

О. Л. Петросян

РЕШЕНИЕ С ИНФОРМАЦИОННОЙ ДИСКРИМИНАЦИЕЙ В КООПЕРАТИВНЫХ ДИФФЕРЕНЦИАЛЬНЫХ ИГРАХ С БЕСКОНЕЧНОЙ ПРОДОЛЖИТЕЛЬНОСТЬЮ*

Санкт-Петербургский государственный университет, Российская Федерация,
199034, Санкт-Петербург, Университетская наб., 7–9

Предложен новый подход к определению решения дифференциальных игр с бесконечной продолжительностью для случая, когда игроки не имеют точную информацию об игре (уравнения движения, функция выигрыша) на временном интервале, на котором задана игра. В любой момент времени игроки принимают решение, используя информацию на временном интервале с конечной продолжительностью. Информация об игре обновляется в определенные моменты времени и неизвестна заранее. Согласно описанному подходу решение в игре определяется как комбинация решений в усеченных играх. Рассмотрен пример игры управления природными ресурсами, в котором приведено сравнение кооперативной траектории, дележей и процедуры распределения дележа в исходной игре с бесконечной продолжительностью и в игре с представленным подходом. Библиогр. 15 назв. Ил. 3.

Ключевые слова: дифференциальные игры, решение с информационной дискриминацией, процедура распределения дележа, динамическая устойчивость, сильная динамическая устойчивость.

O. L. Petrosian

LOOKING FORWARD APPROACH IN COOPERATIVE DIFFERENTIAL GAMES WITH INFINITE-HORIZON

St. Petersburg State University, 7–9, Universitetskaya nab.,
St. Petersburg, 199034, Russian Federation

A novel approach to definition and computation of a solution of a differential with an infinite-horizon game is presented for the case when players do not have certain information about the game structure on the infinite time interval. At any instant of time players have certain information about motion equations and payoff functions on a subinterval with fixed duration. The information about the game structure updates at fixed instants of time and is completely unknown in advance. A new solution is defined as a recursive combination of sets of imputations in the truncated subgames that are analyzed by the Looking Forward Approach. An example of a resource extraction game illustrates a comparison of cooperative trajectory, imputation, imputation distribution procedure in the original game with infinite-horizon and in the corresponding game with Looking Forward Approach. Refs. 15. Figs 3.

Keywords: differential game, looking forward approach, imputation distribution procedure, time-consistency, strong time-consistency.

1. Введение. Теория кооперативных дифференциальных игр изучает вопросы построения оптимальных решений в процессах со многими участниками. Оптимальное решение включает в себя кооперативную траекторию, стратегии, ее порождаю-

Петросян Ованес Леонович — аспирант; petrosian.ovanes@yandex.ru

Petrosian Ovanes Leonovich — postgraduate student; petrosian.ovanes@yandex.ru

* Работа выполнена при финансовой поддержке Санкт-Петербургского государственного университета (НИР № 9.38.205.2014).

© Санкт-Петербургский государственный университет, 2016

щие, выигрыш вдоль кооперативной траектории, распределение выигрыша между игроками и анализ динамической устойчивости выбранного решения. Проблема динамической неустойчивости вектора Шепли для переговоров была озвучена в работе [1]. Л. А. Петросян впервые математически сформулировал понятие динамической устойчивости [2] и сильной динамической устойчивости [3] решения в кооперативных дифференциальных играх. С целью предотвращения нарушения устойчивости позднее в [4] была предложена схема выплат, получившая название процедуры распределения дележа (ПРД), которая обеспечивает реализацию решения.

В данной работе внимание уделяется специальному классу дифференциальных игр с бесконечной продолжительностью, когда игроки не имеют полной информации об игре на всем временном интервале, на котором определена игра. В каждый момент времени игрокам доступна информация только на фиксированном временном интервале. В определенные моменты времени информация обновляется. Для того чтобы смоделировать поведение игроков в такой игре, предлагается использовать специальный подход, который будем называть подходом решения с информационной дискриминацией. Одна из вариаций подхода описана в [5, 6].

Рассмотрим кооперативную дифференциальную игру, заданную на временном интервале $[t_0, +\infty)$. Предположим, что информация об игре обновляется в моменты времени $t = t_0 + j\Delta t$, $j = 0, \dots, +\infty$, здесь $0 < \Delta t < +\infty$ задает время между моментами обновления информации. В эти моменты игроки получают точную информацию об уравнениях движений и функции выигрыша на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, где $\Delta t < \bar{T} < +\infty$ задает временной горизонт, на котором игрокам известна информация об игре. На интервалах $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$, $j = 0, \dots, +\infty$, строится игра. С помощью уравнений Гамильтона—Якоби—Беллмана [7] можно установить кооперативное поведение (кооперативные стратегии, траекторию) в каждой подобной игре. Вопрос определения оптимального в некотором смысле информационного горизонта \bar{T} для подхода с информационной дискриминацией рассматривался в работе [8].

Характеристическую функцию будем считать также наилучшим ответом коалиции на стратегии, входящие в ситуацию равновесия по Нэшу остальных игроков [9]. Такой подход к получению характеристической функции требует определения ситуации равновесия по Нэшу в дифференциальных играх, вопрос построения равновесия по Нэшу детально описан в [9]. Множество дележей или решение в игре находится для каждого момента времени $t = t_0 + j\Delta t$. Для каждого дележа из выбранного решения определяется ПРД, введенная Л. А. Петросяном в [4]. Некоторые последние публикации на эту тему [9–11]. Для того чтобы определить решение для всей игры, необходимо скомбинировать решения и соответствующие ПРД на временных интервалах $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. Также в работе проведено исследование свойств динамической устойчивости и сильной динамической устойчивости, впервые рассмотренных Л. А. Петросяном в [2, 12].

