i=1}^2 \Xi_i$ we can calculate the expected players' payoffs for subgames which are denoted as: $E_i(\eta) = (E_i^{(0,0)}(\eta), E_i^{(0,1)}(\eta), E_i^{(1,0)}(\eta),$

$E_i^{(1,1)}(\eta))'$, and

$$E_i(\eta) = (\mathbb{I}_t - (1 - q)II(\eta))^{-1}K_i(x), \quad (26)$$

where $K_i(x)$, $II(\eta)$ are determined above.

The expected payoff to player i in the whole game including the chance move is

$$\bar{E}_i(\eta) = \pi^0 E_i(\eta), \quad (27)$$

where $\pi^0 = (\pi_{(0,0)}^0, \pi_{(0,1)}^0, \pi_{(1,0)}^0, \pi_{(1,1)}^0)$ is a vector of the initial probabilities, and π_k^0 is the probability that the first state in the stochastic game will be $k \in Q$. Vector π^0 is given.

6.4. Algorithm of solving cooperative stochastic game

In this section we describe the steps of solving cooperative stochastic game of data transmission in a wireless network of topology represented in Fig. 1.

1. For any state $k \in Q$ and any pure strategy profile $\eta = (\eta_1, \eta_2)$, $\eta_i \in \Xi_i$, $i = 1, 2$, calculate the expected players' payoffs $E_i^k(\eta)$ in subgame G^k by equation (5) and their expected payoffs in the whole game $\bar{E}_i(\eta)$ by equation (6).
2. Find the cooperative strategy profile $\bar{\eta}$ by equation (7).
3. Calculate the values of the characteristic functions $v^k(S)$ for any state $k \in Q$ and any coalition $S \subset N$ using equations (9), (10), (11). Then calculate the values of the characteristic function $\bar{v}(S)$ for any $S \subset N$ by (8).
4. Calculate the Shapley values $\alpha^k = (\alpha_1^k, \dots, \alpha_n^k)$ for any subgame G^k starting from state $k \in Q$ using formula (Shapley, 1953b):

$$\alpha_i^k = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} (v^k(S \cup \{i\}) - v^k(S)). \quad (28)$$

Then calculate the Shapley value for the whole game $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2)$ using equation $\bar{\alpha}_i = \pi^0 \alpha_i$.

5. Calculate the components of the IDP β_i^k , $i = 1, 2$ and $k \in Q$ by equation (12).

6. To construct the subgame-consistent Shapley value we determine the α -regularization G_α re-defining the payoff functions by equation (16).
7. Verify if there exists the Nash equilibrium in behavior strategies with payoffs $(\bar{\alpha}_1, \bar{\alpha}_2)$ using inequality (18).

6.5. Numerical illustration

We introduce the numerical example of the data transmission game for wireless network. We identify the parameters of the simulation. The probability of package appearance at node 1 is higher than in node 2: $a_1 = 0.4$, $a_2 = 0.1$. The probability of a game end is $q = 0.01$ which is equivalent to the discount rate 0.99. The rewards and costs are $f = 1$, $d = 0.1$, $c = 0.3$, $D_{12} = 0.1$, $D_{13} = 0.6$, $D_{23} = 0.2$. We may notice that the cost of package transmission from node 1 to node 3 is three times more than the cost of package transmission from node 2 to node 3. Therefore, the cooperation of nodes 1 and 2 may be profitable. Let the game begin from any state with equal probability, i. e., $\pi^0 = (0.25, 0.25, 0.25, 0.25)$.

Table 1 represents the expected players' payoffs

$$E_i(\eta) = \left(E_i^{(0,0)}(\eta), E_i^{(0,1)}(\eta), E_i^{(1,0)}(\eta), E_i^{(1,1)}(\eta) \right)'$$

for any pure stationary strategy profile η for any player $i = 1, 2$, and the sum of the expected payoffs. The last column in Table 1 is $\bar{E}_1 + \bar{E}_2$ which is the total expected players' payoff in the whole game taking into account the vector of initial probabilities π^0 .

The cooperative strategy profile maximizing the total players' payoff is

$$\bar{\eta} = \eta^{11} = (0, 0, 1, 1),$$

in which the player 1's strategy $\eta_1^{11} = (0, 0)$ assigns him "not to transmit to node 3, but transmit to node 2" in state (1,0) when there is a package at node 1 and there is no package at node 2. In this state the player 2's strategy $\eta_2^{11} = (1, 1)$ assigns her "to accept the package" from player 1 in state (1, 0). When the game in state (1, 1), player 1 "waits" and player 2 "transmits" to node 3.

