УДК 517.977.8

УСЛОВИЯ ОПТИМАЛЬНОСТИ ДЛЯ ДИФФЕРЕНЦИАЛЬНЫХ ИГР С НЕПРЕРЫВНЫМ ОБНОЛЕНИЕМ ИНФОРМАЦИИ

О.Л. Петросян

Санкт-Петербургский государственный университет

Россия, 198504, Санкт-Петербург, Университетский просп., 35

E-mail: petrosian.ovanes@yandex.ru

Ключевые слова: дифференциальные игры с непрерывным обновлением информации, равновесие по Нэшу, уравнения Гамильтона-Якоби-Беллмана.

Аннотация: В настоящей работе исследуется новый класс дифферециальных игр с непрерывным обновлением информации. Предполагается, что в каждый текущий момент времени игроки имеют информацию на замкнутом временно интервале. Однако, с течением времени информация об игре обновляется, а именно происходит непрерывное смещение временного интервала, который определяет информацию доступную игрокам. Под информацией про игру понимается информация об уравнениях движения и функциях выигрыша игроков. Для подобного класса игр прямое применение классических подходов для определения таких принципов оптимальности как равновесие по Нэшу не представляется возможным. Предметом работы является вопрос построения принципов оптимальности в подобном классе дифференциальных игр с помощью модернизированных уравнений Гамильтона-Якоби-Беллмана.

1. Введение

Основные модели, рассматриваемые в классической теории дифференциальных игр, связаны с задачами, определенными на фиксированном интервале времени [4], задачами заданными на бесконечном интервале времени с дисконтированием [1], задачами, определенными на случайном интервале времени [3]. Одна из первых работ в теории дифференциальных игр посвящена дифференциальной игре преследования [8]. Во всех вышеупомянутых моделях и подходах предполагается, что игроки в начале игры имеют всю информацию об уравнениях движения и о функциях выигрыша. Однако, эти подходы не учитывают тот факт, что во многих реальных конфликтно-управляемых процессах игроки в начальный момент времени не имеют всю информации об игре. Из-за этого классические подходы для поиска в некотором смысле оптимальных стратегий (например, равновесия по Нэшу), такие как уравнение Гамильтона-Якоби-Беллмана [2] или принцип максимума Понтрягина [5], нельзя напрямую использовать для построения достаточно большого диапазона реальных теоретико-игровых моделей.

Большинство реальных конфликтно-управляемых процессов непрерывно развиваются во времени, а их участники постоянно получают обновленную информацию и адаптируются. В данной работе предложен подход для построения оптимальных стратегий и изучен вопрос построения оптимального управления для игровых моделей с непрерывным обновлением информации. В игровых моделях с непрерывным обновлением предполагается, что игроки

1. в текущий момент времени t имеют информацию об уравнениях движения и функциях выигрыша на усеченном временном интервале $[t, t+\overline{T}]$ с величиной информационного горизонта равной $\overline{T}$,

2. непрерывно получают обновленную информацию об уравнениях движения и функциях выигрыша с течением времени $t \in [t_0, +\infty)$, и как результат постоянно адаптируются к обновленной информации.

Очевидно, что понятие оптимальности для класса дифференциальных игр с непрерывным обновлением информации из-за отсутствия фундаментального подхода не является тривиальным. До настоящего времени изучался только класс игр с динамическим обновлением информации. В работах [6], [7] авторы заложили основу для дальнейшего изучения класса игр с динамическим обновлением. Там предполагается, что информация об уравнениях движения и функциях выигрыша обновляется в дискретные моменты времени, а интервал, на котором игроки имеют информацию, определяется значением информационного горизонта.

В данной работе определена игровая модель с непрерывным обновлением и представлены необходимые и достаточные условия оптимальности в виде уравнений Гамильтона-Якоби-Беллмана для модифицированного равновесия по Нэшу в позиционных стратегиях.

2. Модель дифференциальной игры с непрерывным обновлением информации.

Дифференциальная игра с непрерывным обновлением информации. Рассмотрим дифференциальную игру n-лиц $\Gamma(x_0, t_0, \overline{T})$, определенную на интервале $[t_0, t_0+\overline{T}]$, где $0 < \overline{T} < +\infty$.

Уравнения движения имеют вид

$$\begin{aligned} &\dot{x}_{t_0}(s) = f(s, x_{t_0}, u^{t_0}), \\ &x_{t_0}(t_0) = x_0, \\ &x_{t_0} \in \mathbb{R}^l, \ u^{t_0} = (u_1^{t_0}, \ldots, u_n^{t_0}), \ u_i^{t_0} = u_i^{t_0}(s, x) \in U_i \subset \mathrm{comp}\mathbb{R}^k. \end{aligned} \tag{1}$$

Функция выигрыша игрока $i \in N$ определена в виде

$$K_i^{t_0}(x_0, t_0; u^{t_0}) = \int\limits_{t_0}^{t_0+\overline{T}} g^i[s, x_{t_0}(s), u^{t_0}(s, x)]ds, \ i \in N, \tag{2}$$

где $x_{t_0}(s)$, $u^{t_0}(s, x)$ - траектория и стратегии в игре $\Gamma(x_0, t_0, \overline{T})$.