Для того чтобы продемонстрировать предлагаемый подход, была изучена игра добычи ограниченных ресурсов с бесконечной продолжительностью. Исходная игра описана в [13]. В настоящей работе представлено аналитическое и численное решение игры, которое сравнивается с решением в исходной игре. В качестве дележа в игре используется пропорциональное решение.

Статья имеет следующую структуру. В п. 2 описывается исходная игра, в п. 3 — усеченная игра, с помощью которой можно моделировать поведение игроков с предложенным подходом. В п. 4 дается решение усеченных игр. В п. 5 характеризуется

предлагаемый подход для определения решения во всей игре. В п. 6 подход применяется к дифференциальной игре добычи ограниченного ресурса с бесконечной продолжительностью. В п. 7 приведены выводы.

2. Модель исходной игры. Рассмотрим дифференциальную игру n -лиц с бесконечной продолжительностью $\Gamma(x_0, t_0)$, начинающуюся в позиции $x_0 \in R^m$ в момент времени t_0 . Уравнения движения для этой игры имеют следующий вид:

$$\dot{x} = g(t, x, u), \quad x(t_0) = x_0, \quad (1)$$

где $x \in R^m$; $u = (u_1, \dots, u_n)$. Множество игроков обозначим через $N = \{1, 2, \dots, n\}$. Игрок $i = 1, \dots, n$ выбирает стратегии u_i , как функции текущего состояния и времени со значениями в множестве $U_i \subset \text{Comp}R^k$ (подробней см. в [14]).

Функция выигрыша игрока i

$$K_i(x_0, t_0; u) = \int_{t_0}^{+\infty} h_i(x(\tau), u(\tau))e^{-r(\tau-t_0)}d\tau,$$

здесь $x(\tau)$ — траектория (решение) при заданных уравнениях движения (1) и стратегиях u на бесконечном временном интервале.

Предлагаемый подход предполагает использование игроками усеченной информации об игре. В моменты времени $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ игроки обладают информацией об уравнениях движения и функции выигрыша на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. В моменты времени $t = t_0 + j\Delta t$ информация об игре обновляется и игроки переопределяют свое поведение. Подобные задачи часто появляются в реальной жизни, так как информация о конфликтном процессе на длительном временном интервале не всегда определена и необходимо принимать решения в реальном времени, подстраиваясь под изменения.

3. Понятие усеченной подыгры. В течение первого временного интервала $[t_0, t_0 + \Delta t]$ игроки владеют информацией об игре на временном интервале $[t_0, t_0 + \bar{T}]$. В момент времени $t = t_0 + \Delta t$ информация обновляется, и в течение второго временного интервала $[t_0 + \Delta t, t_0 + 2\Delta t]$ игроки имеют информацию об игре на $[t_0 + \Delta t, t_0 + \Delta t + \bar{T}]$. Для того чтобы учесть это в модели, введем определение усеченной игры. Примем следующее обозначение: $x_{j,0} = x(t_0 + j\Delta t)$.

Определение 1. Пусть $j = 0, \dots, +\infty$. Усеченная игра $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ определена на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ следующим образом. На временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ уравнения движения и функция выигрыша в усеченной игре и исходной игре $\Gamma(x_0, t_0)$ совпадают:

$$\dot{x} = g(t, x, u), \quad x(t_0 + j\Delta t) = x_{j,0},$$

$$K_i^j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u) = \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} h_i(x(\tau), u(\tau))e^{-r(\tau-t_0)}d\tau, \quad (2)$$

где $e^{-r(\tau-t_0)}$ — функция, дисконтирующая выигрыш начиная с момента времени t_0 в игре $\Gamma(x_0, t_0)$.

Предполагается, что выигрыш в игре (в любой усеченной подыгре) рассчитывается от момента времени t_0 ; в формуле (2) дисконтирование выигрыша начинается с момента времени t_0 .

4. Решение кооперативной усеченной подыгры. Рассмотрим усеченную кооперативную подыгру $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ с начальным условием $x(t_0 + j\Delta t) = x_{j,0}$. В кооперативной постановке игрокам необходимо максимизировать суммарный выигрыш

$$\sum_{i \in N} K_i^j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u^j) = \sum_{i \in N} \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} h_i(x(\tau), u(\tau)) e^{-r(\tau - t_0)} d\tau \quad (3)$$

при условии

$$\dot{x} = g(t, x, u), \quad x(t_0 + j\Delta t) = x_{j,0}. \quad (4)$$

Для решения подобной задачи может быть использована система уравнений Гамильтона—Якоби—Беллмана [15, теорема А.1]:

Теорема 1. Пусть существует непрерывно дифференцируемая функция $W^{(j\Delta t)}(t, x) : [t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}] \times R^m \rightarrow R$, удовлетворяющая системе уравнений в частных производных

$$-W_t^{(j\Delta t)}(t, x) = \max_u \left\{ \sum_{i=1}^n h_i(t, x, u) e^{-r(t-t_0)} + W_x^{(j\Delta t)}(t, x) g(t, x, u) \right\} \quad (5)$$

при условии

$$W^{(j\Delta t)}(t_0 + j\Delta t + \bar{T}, x) = 0.$$

Предположим, что максимум в (5) достигается при $u = u^{*j}(t)$. Тогда $u = u^{*j}(t)$ является оптимальным в задаче управления, определяемой (3), (4).

Траекторию, соответствующую $u = u^{*j}(t)$, будем называть кооперативной и обозначать через $x_j^*(t)$.