The maximum of the total expected players' payoff in the whole game is

$$\max_{\eta \in \Xi} \sum_{i \in N} \bar{E}_i(\eta) = \sum_{i \in N} \bar{E}_i(\bar{\eta}) = 26.9472.$$

Table 1: Expected payoffs in the stochastic game.

η	$E_1(\eta)$	$E_2(\eta)$	$\bar{E}_1 + \bar{E}_2$
$\eta^1 = (1, 1, 1, 1)$	-53.0129	-22.9935	-81.7048
	-53.0129	-22.1935	
	-52.6129	-22.9935	
	-70.1000	-30.0000	
$\eta^2 = (1, 1, 1, 0)$	15.8400	6.76818	23.1355
	15.8400	7.56818	
	16.2400	6.76818	
	16.2400	7.27732	
$\eta^3 = (1, 0, 1, 1)$	14.7353	7.9200	23.1855
	14.7353	8.7200	
	15.1353	7.9200	
	14.8563	8.7200	
$\eta^4 = (1, 0, 1, 0)$	-5.10968	-7.02581	-13.8016
	-5.10968	-6.22581	
	-4.70968	-7.02581	
	-10.0000	-10.0000	
$\eta^5 = (1, 1, 0, 1)$	-53.0129	-22.9935	-81.7048
	-53.0129	-22.1935	
	-52.6129	-22.9935	
	-70.0000	-30.0000	
$\eta^6 = (1, 1, 0, 0)$	15.8400	6.76818	23.1355
	15.8400	7.56818	
	16.2400	6.76818	
	16.2400	7.27732	
$\eta^7 = (1, 0, 0, 1)$	14.7353	7.9200	23.1855
	14.7353	8.7200	
	15.1353	7.9200	
	14.8563	8.7200	
$\eta^8 = (1, 0, 0, 0)$	-5.10968	-7.02581	-13.8016
	-5.10968	-6.22581	
	-4.70968	-7.02581	
	-10.0000	-10.0000	
$\eta^9 = (0, 1, 1, 1)$	-64.7464	-27.3398	-94.3241
	-64.7464	-26.5398	
	-65.9794	-27.9446	
	-70.0000	-30.0000	

η	$E_1(\eta)$	$E_2(\eta)$	$\bar{E}_1 + \bar{E}_2$
$\eta^{10} = (0, 1, 1, 0)$	11.5347	13.4228	25.4926
	11.5347	14.2228	
	11.8060	13.6218	
	12.0060	13.8218	
$\eta^{11} = (0, 0, 1, 1)$	4.90248	21.4759	26.9472
	4.90248	22.2759	
	5.04298	21.8337	
	4.87602	22.4792	
$\eta^{12} = (0, 0, 1, 0)$	-8.93504	-8.73596	-18.1458
	-8.93504	-7.93596	
	-9.06741	-8.97396	
	-10.0000	-10.0000	
$\eta^{13} = (0, 1, 0, 1)$	-64.2491	-26.728	-93.4537
	-64.2491	-25.928	
	-65.4128	-27.248	
	-70.0000	-30.000	
$\eta^{14} = (0, 1, 0, 0)$	-8.4855	5.60847	-2.43475
	-8.4855	6.40847	
	-8.8128	5.57380	
	-7.6828	6.13680	
$\eta^{15} = (0, 0, 0, 1)$	-18.532	7.9200	-10.426
	-18.532	8.7200	
	-19.010	7.9200	
	-18.910	8.7200	
$\eta^{16} = (0, 0, 0, 0)$	-10.559	-8.8313	-19.49526
	-10.559	-8.0313	
	-10.917	-9.0826	
	-10.000	-10.000	

We calculate the values of the characteristic functions for subgames by equations (9), (10), (11):

$$\begin{aligned}
 v(\{1\}) &= (-5.10968, -5.10968, -4.70968, -10.0)', \\
 v(\{2\}) &= (-8.735960, -7.93596, -8.97396, -10.0)', \\
 v(\{1, 2\}) &= (26.3784, 27.1784, 26.8766, 27.3553)',
 \end{aligned}$$

The characteristic function of the whole game is found by equation (8):

$$\begin{aligned}
 \bar{v}(\{1\}) &= -6.23226, \\
 \bar{v}(\{2\}) &= -8.91147, \\
 \bar{v}(\{1, 2\}) &= 26.9472.
 \end{aligned}$$

Then we may calculate the Shapley Values for the subgames and the whole game using equation (28):

- for subgames:
 - $\alpha_1 = (15.0023, 15.0023, 15.5705, 13.6776)'$,
 - $\alpha_2 = (11.376, 12.176, 11.3062, 13.6776)'$,

- for the whole game:
 - $\bar{\alpha}_1 = 14.8132$,
 - $\bar{\alpha}_2 = 12.134$.