Подыгра дифференциальной игры с непрерывным обновлением информации. Рассмотрим дифференциальную игру n-лиц $\Gamma(x, t, \overline{T})$, $t \in [t_0, +\infty)$ определенную на интервале $[t, t+\overline{T}]$, где $0 < \overline{T} < +\infty$.

Уравнения движения для подыгры $\Gamma(x, t, \overline{T})$ имеют вид

$$
(3) \qquad \begin{aligned}
&\dot{x}_t(s) = f(s, x_t, u^t),\\
&x_t(t) = x,\\
&x_t \in \mathbb{R}^l,\ u^t = (u_1^t, \ldots, u_n^t),\ u_i^t = u_i^t(s, x) \in U_i \subset \mathrm{comp}\mathbb{R}^k,\ s \in [t, t + \overline{T}].
\end{aligned}
$$

Функция выигрыша игрока $i \in N$ для подыгры $\Gamma(x, t, \overline{T})$ имеет вид

$$
(4) \qquad K_i^t(x, t; u^t) = \int\limits_t^{t+\overline{T}} g^i[s, x_t(s), u^t(s, x)]ds,\ i \in N,
$$

где $x_t(s)$, $u^t(s, x)$ - траектория и стратегии в игре $\Gamma(x, t, \overline{T})$.

Дифференциальная игра с непрерывным обновлением информации развивается по следующему правилу:

Текущее время $t \in [t_0, +\infty)$ развивается непрерывно, и как результат игроки непрерывно получают обновленную информацию об уравнениях движения и функциях выигрыша в игре $\Gamma(x, t, \overline{T})$.

Набор стратегий $u(t, x)$ в дифференциальной игре с непрерывным обновлением информации имеет вид:

$$
(5) \qquad u(t, x) = u^t(t, x),\ t \in [t_0, +\infty),
$$

где $u^t(s, x)$, $s \in [t, t + \overline{T}]$ - стратегии определенные для подыгры $\Gamma(x, t, \overline{T})$.

Траектория $x(t)$ в дифференциальной игре с непрерывным обновлением информации определяется в соответствии с (3) при фиксированных $u(t, x)$ (5).

Существенная разница между игровой моделью с непрерывным обновлением и классической дифференциальной игрой $\Gamma(x_0, T - t_0)$ с предписанной продолжительностью заключается в том, что игроки в начальной игре ориентируются на выигрыши, которые они в конечном итоге получат на интервале $[t_0, T]$, но в случае игры с непрерывным обновлением они в момент времени t ориентируются на ожидаемые выигрыши (4), которые рассчитаны на основе информации о процессе на интервале $[t, t + \overline{T}]$.

3. Равновесие по Нэшу в игре с непрерывным обновлением инфомрации.

В рамках непрерывно обновляющейся информации для описанного класса дифференциальных игр интересным является вопрос моделирования поведения игроков. Для этого воспользуемся понятием равновесия по Нэшу в позиционных стратегиях. Однако, для класса дифференциальных игр с непрерывным обновлением информации хотелось бы, чтобы оно имело следующий вид:

- для каждого фиксированного $t \in [t_0, T]$, $u^{NE}(t,x) = (u_1^{NE}(t,x), \ldots, u_n^{NE}(t,x))$ совпадает с равновесием по Нэшу в игре (3),(4) заданной на интервале $[t, t+\overline{T}]$ в момент t.

Однако, напрямую применение классических подходов для определения равновесия по Нэшу в позиционных стратегиях (РН) не представляется возможным, покажем это рассмотрев два интервала $[t, t+\overline{T}]$ и $[t+\epsilon, t+\overline{T}+\epsilon]$, $\epsilon << \overline{T}$, в соответствии с постановкой задачи:

- $u^{NE}(t,x)$ должно совпадать с РН в задаче на интервале $[t, t+\overline{T}]$ в момент t,

- $u^{NE}(t+\epsilon, x)$ должно совпадать с РН в задаче на интервале $[t+\epsilon, t+\epsilon+\overline{T}]$ в момент $t+\epsilon$.

Для построения подобных стратегий рассмотрим в качестве принципа оптимальности понятие обобщенного равновесия по Нэшу в позиционных стратегиях

$$\widetilde{u}^{NE}(t,s,x) = (\widetilde{u}_1^{NE}(t,s,x), \ldots, \widetilde{u}_n^{NE}(t,s,x)),\ t \in [t_0, T],\ s \in [t, t+\overline{T}], \tag{6}$$

которое мы в дальнейшем используем для построения желаемых стратегий $u^{NE}(t,x)$.

Определение. *Набор стратегий* $\widetilde{u}^{NE}(t,s,x) = (\widetilde{u}_1^{NE}(t,s,x), \ldots, \widetilde{u}_n^{NE}(t,s,x))$ *является обобщенным равновесием по Нэшу в игре с непрерывным обновлением информации, если для любого фиксированного* $t \in [t_0, +\infty)$ *набор стратегий* $\widetilde{u}^{NE}(t,s,x)$ *является равновесием по Нэшу в позиционных стратегиях в игре* $\Gamma(x, t, \overline{T})$.