В соответствии с рассматриваемым подходом в каждый момент времени игрокам доступна ограниченная информация об игре $\Gamma(x_0, t_0)$. Этой информации недостаточно, чтобы определить кооперативное поведение для игроков во всей игре $\Gamma(x_0, t_0)$. Вместо кооперативной траектории в игре $\Gamma(x_0, t_0)$ будем строить условно кооперативную траекторию $\{\hat{x}^*(t)\}_{t=t_0}^{+\infty}$

$$\{\hat{x}^*(t)\}_{t=t_0}^{+\infty} = \begin{cases} x_0^*(t), & t \in [t_0, t_0 + \Delta t], \\ x_1^*(t), & t \in (t_0 + \Delta t, t_0 + 2\Delta t], \\ \dots & \\ x_j^*(t), & t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t], \\ \dots & \end{cases}$$

На временном интервале $[t_0, t_0 + \Delta t]$ траектория $x_0^*(t)$ является кооперативной в усеченной подыгре $\bar{\Gamma}_0^c(x_0, t_0, t_0 + \bar{T})$. В момент времени $t = t_0 + \Delta t$ в позиции $x_0^*(t_0 + \Delta t)$ информация об игре обновляется. На временном интервале $[t_0 + \Delta t, t_0 + 2\Delta t]$ игроки двигаются вдоль кооперативной траектории $x_1^*(t)$ в усеченной подыгре $\bar{\Gamma}_1^c(x_0^*(t_0 + \Delta t), t_0 + \Delta t, t_0 + \Delta t + \bar{T})$. В момент времени $t = t_0 + j\Delta t$ в позиции $x_{j-1}^*(t_0 + j\Delta t)$ информация об игре обновляется. Условно кооперативная траектория $\hat{x}^*(t)$ на временном интервале $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ определена как комбинация частей кооперативных траекторий $x_j^*(t)$ в усеченных подыграх $\bar{\Gamma}_j^c(x_{j-1}^*(t_0 + j\Delta t), t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. Введем следующие обозначения: $x_{j,0}^* = x_{j-1}^*(t_0 + j\Delta t) = x_j^*(t_0 + j\Delta t)$. Тогда усеченная подыгра может быть записана в таком виде: $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$.

Для каждой коалиции $S \subset N$ и усеченной подыгры с номером $j = 0, \dots, +\infty$ найдем значения характеристической функции так, как это сделано в [14]:

$$V_j(S, x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) = \begin{cases} 0, & S = \{\emptyset\}, \\ \max_{u_i, i \in S} \sum_{i \in S} K_i^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u_{j,S}^*, u_{j, N \setminus S}^{NE}), & S \subset N, \\ \max_u \sum_{i=1}^n K_i^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u_j^*), & S = N. \end{cases} \quad (6)$$

В этом подходе предполагается, что фиксируется некоторая ситуация равновесия по Нэшу $u_j^{NE} = (u_{1,j}^{NE}, \dots, u_{n,j}^{NE})$, игроки k , не входящие в коалицию $k \notin S$, используют равновесные по Нэшу стратегии $\{u_{k,j}^{NE}\}$, тогда как игроки из коалиции S максимизируют свой суммарный выигрыш.

Любой дележ $\xi^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ в кооперативной усеченной подыгре $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ должен удовлетворять следующей системе неравенств:

$$\begin{aligned} \xi_i^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) &\geq V_j(\{i\}, x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}), \quad i \in N, \\ \sum_{i \in N} \xi_i^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) &= V_j(N, x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}). \end{aligned}$$

Обозначим множество всевозможных дележей для усеченной подыгры $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ через $E_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$.

Предположим, что для каждой усеченной подыгры выбрано непустое решение

$$W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \subset E_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}).$$

Это может быть С-ядро, НМ-решение, N-ядро или вектор Шепли.

Логично предположить, что распределение суммарного выигрыша между игроками в игре $\Gamma(x_0, t_0)$ вдоль условно кооперативной траектории $\{\hat{x}^*(t)\}_{t=t_0}^{+\infty}$ найдено как комбинация дележей на временных интервалах $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, +\infty$. Эту конструкцию будем называть новой концепцией решения.

Комбинация множеств $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ не позволяет получить решение в игре $\Gamma(x_0, t_0)$ напрямую. Для каждого $j = 0, \dots, +\infty$ решение в усеченной подыгре $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ построено для временного интервала $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. Использование такого решения на временном интервале $[t_0 + j\Delta t, t_0 + (j+1)\Delta t]$ не имеет смысла, так как информация об игре обновляется каждый Δt временной интервал. Необходимая часть решения может быть получена с помощью процедуры распределения дележа для каждой усеченной подыгры. ПРД также обеспечивает свойство динамической устойчивости новой концепции решения и возможность определять решения внутри временного интервала $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$.

5. Концепция решения. Для того чтобы построить предложенное решение в игре $\Gamma(x_0, t_0)$, необходимо установить ПРД для всех усеченных подыгр $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, +\infty$.

Обозначим семейство подыгр для $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ вдоль кооперативной траектории $x_j^*(t)$ через $\bar{\Gamma}_j^c(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$, где $t \in (t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ — начальный момент времени подыгры.

Характеристическая функция вдоль $x_j^*(t)$ в семействе подыгр $\bar{\Gamma}_j^c(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ определена так же, как и в (6). Обозначим через $E_j(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ множество дележей в подыгре $\bar{\Gamma}_j^c(x_j^*(t), t, t_0 + j\Delta t + \bar{T})$ вдоль $x_j^*(t)$.

Предположим, что в каждой усеченной подыгре $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ решение $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \neq \emptyset$ вдоль кооперативной траектории $x_j^*(t)$ выбрано.