The cooperative payoff distribution procedure β_1 for player 1 and β_2 for player 2 are found by equations (12) using the Shapley values:

- $\beta_1 = (0, 0, 1.24274, -1.54974)'$,
- $\beta_2 = (0, 0.8, -1.34274, 2.24974)'$,

where β_i^k is a payment to player i in the state k . Remind that the payoffs of the players in the states defined in the matrix forms are as follows:

- $K_1 = (0, 0, 0.2, -0.1)'$,
- $K_2 = (0, 0.8, -0.3, 0.8)'$,

We may notice that in states $(0, 0)$ and $(0, 1)$ the components of IDP coincide with the payoffs according to the payoff functions K_1 and K_2 . But in states $(1, 0)$ and $(1, 1)$ there is a redistribution of the total payoffs among the players. In state $(1, 0)$ players obtain -0.1 together and according to the IDP player 1 receives 1.24274 instead of 0.2 , and player 2 receives -1.34274 instead of -0.3 . Therefore, player 2 gives 1.04274 to player 1 to make IDP subgame-consistent. In state $(1, 1)$ players obtain 0.7 together and according to the IDP player 1 receives -1.54974 instead of -0.1 , and player 2 receives 2.24974 instead of 0.8 . Therefore, player 1 gives 1.44974 to player 2 to make IDP subgame-consistent.

Thus, the Shapley Value $\bar{\alpha} = (14.8132, 12.134)'$ is subgame consistent if the payoffs to the players in the states are made according to the IDP $\beta_1 = (0, 0, 1.24274, -1.54974)'$, and $\beta_2 = (0, 0.8, -1.34274, 2.24974)'$.

Now we need to examine the problem of strategic support of the cooperative strategy profile. First, calculate values $F^k(\{i\})$ for $i = 1, 2$ and $k \in Q$ determined in Theorem 2:

1. $F(\{1\}) = (-5.10968, -5.10968, -4.70968, -5.28632)'$,
2. $F(\{2\}) = (-8.73596, -7.93596, -8.97396, -8.1858)'$.

Second, verify if the inequalities (18) are true. For player 1 the inequality (18) takes the form:

$$\begin{pmatrix} 0 \\ 0 \\ 1.24274 \\ -1.54974 \end{pmatrix} \geq \begin{pmatrix} -0.186662 \\ -0.186662 \\ 0.418853 \\ -0.566649 \end{pmatrix}$$

We notice that the inequality is not true. In state $(1, 1)$ there is an intense for deviation of player 1 as his payoff according to IDP in this state -1.54974 is less than his payoff in case of deviation -0.566649 . This means that the cooperation cannot be supported strategically by the behavior strategy profile (19).

For player 2 the inequality (18) takes the form:

$$\begin{pmatrix} 0 \\ 0.8 \\ -1.34274 \\ 2.24974 \end{pmatrix} \geq \begin{pmatrix} -0.0718427 \\ 0.728157 \\ -1.01842 \\ 0.620393 \end{pmatrix}$$

Again, we notice that the inequality is not true. In state $(1, 0)$ there is an intense for deviation of player 2 as her payoff according to IDP in this state -1.34274 is lower than her payoff in case of deviation -1.01842 . We state that the behavior strategy profile determined by (19) cannot strategically support the cooperative payoffs.

7. Conclusion

We have examined the problem of cooperation in a dynamic game having a stochastic structure. First, we construct the subgame consistent cooperative solution of the game by redefining players' state payoff functions using the imputation distribution procedure. Second, we provide the conditions to verify if the cooperative solution can be supported strategically. All theoretical results are demonstrated by the example of a stochastic game modeling data transmission in an ad hoc wireless network with a simple topology. Numerical simulations show the actuality of an application of a game-theoretical model to telecommunication problems because it proposes the method of cost reduction.

