Trudy Instituta Matematiki i Mekhaniki UrO RAN

Vol. 27 No. 3 2021

AN EFFECTIVE PUNISHMENT FOR AN n-PERSON PRISONER'S DILEMMA ON A NETWORK[1]

A. L. Grinikh, L. A. Petrosyan

The paper considers an n-person prisoner's dilemma game. We present a modification of this model for the network interaction of players. A set of grim trigger strategies is a Nash equilibrium in the repeated n-person prisoner's dilemma on a network, just as in the two-player game. However, even a slight deviation leads to the case where players get low payoffs in perpetuity without the possibility of returning to the Pareto optimal payoffs. A solution to this problem is proposed. The players' payoff functions in a game of an n-person prisoner's dilemma type on a network are described. A strategy involving a punishment on a limited interval of the game is proposed. The number of steps required for an effective punishment is found. An example of a network for this game is given. The number of steps for an effective punishment is found for the given example.

Keywords: prisoner's dilemma, network game, effective punishment.

А. Л. Гриних, Л. А. Петросян. Эффективное наказание в дилемме заключенного для n лиц на сети.

В работе рассматривается дилемма заключенного для n лиц. Приводится модификация данной модели для сетевого взаимодействия игроков. Набор стратегий вечной кары является равновесием по Нэшу в повторяющейся дилемме заключенного для n игроков на сети аналогично случаю двух игроков. Однако даже незначительное отклонение приводит к ситуации, когда игроки получают малые выигрыши в бесконечной перспективе без возможности возвращения к оптимальным по Парето выигрышам. В статье рассматривается вариант решения данной проблемы. Описаны функции выигрыша игроков в игре типа дилемма заключенного для n лиц на сети. Приводится стратегия, предусматривающая наказание на ограниченном интервале игры. Найдено количество шагов, необходимое для эффективного наказания. Приведен пример сети для данной игры. Продемонстрировано нахождение количества шагов для эффективного наказания.

Ключевые слова: дилемма заключенного, игра на сети, эффективное наказание.

MSC: 91A06, 91A43

DOI: 10.21538/0134-4889-2021-27-3-256-262

Introduction

The two-person prisoner's dilemma is a classical example of a problem analized in game theory. This problem exhibits the confrontation between personal gains and social welfare. All players are asked about their joint crime. Each of them has two pure strategies: "to stay silent" or "to defect". Henry Hamburger [1] proposed a modification of this model for a game of n players. His problem statement represents the cumulative effect of the players' influence on each other as follows. Each player gives all other players smaller payoffs by choosing the strategy "to defect" in contrast to the strategy "to stay silent", while his own payoff increases, all things being equal. The table variant of the payoff function for the n-person prisoner's dilemma static game was considered by Straffin [2].

In our previous paper [3], we considered a generalized payoff function for the one-step and multi-step n-person prisoner's dilemma games. This work continues the research started in [3]. Here we will consider a modified version of the payoff function for the network game.

Suppose that each player knows only a part of the details of the other players' involvement in the crime. For example, the organizer knows what he told his associates to do. They, in turn,

[1]The research was supported by the Russian Science Foundation, the grant "Optimal Behavior in Conflict-Controlled Systems" (project no. 17-11-01079).

decided to enlist the help of their fellows. In this case, the judge can increase the prison term of such prisoners, if the closest accomplice betrays them. Distant accomplices have little knowledge about each other and have little influence over each other's prison terms.

Let each node of the network be a player, and let each player have two possible pure strategies: "to stay silent" and "to defect". The edges of the network represent the connections between the players. The power of influence on each other depends on the distance between the players. The more the distance, the less the players can influence each other's payoffs. This effect is achieved by introducing a discount factor.

1. Preliminaries

Let M be a network whose vertices correspond to the players in an n-person prisoner's dilemma game. A path from i to j is a sequence of players connected by edges of the network M. The length of a path is the number of edges in it. If the path from i to j contains the minimal possible number of edges, then it is a shortest path from i to j. The distance between players i and j is the length of a shortest path between the corresponding vertices in M.