Примем, что для любой усеченной подыгры $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ в начальной позиции $x_{j,0}^*$ игроки договорились о выборе дележа:

$$\xi^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \in W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$$

и соответствующего ПРД

$$B_j(t, x_j^*) = [B_1^j(t, x_j^*) \dots B_n^j(t, x_j^*)], \quad t \in (t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}),$$

что гарантирует динамическую устойчивость выбранного дележа [2]:

$$\xi^j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) = \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} B_j(t, x_j^*) e^{-r(t-t_0)} dt,$$

где $e^{-r(t-t_0)}$ — функция, обеспечивающая дисконтирование выигрыша начиная с момента времени t_0 в игре $\Gamma(x_0, t_0)$. ПРД $B_j(t, x_j^*)$ может быть получена путем дифференцирования дележа $\xi_t^j(x_j^*, t, t_0 + j\Delta t + \bar{T})$ [15, лемма 4.1]:

Теорема 2. Если функция $\xi^j(x_j^*, t, t_0 + j\Delta t + \bar{T})$ является непрерывно дифференцируемой по t и x_j^* , тогда

$$B_j(t, x_j^*) = - \left[\xi_t^j(x_j^*, t, t_0 + j\Delta t + \bar{T}) \right] - \left[\xi_{x_j^*}^j(x_j^*, t, T) \right] g_j \left[t, x_j^*, \psi_1^{*j}(\tau, x) \dots \psi_n^{*j}(\tau, x) \right].$$

Новая концепция решения в игре $\Gamma(x_0, t_0)$ состоит из комбинации решений $W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ (соответствующих ПРД) в усеченных подыграх $\bar{\Gamma}_c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $j = 0, \dots, +\infty$. Пусть для каждого дележа $\xi_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}) \in W_j(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ существует ПРД $B_j(t, x_j^*)$. Определим условное ПРД для всей игры $\Gamma(x_0, t_0)$ следующим образом:

$$\hat{B}(t, \hat{x}^*) = B_j(t, x_j^*), \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \quad j = 0, \dots, +\infty. \quad (7)$$

С помощью условного ПРД $\hat{B}(t, \hat{x}^*)$ определим вектор

$$\begin{aligned} \hat{\xi}^j(x_{j,0}^*, t_0 + j\Delta t) &= \int_{t_0 + j\Delta t}^{+\infty} \hat{B}(\tau, \hat{x}^*(\tau)) e^{-r(\tau-t_0)} d\tau = \\ &= \sum_{m=j}^{+\infty} \left[\int_{m\Delta t}^{(m+1)\Delta t} B_m(\tau, x_m^*(\tau)) e^{-r(\tau-t_0)} d\tau \right], \end{aligned} \quad (8)$$

в котором $j = 0, \dots, +\infty$. Через $\hat{W}_j(x_{j,0}^*, t_0 + j\Delta t)$ обозначим множество векторов $\hat{\xi}^j(x_{j,0}^*, t_0 + j\Delta t)$, построенных с помощью (7), (8). В соответствии с новой концепцией

решение в игре $\Gamma(x_0, t_0)$ определено, как $\hat{W} = (\hat{W}_j(x_{j,0}^*, t_0 + j\Delta t))_{j=0}^{+\infty}$. Решение \hat{W} является динамически устойчивым. Оказывается, что \hat{W} обладает и другим свойством.

Определение 2. Решение $W = (W_j(x_{j,0}^*, t_0 + j\Delta t))_{j=0}^{+\infty}$ называется сильно Δt -динамически устойчивым, если для каждого $j = 0, \dots, +\infty$ и каждого $\xi \in W$ соответствующее ПРД $B(t, x^*)$ удовлетворяет условию

$$\int_{t_0}^{t_0+j\Delta t} B(\tau, x^*(\tau))e^{-r(\tau-t_0)} d\tau \oplus W_j(x_{j,0}^*, t_0 + j\Delta t) \subset W_0(x_0, t_0),$$

в котором $a \oplus A = \{a + a' : a' \in A\}$.

Теорема 3. Решение \hat{W} является сильно Δt -динамически устойчивым в игре $\Gamma(x_0, t_0)$.

Доказательство. Пусть $0 \leq j \leq +\infty$ и дележ $\hat{\xi}^0(x_0, t_0) \in \hat{W}_0(x_0, t_0)$ порождает ПРД $\hat{B}(t, \hat{x}^*)$. Тогда для любого $0 \leq k < j$ существует $\xi^k(x_{k,0}^*, t_0 + k\Delta t) \in W_k(x_{k,0}^*, t_0 + k\Delta t)$ с ПРД $B_k(t, x_k^*)$ таким, что

$$\hat{B}(t, \hat{x}^*) = B_k(t, x_k^*), \quad t \in [k\Delta t, (k+1)\Delta t), \quad 0 \leq k \leq j-1.$$

Следовательно,

$$\int_{t_0}^{t_0+j\Delta t} \hat{B}(\tau, \hat{x}^*(\tau))e^{-r(\tau-t_0)} d\tau = \sum_{k=0}^{j-1} \int_{t_0+k\Delta t}^{t_0+(k+1)\Delta t} B_k(t, x_k^*(t))e^{-r(\tau-t_0)} dt.$$

Предположим, что $\xi'' \in W_j(x_{j,0}^*, t_0 + j\Delta t)$. Тогда для любого $j \leq k \leq l-1$ существует $\xi^k(x_{k,0}^*, t_0 + k\Delta t) \in W_k(x_{k,0}^*, t_0 + k\Delta t)$ с ПРД $B_k(t, x_k^*)$ такое, что $\hat{B}(t, \hat{x}^*) = B_k(t, x_k^*)$ для $t \in [t_0 + k\Delta t, t_0 + (k+1)\Delta t)$ и

$$\xi'' = \sum_{m=j}^{l-1} \left[\int_{t_0+m\Delta t}^{t_0+(m+1)\Delta t} B_m(\tau, x_m^*(\tau))e^{-r(\tau-t_0)} d\tau \right].$$

Таким образом,

$$\begin{aligned} & \int_{t_0}^{t_0+j\Delta t} \hat{B}(\tau, \hat{x}^*(\tau))e^{-r(\tau-t_0)} d\tau + \xi'' = \\ & = \sum_{m=0}^{l-1} \left[\int_{t_0+m\Delta t}^{t_0+(m+1)\Delta t} B_m(\tau, x_m^*(\tau))e^{-r(\tau-t_0)} d\tau \right] \in \hat{W}_0(x_0, t_0). \end{aligned}$$

Теорема доказана.

