Прикладная математика и механика. Том 34, 1970

ДИФФЕРЕНЦИАЛЬНЫЕ ИГРЫ С ЗАПАЗДЫВАНИЕМ ИНФОРМАЦИИ

Б. Н. Соколов, Ф. Л. Черноусько

(Москва)

Рассматриваются дифференциальные игры двух игроков в предположении, что одна из сторон получает информацию о фазовых координатах другой стороны с запаздыванием. В прикладных задачах запаздывание информации обусловлено временем, необходимым для получения и обработки данных измерения. Наличие запаздывания информации отличает рассмотренные в работе задачи от обычных задач теории дифференциальных игр (см., например, [1-3]). В работе показано (см. также [4]), что при некоторых условиях общего характера каждая дифференциальная игра с запаздыванием информации эквивалентна некоторой дифференциальной игре без запаздывания. Это дает возможность применить известные методы теории дифференциальных игр для решения игр с западыванием информации. Рассмотрен ряд конкретных задач преследования, в которых одна из сторон получает информацию с запаздыванием. Получено и проанализировано решение этих задач, определены условия возможности захвата.

§ 1. Постановка задачи. Рассматривается дифференциальная игра двух управляемых систем, называемых сторонами (или игроками) P и E. Векторы фазовых координат этих систем обозначим соответственно через $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_s)$, а векторы управляющих функций — через $u = (u_1, \ldots, u_m)$ и $v = (v_1, \ldots, v_k)$. Размерности n, s, m, k этих векторов произвольны.

Уравнения движения систем имеют вид

$$dx \,/\, dt = f(x, u, t), \qquad dy \,/\, dt = g(y, v, t) \tag{1.1}$$

Здесь t — время, $f = (f_1, \ldots, f_n)$ и $g = (g_1, \ldots, g_s)$ — заданные вектор-функции. Момент начала игры t_0 предполагаем заданным, а момент T ее окончания определяется условием вида

$$h(x(T), \; y(T), \; T) = 0 \tag{1.2}$$

Здесь h — заданная скалярная функция, T — наименьший корень уравнения (1.2), для которого $T > t_0$. Не нарушая общности, можно принять, что $h > 0$ при $t < T$; в противном случае просто умножим функцию h на -1. В частности, если время окончания игры задано и равно T_0, то $h = T_0 - t$.

На управляющие функции наложены ограничения

$$u(t) \in U, \quad v(t) \in V \qquad (t \geqslant t_0) \tag{1.3}$$

Здесь U, V — замкнутые множества в m- и r-мерном пространствах соответственно. Функционал (плата) задан в форме

$$F = F(x(T), \; y(T), \; T) \tag{1.4}$$

где $F(x, y, t)$ — известная скалярная функция. Будем считать, не нарушая общности, что сторона P стремится минимизировать, а E — максимизировать функционал F. В частности, если платой служит длительность игры (в задачах преследования), то $F = t$ или $F = -t$ в зависимости от того, является ли P преследующей или преследуемой стороной. Введенные условия и ограничения обычны в теории дифференциальных игр.

Примем еще следующее предположение об информированности сторон. Пусть в каждый момент t стороне P становится известным текущее значение своего фазового вектора $x(t)$, а также значение фазового вектора стороны E в момент $t - \tau$, т. е. вектор $y(t - \tau)$. Здесь постоянная $\tau > 0$ характеризует запаздывание информации для стороны P и равна времени, необходимому этой стороне для получения и обработки данных измерений. Запаздывание естественно считать достаточно малым (меньше длительности процесса), т. е. $\tau < T - t_0$. Кроме того, стороне P до начала игры известны функции f, g, h, F и множества U, V в соотношениях (1.1) — (1.4), т. е. правила и цель игры. Задачу будем решать для стороны P, т. е. искать оптимальное управление u стороны P в зависимости от времени t на интервале $[t_0, T]$ и от текущих данных измерений — от векторов $x(t)$ и $y(t - \tau)$. Относительно информированности стороны E можно не делать специальных предположений и считать, что E действует наихудшим для P образом. Тем самым будет определено наименьшее гарантированное значение функционала F.