Let γ_M be a static noncooperative n-person prisoner's dilemma network game. We denote the set of all players as N. Each of them has two possible pure strategies:

- the strategy "C" means "to cooperate";
- the strategy "D" means "to defect".

Therefore, the set of pure strategies of each player in the static n-person prisoner's dilemma network game can be represented as $X_i = \{C, D\}$, $\forall i \in N$.

Let $x^i_{m,S}$ be the number of players from a set S for which the following conditions are fulfilled:

- they use the strategy "C";
- the distance between them and the player i is m.

The payoff function $h_i(x_1, \ldots, x_i, \ldots, x_n)$, of the player i in the n-person prisoner's dilemma network game depends entirely on his strategy and the number $x^i_{m,N}$:

$$h_i(x_1, \ldots, x_i, \ldots, x_n) = \begin{cases} \left(\sum\limits_{m=0}^{\infty} a_1 \delta^m x^i_{m,N}\right) + b_1, & \text{if } x_i = C; \\ \left(\sum\limits_{m=1}^{\infty} a_2 \delta^m x^i_{m,N}\right) + b_2, & \text{if } x_i = D. \end{cases}$$

Hereinafter, the parameters $a_1 > 0$, $a_2 > 0$, b_1, b_2 and $\delta \in (0;1)$ are the same for all players.

The payoff function of the n-person prisoner's dilemma network game meets the following conditions:

(1) $\left(\sum\limits_{m=0}^{\infty} a_1 \delta^m x^i_{m,N}\right) + b_1 < \left(\sum\limits_{m=1}^{\infty} a_2 \delta^m x^i_{m,N}\right) + b_2, \forall i \in N$; i.e., so the strategy D strictly dominates the strategy C;

(2) $\left(\sum\limits_{m=0}^{\infty} a_1 \delta^m \bar{x}^i_{m,N}\right) + b_1 > b_2$, where $\bar{x}^i_{m,N}$is the number of players at distance m from i from the set N. This inequality shows that the joint "silence" brings a bigger total payoff than the joint "defection";

(3) $\left(\sum\limits_{m=0}^{\infty} a_1 \delta^m x^i_{m,N}\right) + b_1 \geq a_1 + b_1$ and $\left(\sum\limits_{m=1}^{\infty} a_2 \delta^m x^i_{m,N}\right) + b_2 \geq b_2$; i.e., the "defection" of any of the other players decreases the payoff of player i.

2. Problem Statement

Let Γ_M be the infinitely repeated prisoner's dilemma game γ_M on the network M. The payoff of the player i in the game Γ_M is the sum of his payoffs at all of the steps.

It can be seen that a subgame perfect Nash equilibrium for Γ_M is a set of the players' strategies that consists of the repetition of the action "D" in all subgames of Γ_M, since for each step of this game the strategy "D" is strictly dominant.

Definition 1. The *grim trigger strategy* of player i in the n-person prisoner's dilemma game on the network M is the strategy of choosing the action "C" at all steps of the game Γ_M untill the step when one of the other players chooses the action "D". After that step, player i always chooses the action "D" regardless of the actions of all other players.

Lemma 1. *The set of grim trigger strategies of all the players constitutes a Nash equilibrium in the n-person prisoner's dilemma game Γ_M on the network M.*

Proof. Suppose that all the players choose the grim trigger strategies. Then the difference between the payoffs that player i can get at each step using the grim trigger strategy and that he can achive after deviating from grim trigger strategy equals $\sum\limits_{m=0}^{\infty} a_1\delta^m\bar{x}^i_{m,N} + b_1 - b_2$. This difference is greater than 0; so, for the infinite period the player i loses an infinite gain. The future loss is greater than the benefit from the deviation; therefore, the set of grim trigger strategies is a Nash equilibrium in the n-person prisoner's dilemma game Γ_M on the network M. □

However, the set of grim trigger strategies leads to uncertainty as a consequence of a slight deviation from the current strategies (for example, when one of the players uses "D" only once during an infinite period of the game). As a result, all of the players will get smaller payoffs for an infinite period.