6. Решения с информационной дискриминацией в кооперативной игре добычи ограниченного ресурса с бесконечной продолжительностью. Рассмотрим бесконечную игру добычи ограниченного ресурса. Решение игры в классическом виде представлено в [13]. Применим предлагаемый нами подход для этой игры.

Пусть в игре участвуют два игрока. Уравнения движения, описывающие изменение запаса ресурса $x(t) \in X \subset R$, имеют следующий вид:

$$\dot{x} = a\sqrt{x(t)} - bx(t) - u_1 - u_2, \quad x(t_0) = x_0,$$

где u_i — уровень добычи игрока, $i = 1, 2$.

Запишем функцию выигрыша игрока i :

$$K_i(x_0, t_0; u) = \int_{t_0}^{+\infty} h_i(x(\tau), u(\tau)) e^{-r(\tau-t_0)} d\tau, \quad i = 1, 2,$$

здесь

$$h_i(x(\tau), u(\tau)) = \sqrt{u_i(\tau)} - \frac{c_i}{\sqrt{x(\tau)}} u_i(\tau), \quad i = 1, 2,$$

где r — параметр дисконтирования (дисконтирование выигрыша происходит начиная с момента времени t_0); c_1, c_2 — константы $c_1 \neq c_2$.

Исходная игра $\Gamma(x_0, t_0)$ определена на временном интервале $[t_0, +\infty]$. Предположим, что для любого $t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t]$, $j = 0, \dots, +\infty$, игроки имеют усеченную информацию об игре. Она включает в себя информацию об уравнениях движения и функциях выигрыша на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$. Смоделируем это с помощью усеченной подыгры $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. Уравнения движения и начальные данные имеют такой вид:

$$\dot{x} = a_j\sqrt{x(t)} - b_jx(t) - u_1 - u_2, \quad x(t_0 + j\Delta t) = x_{j,0}, \quad (9)$$

функция выигрыша игрока i :

$$K_i^j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}; u) = \int_{t_0 + j\Delta t}^{t_0 + j\Delta t + \bar{T}} h_i(x(\tau), u(\tau)) e^{-r(\tau-t_0)} d\tau.$$

Усеченная подыгра $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ является конечной дифференциальной игрой, подобная игра была детально изучена в [13, 15]. Стратегии, входящие в ситуацию равновесия по Нэшу в усеченной подыгре $\bar{\Gamma}_j(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, могут быть вычислены так:

$$u_i^j(t, x) = \frac{x}{4[c_i + A_i^j(t)/2]^2}, \quad i = 1, 2.$$

Здесь функции $A_i^j(t)$ определены с помощью системы дифференциальных уравнений

$$\begin{aligned} \dot{A}_i^j(t) &= A_i^j(t) \left[r + \frac{b_j}{2} + \frac{1}{8(c_j + A_{3-i}^j(t)/2)^2} \right] - \frac{1}{4(c_i + A_i^j(t)/2)}, \\ \dot{C}_i^j(t) &= rC_i^j(t) - \frac{a_j}{2} A_i^j(t) \end{aligned}$$

для $i = 1, 2$ с граничными условиями $A_i^j(t_0 + j\Delta t + \bar{T}) = 0$ и $C_i^j(t_0 + j\Delta t + \bar{T}) = 0$.

Выигрыш игрока $i = 1, 2$ в ситуации равновесия по Нэшу [13, 15]

$$V_i^j(t, x) = e^{-r(t-t_0)} \left[A_i^j(t) \sqrt{x} + C_i^j(t) \right], \quad i = 1, 2.$$

Рассмотрим теперь случай, когда игроки договорились кооперироваться в усеченной подыгре $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$. Вычислим максимальный суммарный выигрыш в игре $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ [13, 15]:

$$W^j(t, x) = e^{-r(t-t_0)} [A^j(t)\sqrt{x} + C^j(t)],$$

где функции $A^j(t)$, $C^j(t)$ удовлетворяют следующей системе дифференциальных уравнений:

$$\dot{A}^j(t) = \left[r + \frac{b_j}{2} \right] A^j(t) - \frac{1}{4 \left[c_1 + \frac{A^j(t)}{2} \right]} - \frac{1}{4 \left[c_2 + \frac{A^j(t)}{2} \right]},$$

$$\dot{C}^j(t) = rC^j(t) - \frac{a_j}{2}A^j(t), \quad A^j(t_0 + j\Delta t + \bar{T}) = 0, \quad C^j(t_0 + j\Delta t + \bar{T}) = 0.$$

Кооперативная траектория $x_j^*(t)$ в усеченной подыгре $\bar{\Gamma}_j^c(x_{j,0}, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ определена на временном интервале $[t_0 + j\Delta t, t_0 + j\Delta t + \bar{T}]$ [13, 15]:

$$x_j^*(t) = \varpi_j^2(t_0 + j\Delta t, t) \left[\sqrt{x_{j,0}^*} + \frac{1}{2}a \int_{t_0 + j\Delta t}^t \varpi_j(t_0 + j\Delta t, \tau)^{-1} d\tau \right]^2,$$

$$t \in (t_0 + j\Delta t, t_0 + (j+1)\Delta t],$$

$$\text{здесь } \varpi_j(t_0 + j\Delta t, t) = \exp \int_{t_0 + j\Delta t}^t - \left[\frac{1}{2}b + \frac{1}{8 \left[c_1 + \frac{A^j(\tau)}{2} \right]^2} + \frac{1}{8 \left[c_2 + \frac{A^j(\tau)}{2} \right]^2} \right] d\tau.$$

Начальное положение для кооперативной траектории в каждой усеченной подыгре устанавливается из предыдущей усеченной подыгры: $x_{0,0}^* = x_0$ и $x_{j,0}^* = x_{j-1}^*(t_0 + j\Delta t)$ для $1 \leq j \leq +\infty$. Определим условно кооперативную траекторию $\hat{x}^*(t)$ в игре $\Gamma(x_0, t_0)$:

$$\hat{x}_j^*(t) = x_j^*(t), \quad t \in [t_0 + j\Delta t, t_0 + (j+1)\Delta t], \quad 0 \leq j \leq +\infty.$$

Рассмотрим пример, в котором информация об игре известна на временном интервале с продолжительностью $\bar{T} = 2$ и обновляется каждый $\Delta t = 1$ временной интервал. Зафиксируем следующие параметры для уравнений движений: $a = 5$, $b = 0.3$, для функции выигрыша: $c_1 = 0.15$, $c_2 = 0.65$ и для начальных условий: $t_0 = 0$, $x_0 = 250$. Также положим дисконт-фактор $r = 0.01$.