На основе обобщенного равновесия по Нэшу $\widetilde{u}^{NE}(t,s,x)$ представляется возможным определить искомый набор стратегий $u^{NE}(t,x)$ для класса дифференциальных игр с непрерывным обнолвением информации по следующему правилу:

$$u^{NE}(t,x) = \widetilde{u}^{NE}(t,s,x)|_{s=t} = (\widetilde{u}_1^{NE}(t,s,x)|_{s=t}, \ldots, \widetilde{u}_n^{NE}(t,s,x)|_{s=t}),\ t \in [t_0, +\infty). \tag{7}$$

4. Уравнения Гамильтона-Якоби-Беллмана с непрерывным обновлением информации.

Для того, чтобы найти набор стратегий вида $u^{NE}(t,x)$ необходимо на первом шаге определить обобщенное равновесие по Нэшу в позиционных стратегиях $\widetilde{u}^{NE}(t,s,x)$ в игре с непрерывным обновлением информации $\Gamma(x_0, t_0, \overline{T})$. Для этого мы будем использовать модернизированный подход динамического программирования. В рамках этого подхода функция Беллмана $V^i(t,s,x)$ определяется как выигрыш игрока i в равновесии по Нэшу в позиционных стратегиях $\widetilde{u}^{NE}(t,s,x)$ в подыгре начинающейся в момент времени s в состоянии x в игре определенной на интервале $[t, t+\overline{T}]$:

$$V^i(t,s,x) = \int_s^{t+\overline{T}} g^i[\tau, x_t(\tau), \widetilde{u}^{NE}(t,\tau,x)]d\tau,\ i \in N \tag{8}$$

при условии

$$\dot{x}(\tau) = f(\tau, x, u),$$
$$x(s) = x. \tag{9}$$

Справедлива следующая теорема:

Теорема 1. $\widetilde{u}^{NE}(t,s,x)$ *является обобщенным равновесием по Нэшу в позиционных стратегиях в дифференциальной игре с непрерывным обновлением информации* $\Gamma(x_0, t_0, \overline{T})$ *тогда и только тогда, когда существуют непрерывно дифференцируемые функции* $V^i(t,s,x) : [t_0, +\infty) \times [t, t+\overline{T}] \times R \to R$, $i \in N$, *удовлетворяющие следующей системе уравнений:*

$$\begin{aligned} -V^i_s(t,s,x) = \max_{\phi_i} \left\{ g^i(s,x,\widetilde{u}^{NE}_{-i}) + V^i_x(t,s,x) f(s,x,\widetilde{u}^{NE}_{-i}) \right\} = \\ = g^i(s,x,\widetilde{u}^{NE}) + V^i_x(t,s,x) f(s,x,\widetilde{u}^{NE}), \\ V^i(t, t+\overline{T}, x) = 0, \quad i \in N, \end{aligned} \tag{10}$$

где $\widetilde{u}^{NE}_{-i} = (\widetilde{u}^{NE}_1, \ldots, \phi_i, \ldots, \widetilde{u}^{NE}_n)$.

В случае, если удается получить обощенное равновесие по Нэшу $\widetilde{u}^{NE}(t,s,x)$ с помощью уравнения (10), то используя процедуру (7) мы получаем искомый набор стратегий $u^{NE}(t,x)$.

5. Заключение

Целью данной работы является построение условий оптимальности для нового класса дифференциальных игр с непрерывным обновлением информации. Важным является вопрос построения уравнений Гамильтона-Якоби-Беллмана и Принципа максимума Понтрягина для подобных задач. Вид уравнения Гамильтона-Якоби-Беллмана представлен в работе выше, а условия Принципа максимума Понтрягина еще предстоит получить.

Исследование выполнено за счет гранта Российского научного фонда (проект No 18-71-00081)

Список литературы

1. Basar T., Olsder G.J. Dynamic noncooperative game theory. Academic Press, 1995.
2. Bellman R. Dynamic Programming. Princeton University Press, 1957.
3. Shevkoplyas E.V. Optimal Solutions in Differential Games with Random Duration // Journal of Mathematical Sciences. 2014. Vol 199, No. 6. P. 715-722.
4. Kleimenov A.F. Non-antagonistic positional differential games. Science, 1993.
5. Pontryagin L.S., Boltyansky V.G., Gamkrelidze R.V., Mishchenko E.F. Mathematical theory of optimal processes. Science, 1976.
6. Petrosian O.L. Looking Forward Approach in Cooperative Differential Games // International Game Theory Review. 2016. Vol. 18. P. 1-14.
7. Petrosian O.L., Barabanov A.E. Looking Forward Approach in Cooperative Differential Games with Uncertain-Stochastic Dynamics // Journal of Optimization Theory and Applications. 2017. Vol. 172. P. 328-347.
8. Petrosyan L.A., Murzov N.V. Game-theoretic Problems in Mechanics // Lithuanian Mathematical Collection. 1966. Vol. 3. P. 423-433.