3. An Effective Punishment in the Repeated Game Γ_M on the Network M

Let introduce an "effective punishment" for the game Γ_M on the network M.

Definition 2. A *"punishment"* is a choice by the non-deviating players of actions that gives the minimal possible payoffs to the deviating players.

We will call an "effective punishment" such punishment that makes the sum of the payoffs of the deviating player at most such that he can achieve without deviating.

Denote by $\lceil\alpha\rceil$ the ceiling, of α, i.e., the smallest integer that is not less than α.

Theorem 1. *The number of steps that provides an "effective punishment" in the game Γ_M on the network M and makes all the players follow the actions "to stay silent" during all steps of the game is*

$$\bar{k} = \max_{i\in N}\left\lceil \frac{a_2\sum_{m=1}^{\infty}\delta^m\bar{x}^i_{m,N}}{a_1 + b_1 - b_2 + a_1\sum_{m=1}^{\infty}\delta^m\bar{x}^i_{m,N}} \right\rceil.$$

Proof. Let k_i be the minimum number of steps that provides an "effective punishment" for player i. If he decides to use the action "D" at any step of the game Γ_M, all the other players will try to minimize his payoffs at the next k_i steps to punish him for this bechaviour. Since the strategy "to defect", in contrast to the strategy "to stay silent", brings smaller payoffs to all the other players, non-deviating players will choose it as a punishment. Then the number of steps k_i satisfies the inequality

$$k_i b_2 + \sum_{m=1}^{\infty} a_2\delta^m\bar{x}^i_{m,N} < \Big(\sum_{m=0}^{\infty} a_1\delta^m\bar{x}^i_{m,N} + b_1\Big)k_i.$$

The minimum number k_i satisfying this inequality is

$$k_i = \left\lceil \frac{a_2 \sum_{m=1}^{\infty} \delta^m \bar{x}^i_{m,N}}{a_1 + b_1 - b_2 + a_1 \sum_{m=1}^{\infty} \delta^m \bar{x}^i_{m,N}} \right\rceil.$$

Therefore, the number of steps $\bar{k}$ that provides an "effective punishment" in the n-person prisoner's dilemma Γ_M game on the network M is the maximum number of steps that provides an "effective punishment" to the players from the set N:

$$\bar{k} = \max_{i \in N} k_i. \qquad \square$$

Example 1. Consider the three-person prisoner's dilemma Γ_M on the network M shown in Fig. 1, where the second player is a head of the criminal group:

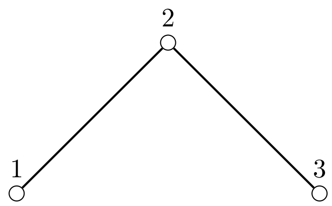

Fig. 1. An example of the network M for three-person prisoner's dilemma game Γ_M.

The one-step payoff functions are

$$h_i(x_1, x_2, x_3) = \begin{cases} \left(\sum\limits_{m=0}^{\infty} 0.8^m x^i_{m,N} \right) + 3, & \text{if } x_i = C, \\ \left(\sum\limits_{m=1}^{\infty} 2 \times 0.8^m x^i_{m,N} \right) + 5, & \text{if } x_i = D. \end{cases}$$

The numbers of steps for an "effective punishment" of each of the players are

$$k_1 = \left\lceil \frac{2 \times (0.8 + 0.64)}{1 + 3 - 5 + 1 \times (0.8 + 0.64)} \right\rceil = 7, \quad k_2 = \left\lceil \frac{2 \times (0.8 \times 2)}{1 + 3 - 5 + 1 \times (0.8 \times 2)} \right\rceil = 6,$$

$$k_3 = \left\lceil \frac{2 \times (0.8 + 0.64)}{1 + 3 - 5 + 1 \times (0.8 + 0.64)} \right\rceil = 7.$$

Therefore, an "effective punishment" in the game Γ_M can be realized in $\bar{k}$ steps, where

$$\bar{k} = \max\{7, 6, 7\} = 7. \tag{1}$$

Thus, a deviation from the strategy "to stay silent" during all steps of the game Γ_M on the network M that maximizes the sum of all players' payoffs can be punished in $\bar{k} = 7$ steps.