Условно кооперативная траектория $\hat{x}^*(t)$ составлена из кооперативных траекторий в усеченных подыграх $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$ с уравнениями движения (9). На рис. 1 представлено сравнение условно кооперативной траектории $\hat{x}^*(t)$ и кооперативной траектории $x^*(t)$ в исходной игре с бесконечной продолжительностью $\Gamma(x_0, t_0)$ [13]. Видно, что в случае ограниченной информации выработка ресурсов происходит быстрее, ибо игроки ориентируются на урезанный временной интервал. Ось абсцисс на рис. 1 определяет время t , ось ординат — запас ресурса x .

В качестве дележа в каждой кооперативной усеченной подыгре $\bar{\Gamma}_j^c(x_{j,0}^*, t_0 + j\Delta t, t_0 + j\Delta t + \bar{T})$, $0 \leq j \leq +\infty$, используется пропорциональное решение

$$\xi^{(j)i}(\hat{x}_j^*(t), t, t_0 + j\Delta t + \bar{T}) = \frac{V_i^j(t, \hat{x}_j^*(t))}{\sum_{k=1}^2 V_k^j(t, \hat{x}_j^*(t))} W^j(t, \hat{x}_j^*(t)) =$$

$$= \frac{[A_i^j(t)\sqrt{\hat{x}_j^*(t)} + C_i^j(t)]}{\sum_{k=1}^2 [A_k^j(t)\sqrt{\hat{x}_j^*(t)} + C_k^j(t)]} [A^j(t)\sqrt{\hat{x}_j^*(t)} + C^j(t)].$$

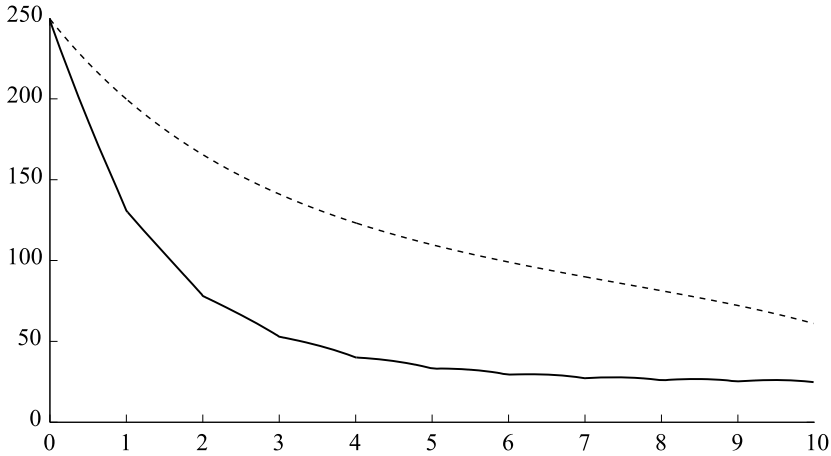


Рис. 1. Условно кооперативная траектория для запаса ресурсов $\hat{x}^*(t)$ (сплошная линия) и кооперативная траектория $x^*(t)$ (пунктирная линия) в исходной игре $\Gamma(x_0, t_0)$ с бесконечной продолжительностью

Соответствующее ПРД обозначим через

$$B_j(t, x_j^*) = [B_1^j(t, x_j^*), B_2^j(t, x_j^*)].$$

Общая формула для ПРД $B_i^j(t, x_j^*)$ в игре добычи ограниченных ресурсов была выведена Д. Янгом и Л. Петросьяном [13]:

$$B_i^j(\tau, x_j^*) = - \left[\xi_t^{(j)i}(x_j^*, t, t_0 + j\Delta t + \bar{T})|_{t=\tau} \right] - \left[\xi_{x_j^*}^{(j)i}(x_j^*, t, t_0 + j\Delta t + \bar{T})|_{t=\tau} \right] \left[a\sqrt{x_j^*} - bx_j^* - \frac{x_j^*}{4[c_1 + A^j(t)]^2} - \frac{x_j^*}{4[c_2 + A^j(t)]^2} \right], \quad i = 1, 2.$$

Условное ПРД $\hat{B}_i(t, \hat{x}^*)$ строится с помощью $B_i^j(\tau, x_j^*)$, $i = 1, 2$ (см. формулу (7)). Условное ПРД, определенное с помощью решений с опережением в игре $\Gamma(x_0, t_0)$, и соответствующее ПРД в исходной игре с бесконечной продолжительностью изображены на рис. 2. Условное ПРД $\hat{B}_i(t, \hat{x}^*)$ имеет ломаный вид, что является результатом обновления информации об игре и пересчета ожидаемого суммарного выигрыша. Ось абсцисс на рис. 2 определяет время t , ось ординат — ПРД.

На рис. 3 видна разница между результирующим дележом $\hat{\xi}(x_0, t_0)$ (8) и дележом в исходной игре $\Gamma(x_0, t_0)$ с бесконечной продолжительностью, а также между соответствующим суммарным выигрышем. Очевидно, что суммарный выигрыш в рассматриваемом подходе будет всегда меньше, чем в исходной игре. Но стоит заметить, что дележ для некоторого игрока может быть больше. Это значит, что некоторым игрокам может быть выгодно, чтобы все игроки имели ограниченную информацию об игре. Ось абсцисс на рис. 3 определяет время t , ось ординат — дележ.