References

- Altman, E., El-Azouzi, R. and Jimenez, T. (2003). *Slotted Aloha as a stochastic game with partial information*. in Proceedings of Modeling and Optimization in Mobile, Ad-Hoc and Wireless Networks (WiOpt '03), Sophia Antipolis, France, March 2003.
- Benslama, M., Boucenna, L., Batatia, H. (2013) *Ad Hoc Networks Telecommunications and Game Theory*, Wiley-ISTE.
- Buttayan, L. and Hubaux, J.P. (2003) *Stimulating cooperation in self-organizing mobile ad hoc network*, ACM Journal for Mobile Networks (MONET), **8(5)**, 579–592.
- Dutta, P. (1995). *A folk theorem for stochastic games*. Journal of Economic Theory **66**, 1–32. International Game Theory Review, **4(3)**, 255–264.
- Haurie, A., Krawczyk, J.B., Zaccour, G. (2012) *Games and Dynamic Games*, Scientific World, Singapore.
- Herings, P. J.-J. and Peeters, R. J. A. P. (2004). *Stationary Equilibria in Stochastic Games: Structure, Selection, and Computation*. Journal of Economic Theory, **118(1)**, 32–60.
- Jáskiewicz, A. and Nowak, A. (2015). *On pure stationary almost Markov Nash equilibria in nonzero-sum ARAT stochastic games*. Mathematical Methods of Operations Research, **81(2)**, 169–179.
- Kohlberg, E. and Neyman, A. (2015). *The Cooperative Solution of Stochastic Games*. Working Paper 15-071, Harvard Business School.
- Michiardi, P. and Molva, R. (2003). *A game-theoretical approach to evaluate cooperation enforcement mechanisms in mobile ad hoc networks*, In Proc. WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, Sophia-Antipolis, France.
- Parilina, E. M. (2010). *Cooperative data transmission game in wireless network*. Upravlenie bol'simi sistemami, **31.1**, 93–110 (in Russian).
- Parilina, E. M. (2015). *Stable cooperation in stochastic games*. Automation and remote control, **76(6)**, 1111–1122.
- Parilina, E. and Zaccour, G. (2015a). *Node-consistent core for games played over event trees*. Automatica, **53**, 304–311.
- Parilina, E. and Zaccour, G. (2015b). *Approximated cooperative equilibria for games played over event trees*. Operations Research Letters, **43(5)**, 507–513.
- Petrosyan, L. A. (1977). *Stability of the solutions in differential games with several players*. Vestnik Leningrad. Univ. Mat. Mekh. Astronom, **19(4)**, 46–52 (in Russian).
- Petrosjan, L. A. (2006). *Cooperative Stochastic Games*. Advances in dynamic games, Annals of the International Society of Dynamic Games, Vol. 8, Part VI, 139–146.

- Petrosyan, L. A. (2008). Strategically Supported Cooperation. *International Game Theory Review* **10(4)**, 471–480.
- Petrosjan, L. A. and Baranova, E. M. (2006). *Cooperative Stochastic Games in Stationary Strategies*. Game Theory and Applications, **XI**, 7–17.
- Petrosjan, L. A. and Danilov, N. N. (1979). *Stability of the solutions in nonantagonistic differential games with transferable payoffs*. Vestnik Leningrad. Univ. Mat. Mekh. Astronom., **1**, 52–59 .
- Petrosyan, L. A. and Zenkevich, N. A. (2009). *Principles of dynamic stability*. Upravlenie bol'simi sistemami, **26.1**, 100–120 (in Russian).
- Rosenberg, D., Solan, E. and Vieille, N. (2003). *Stochastic Games with Imperfect Monitoring*. International Journal of Game Theory, **32**, 133–150.
- Sagduyu, Y. E. and Ephremides, A. (2006). *A game-theoretic look at simple relay channel*. Wireless Networks, **12(5)**, 545–560.
- Shapley, L. S. (1953). *Stochastic Games*. Proceedings of National Academy of Sciences of the USA, **39(10)**, 1095–1100.
- Shapley, L. S. (1953). *A Value for n -person Games*. In Contributions to the Theory of Games, vol. II, by H.W. Kuhn and A.W. Tucker, editors. Annals of Mathematical Studies, **28**, 307–317. Princeton University Press.
- Srinivasan, V., Nuggehalli, P., Chiasserini, C. F., Rao, R. R. (2003). *Cooperation in wireless ad hoc networks*, Proc. INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications, San Francisco, CA, USA, 808–817.
- Yeung, D. W. K. (2006). *An irrational-behavior-proof condition in cooperative differential games*. International Game Theory Review, **8**, 739–744.