4. Pairwise Influence

Let now assume, that only adjacent nodes have influence on each other's payoffs. There can be the following reasons for this assumptions:

- We can count the impact of the players, whose distance to player i is at most 1.
- Players who are unacquainted with each other are not aware of each other's actions and cannot influence payoffs of each other.

- The head of the gang knows everyone, so he can influence everyone but can also be influenced by everyone.

Let us consider a network M_p that consists of nodes and edges representing the players and their relationships in an n-person prisoner's dilemma game. The set of all players is N.

Let γ_p be the static n-person prisoner's dilemma game with the pairwise network influence that is played on the network M_p. The set of player's pure strategies is $X_i = \{C, D\}$. Let x_S^i be the number of the players from the set S adjacent to player i and using the strategy C. Accordingly, the number of all players from the set S adjacent to player i is denoted by $\bar{x}_S^i$.

The payoff function of player i can be represented as

$$h_i^p(x_1, \ldots, x_i, \ldots, x_n) = \begin{cases} a_1 x_N^i + b_1, & \text{if } x_i = C, \\ \\ a_2 x_N^i + b_2, & \text{if } x_i = D. \end{cases} \tag{2}$$

Since we consider the influence of the players at distance 1, there is no need to include a discount factor in this function.

Given that this game is of the prisoner's dilemma type, it meets the following conditions:

(1) $a_1 x_N^i + b_1 < a_2 x_N^i + b_2$, $\forall i \in N$;

(2) $a_1 \bar{x}_N^i + b_1 > b_2$;

(3) $a_1 x_N^i + b_1 > b_1$ and $a_2 x_N^i + b_2 > b_2$.

The game Γ_p is the infinitely repeated prisoner's dilemma game γ_p with pairwise network influence. Since the first condition of the game holds, each of the players would prefer to use the strategy D, but it leads to a smaller total payoff because of the second condition.

Therefore, we should find a behaviour that maximizes the total payoff without incentives to defect. Here, we can also use the grim trigger strategy for a finite number of steps.

Theorem 2. *The number of steps that provides an "effective punishment" in the infinitely repeated prisoner's dilemma game Γ_p with pairwise network influence and makes all the players to follow the actions "C" during all steps of the game is*

$$\bar{k} = \max_{i \in N} \left\lceil \frac{a_2 \bar{x}_N^i + b_2 - a_1 \bar{x}_N^i - b_1}{a_1 \bar{x}_N^i + b_1 - b_2} \right\rceil.$$

Proof. Each of the players tends to use the "D" strategy at each step of the game Γ_p, as this will increase his payoff at this step. In this case, the losses at the next k_i steps should be greater than the difference in the payoffs from choosing the D and C strategies at this step:

$$a_2 \bar{x}_N^i + b_2 + k_i b_2 < \left(a_1 \bar{x}_N^i + b_1\right)(k+1).$$

Then, the minimum number of steps required to effectively punish the i-th player for choosing the strategy "D" will be

$$k_i = \left\lceil \frac{a_2 \bar{x}_N^i + b_2 - a_1 \bar{x}_N^i - b_1}{a_1 \bar{x}_N^i + b_1 - b_2} \right\rceil.$$

The denominator of the fraction in this formula is nonzero, because of condition 2 of the payoff function.

Since changing the strategy from C to D equally changes the payoffs of the other players, we need to find the number of steps sufficient for the punishment of any of the deviating players. Since the payoff functions increase linearly as the number of players who choose the strategy "C" increases, the number of steps that will ensure an effective punishment of any of the players is

$$\bar{k} = \max_{i \in N} k_i. \qquad \square$$

Table. 1. Payoff functions for the n-person prisoner's dilemma game with the pairwise network influence

x_N^i	0		1		2		3		4	
$x_i =$	C	D	C	D	C	D	C	D	C	D
h_1^p	0	1	5	10	10	19	15	28	20	37
h_2^p	0	1	5	10	10	19				
h_3^p	0	1	5	10	10	19				
h_4^p	0	1	5	10	10	19				
h_5^p	0	1	5	10	10	19				