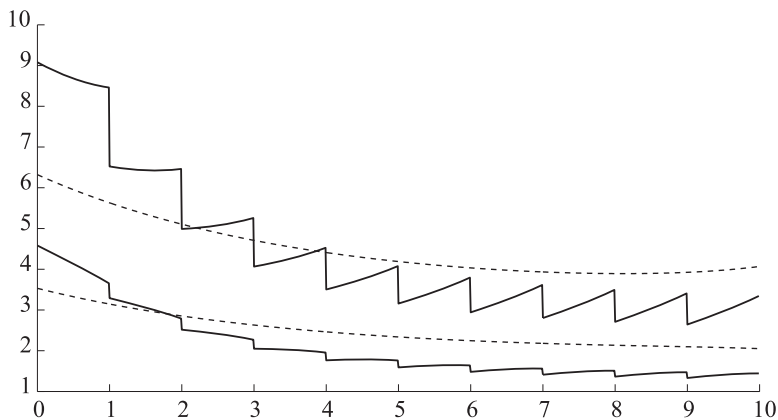


Рис. 2. Условное ПРД $\hat{B}_i(t, \hat{x}^*)$ (сплошные линии) и ПРД $B_i(t, x^*)$ (пунктирные линии) в исходной игре $\Gamma(x_0, t_0)$ с бесконечной продолжительностью ПРД изображено для первого и второго игрока соответственно.

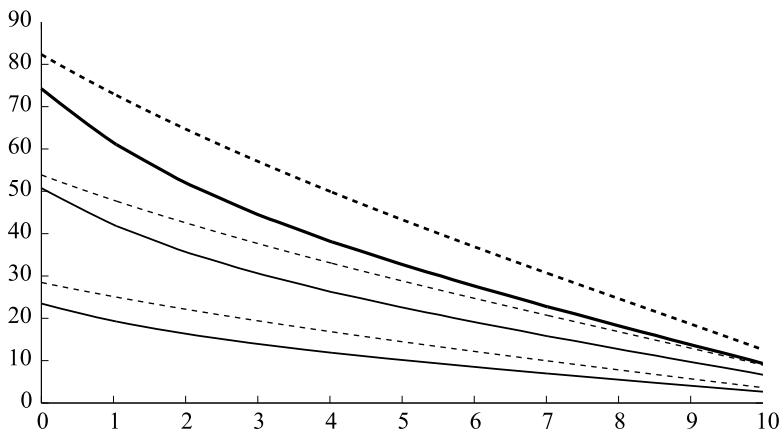


Рис. 3. Дележ $\hat{\xi}(x_0, t_0)$ (сплошные линии), дележ $\xi(x_0, t_0)$ (пунктирные линии) в исходной игре $\Gamma(x_0, t_0)$ с бесконечной продолжительностью и соответствующий суммарный выигрыш игроков (жирные пунктирная и сплошная линии) Дележи изображены для первого и второго игрока соответственно.

7. Заключение. В работе представлен подход для построения решения в реальном времени в кооперативных дифференциальных играх с бесконечной продолжительностью. Игра определена на интервале с бесконечной продолжительностью, разбитом на интервалы конечной длины. Предполагается, что в каждый момент времени игроки имеют информацию об уравнениях движения, функциях выигрыша на конечном временном интервале. В фиксированные моменты времени информация обновляется, таким образом достигается эффект того, что игра решается в реальном времени. Для получения оптимального решения в игре вводится понятие усеченной игры. На его основе определяются условно кооперативная траектория, условная ПРД и концепция решения. Доказывается свойство сильной Δt динамической устойчивости предложенной концепции решения.

Подход проиллюстрирован на бесконечной игре добычи ограниченного ресурса.

Проведено сравнение решения в классической постановке и в постановке с использованием разработанного подхода. Показано, что в случае ограниченной информации выработка ресурсов происходит быстрее, так как игроки ориентируются на узкий временной интервал. Показан ломаный вид условного ПРД по сравнению с ПРД в исходной игре. Также приведено сравнение дележей в исходной игре и в игре с использованием данного подхода. Замечено, что суммарный выигрыш в рассматриваемом подходе будет всегда меньше, чем в исходной игре, но дележ для некоторого игрока может быть больше. Это значит, что некоторым игрокам может быть выгодно, чтобы все игроки имели ограниченную информацию об игре.

Литература

1. Haurie A. A note on nonzero-sum differential games with bargaining solutions // *Journal of Optimization Theory and Applications*. 1976. Vol. 18, issue 1. P. 31–39.
2. Петросян Л. А. Устойчивость решений дифференциальных игр со многими участниками // *Вестн. Ленингр. ун-та. Сер. 1. Математика, Механика, Астрономия*. 1977. Вып. 4. С. 46–52.
3. Петросян Л. А. Сильно динамически устойчивые дифференциальные принципы оптимальности // *Вестн. С.-Петерб. ун-та. Сер. 1. Математика. Механика. Астрономия*. 1993. Вып. 4. С. 35–40.
4. Петросян Л. А., Данилов Н. Н. Устойчивость решений неантагонистических дифференциальных игр с трансферальными выигрышами // *Вестн. Ленингр. ун-та. Сер. 1. Математика. Механика. Астрономия*. 1979. Вып. 1. С. 52–59.
5. Petrosian O. L. Looking forward approach in cooperative differential games // *Intern. Game Theory Review*. 2016. Vol. 18, issue 2. P. 1–14. DOI: 10.1142/S0219198916400077.
6. Petrosian O. L., Barabanov A. E. Looking forward approach in cooperative differential games with uncertain-stochastic dynamics // *Journal of Optimization Theory and Applications*. 2016. P. 1–20. DOI: 10.1007/s10957-016-1009-8.
7. Bellman R. *Dynamic programming*. Princeton: Princeton University Press, 1957. 550 p.
8. Gromova E. V., Petrosian O. L. Control of informational horizon for cooperative differential game of pollution control // 2016 Intern. conference stability and oscillations of nonlinear control systems (Pyatnitskiy's conference). 2016. DOI: 10.1109/STAB.2016.7541187.
9. Petrosyan L. A., Zaccour G. Time-consistent Shapley value allocation of pollution cost reduction // *J. econ. dyn. control*. 2003. Vol. 27, issue 3. P. 381–398.
10. Petrosyan L. A., Yeung D. W. K. Dynamically stable solutions in randomly-furcating differential games // *Trans. Steklov inst. math.* 2006. Vol. 253, issue 1. P. 208–220.
11. Jorgensen S., Martin-Herran G., Zaccour G. Agreeability and time consistency in linear-state differential games // *Journal of Optimization Theory and Applications*. 2003. Vol. 119, issue 1. P. 49–63.
12. Petrosjan L. A. Strongly time-consistent differential optimality principles // *Vestnik of Saint Petersburg University. Series 1. Mathematics. Mechanics. Astronomy*. 1993. Issue 4. P. 40–46.
13. Yeung D. W. K., Petrosyan L. A. *Subgame-consistent economic optimization*. New York: Springer, 2012. 395 p.
14. Basar T., Olsder G. J. *Dynamic noncooperative game theory*. London: Academic Press, 1995. 535 p.
15. Jorgensen S., Yeung D. W. K. Inter- and intergenerational renewable resource extraction // *Annals of Operations Research*. 1999. Vol. 88, issue 0. P. 275–289.

Для цитирования: Петросян О. Л. Решение с информационной дискриминацией в кооперативных дифференциальных играх с бесконечной продолжительностью // *Вестник Санкт-Петербургского университета. Сер. 10. Прикладная математика. Информатика. Процессы управления*. 2016. Вып. 4. С. 18–30. DOI: 10.21638/11701/spbu10.2016.402

References

1. Haurie A. A note on nonzero-sum differential games with bargaining solutions. *Journal of Optimization Theory and Applications*, 1976, vol. 18, issue 1, pp. 31–39.
2. Petrosyan L. A. Ustojchivost' reshenij differencial'nyh igr so mnogimi uchastnikami [Time-consistency of solutions in multi-player differential games]. *Vestnik of Leningrad State University. Series 1. Mathematics. Mechanics. Astronomy*, 1977, issue 4, pp. 46–52. (In Russian).
3. Petrosyan L. A. Sil'no dinamicheski ustojchivye differencial'nye principy optimal'nosti [Strongly

time-consistent differential optimality principles]. *Vestnik of Leningrad State University. Series 1. Mathematics. Mechanics. Astronomy*, 1993, issue 4, pp. 40–46. (In Russian).

4. Petrosyan L. A., Danilov N. N. Ustojchivost' reshenij neantagonisticheskikh differentsial'nyh igr s transferabel'nymi vyigryshami [Stability of solutions in non-zero sum differential games with transferable payoffs]. *Vestnik of Leningrad State University. Series 1. Mathematics. Mechanics. Astronomy*, 1979, issue 1, pp. 52–59. (In Russian).

5. Petrosian O. L. Looking forward approach in cooperative differential games. *Intern. Game Theory Review*, 2016, vol. 18, issue 2, pp. 1–14. DOI: 10.1142/S0219198916400077.

6. Petrosian O. L., Barabanov A. E. Looking forward approach in cooperative differential games with uncertain-stochastic dynamics. *Journal of Optimization Theory and Applications*, 2016, pp. 1–20. DOI: 10.1007/s10957-016-1009-8.

7. Bellman R. *Dynamic programming*. Princeton, Princeton University Press, 1957, 550 p.

8. Gromova E. V., Petrosian O. L. Control of informational horizon for cooperative differential game of pollution control. *Intern. conference stability and oscillations of nonlinear control systems (Pyatnitskiy's conference)*, 2016. DOI: 10.1109/STAB.2016.7541187.

9. Petrosyan L. A., Zaccour G. Time-consistent Shapley value allocation of pollution cost reduction. *J. econ. dyn. control*, 2003, vol. 27, issue 3, pp. 381–398.

10. Petrosyan L. A., Yeung D. W. K. Dynamically stable solutions in randomly-furcating differential games. *Trans. Steklov inst. math.*, 2006, vol. 253, issue 1, pp. 208–220.

11. Jorgensen S., Martin-Herran G., Zaccour G. Agreeability and time consistency in linear-state differential games. *Journal of Optimization Theory and Applications*, 2003, vol. 119, issue 1, pp. 49–63.

12. Petrosjan L. A. Strongly time-consistent differential optimality principles. *Vestnik of Saint Petersburg University. Series 1. Mathematics. Mechanics. Astronomy*, 1993, issue 4, pp. 40–46. (In Russian).

13. Yeung D. W. K., Petrosyan L. A. *Subgame-consistent economic optimization*. New York, Springer Press, 2012, 395 p.

14. Basar T., Olsder G. J. *Dynamic noncooperative game theory*. London, Academic Press, 1995, 535 p.

15. Jorgensen S., Yeung D. W. K. Inter- and intergenerational renewable resource extraction. *Annals of Operations Research*, 1999, vol. 88, issue 0, pp. 275–289.

For citation: Petrosian O. L. Looking forward approach in cooperative differential games with infinite-horizon. *Vestnik of Saint Petersburg University. Series 10. Applied mathematics. Computer science. Control processes*, 2016, issue 4, pp. 18–30. DOI: 10.21638/11701/spbu10.2016.402

Статья рекомендована к печати проф. Л. А. Петросьяном.

Статья поступила в редакцию 7 мая 2016 г.

Статья принята к печати 29 сентября 2016 г.