Example 2. Consider the game Γ_p for five players on the network M_p, shown in the Fig. 2:

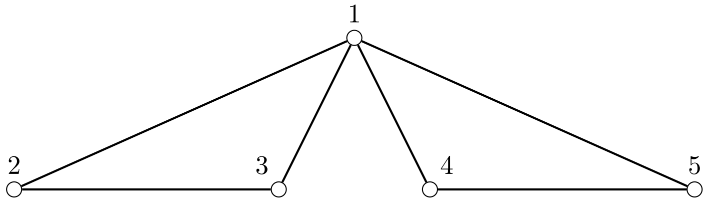

Fig. 2. An example of the network M_p for the five-person prisoner's dilemma Γ_p game.

Set the parameters $a_1 = 5,\ a_2 = 9,\ b_1 = 0$ and $b_2 = 1$.
In this case, the payoff functions for each of the players are represented in Table 1.
Calculate the number of steps for each of the players required for an effective punishment:

$$k_1 = \left\lceil \frac{9 \times 4 + 1 - 5 \times 4 - 0}{5 \times 4 + 0 - 1} \right\rceil = 1,$$

$$k_2 = \left\lceil \frac{9 \times 2 + 1 - 5 \times 2 - 0}{5 \times 2 + 0 - 1} \right\rceil = 1,$$

$$k_3 = \left\lceil \frac{9 \times 2 + 1 - 5 \times 2 - 0}{5 \times 2 + 0 - 1} \right\rceil = 1,$$

$$k_4 = \left\lceil \frac{9 \times 2 + 1 - 5 \times 2 - 0}{5 \times 2 + 0 - 1} \right\rceil = 1,$$

$$k_5 = \left\lceil \frac{9 \times 2 + 1 - 5 \times 2 - 0}{5 \times 2 + 0 - 1} \right\rceil = 1.$$

Thus, the number of steps for an effective punishment in the infinitely repeated prisoner's dilemma game Γ_p with pairwise network influence is

$$\bar{k} = \max_{i \in N} k_i = 1.$$

Conclusion

This paper introduces payoff functions for two types of network games based on the n-person prisoner's dilemma game. In these payoff functions, the incentives for each individual player to use a strategy that leads to the smallest possible total payoff are preserved. Accordingly, the structure of the behaviour is determined that ensures an effective punishment for any of the players who decides to use a strategy harmful to society.

REFERENCES

1. Hamburger H. N-person prisoner's dilemma. *Journal of Mathematical Sociology*, 1973, vol. 3, no. 1, pp. 27–48. doi: 10.1080/0022250X.1973.9989822.
2. Straffin Philip D. Jr. *Game theory and strategy*, Part III, Chapter 21: N-person prisoners dilemma, Ser. Anneli Lax New Mathematical Library, vol. 36, MAA Press, 1993, pp. 139–144.
3. Grinikh A.L. Stochastic n-person Prisoner's dilemma: the time-consistency of core and Shapley value. In: *Contributions to Game Theory and Management*, 2019, vol. 12, pp. 151–158.

Received April 2, 2021
Revised May 2, 2021
Accepted August 2, 2021

Funding Agency: This work was supported by the Russian Science Foundation, the grant "Optimal Behavior in Conflict-Controlled Systems" (project no. 17-11-01079).

Aleksandra Leonidovna Grinikh, PhD student, Saint Petersburg State University, St. Petersburg, 198504 Russia, e-mail: st062331@student.spbu.ru.

Leon Aganesovich Petrosyan, Dr. Phys.-Math. Sci., Prof., Saint Petersburg State University, St. Petersburg, 198504 Russia, e-mail: l.petrosyan@spbu.ru.

Cite this article as: A. L. Grinikh, L. A. Petrosyan. An Effective Punishment for an n-Person Prisoner's Dilemma on a Network, *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, 2021, vol. 27, no. 3, pp. 256–262.