Под начальным моментом t_0 будем понимать момент первого поступления информации, и начальные условия зададим в виде

$$x(t_0) = x^\circ, \qquad y(t_0 - \tau) = y^\circ \tag{1.5}$$

Управляемое движение стороны P начинается в момент t_0, а при $t < t_0$, до поступления информации, управление стороны P должно выбираться из каких-либо априорных соображений (движение при $t < t_0$ можно считать неуправляемым).

§ 2. Эквивалентность игр с запаздыванием и игр без запаздывания информации. Введем обозначение $\eta(t) = y(t - \tau)$ и перепишем уравнения (1.1) и начальные условия (1.5) в виде

$$dx / dt = f(x, u, t), \qquad d\eta / dt = g(\eta, v, t - \tau) \tag{2.1}$$

$$x(t_0) = x^0, \qquad \eta(t_0) = y^0$$

Обозначим через $D(y, t)$ область достижимости для стороны E за время τ при начальном условии $y(t) = y$. Иными словами, $D(y, t)$ — это множество векторов $y(t + \tau)$, которое получится при условии $y(t) = y$ и всевозможных допустимых управлениях $v \in V$ на интервале $[t, t + \tau]$, если функции y, v связаны вторым из уравнений (1.1). Запишем условие окончания игры и выражение для гарантированного значения минимизируемого функционала F, используя введенные обозначения. При этом ограничимся рассмотрением трех важных частных случаев.

1°. Пусть время окончания игры фиксировано и равно T_0, тогда функция h из (1.2) равна $h = T_0 - t$. Гарантированное значение функционала F при заданных $x(T_0)$, $\eta(T_0)$ определится (с расчетом на наихудший случай) путем максимизации функции F из (1.4) по допустимым значениям $y(T_0)$. Получим

$$\begin{gathered} F_0 = F_0(x(T_0),\ \eta(T_0),\ T_0) = \\ = \max_y F(x(T_0),\ y,\ T_0),\ y \in D(\eta(T_0),\ T_0 - \tau) \end{gathered} \tag{2.2}$$

2°. Пусть теперь функция h произвольна (с условием, что $h > 0$ при $t < T$), а функция F из (1.4) равна $F = t$. Это имеет место для игры преследования, когда P — преследователь. Тогда наихудший для P момент окончания игры будет определяться из условия

$$\begin{gathered} h_0(x(t),\ \eta(t),\ t) = \max_y h(x(t),\ y,\ t) = 0 \\ y \in D(\eta(t),\ t - \tau) \end{gathered} \tag{2.3}$$

3°. Пусть функция h по-прежнему произвольна, а $F = -t$, т. е. P является преследуемой стороной. Тогда вместо (2.3) получим

$$\begin{gathered} h_0(x(t),\ \eta(t),\ t) = \min_y h(x(t),\ y,\ t) = 0 \\ y \in D(\eta(t),\ t - \tau) \end{gathered} \tag{2.4}$$

Во всех рассмотренных случаях условие окончания игры и гарантированное значение функционала имеют вид

$$\begin{gathered} h_0(x(T),\ \eta(T),\ T) = 0 \\ F_0 = F_0(x(T),\ \eta(T),\ T) \end{gathered} \tag{2.5}$$

Здесь функция h_0 задается равенствами $h_0 = T_0 - t$, (2.3) и (2.4), а функция F_0 — равенствами (2.2), $F_0 = t$ и $F_0 = -t$ соответственно в трех рассмотренных случаях. В работе [4] показано, что соотношения (2.5) справедливы и в более общих случаях, и приведены выражения для функций h_0, F_0 в этих случаях.

Дифференциальные уравнения и начальные условия (2.1), ограничения (1.3), условие окончания процесса и функционал (2.5) определяют, очевидно, дифференциальную игру без запаздывания информации. В самом деле, в каждый момент t игроку P известны фазовые координаты $x(t)$, $\eta(t)$ обеих сторон. Итак, доказано, что для получения гарантированного результата в игре с запаздыванием информации (1.1) — (1.5) достаточно решить дифференциальную игру (2.1), (1.3), (2.5) без запаздывания информации.

Так как дифференциальная игра с запаздыванием информации эквивалентна дифференциальной игре без запаздывания, то для решения игр с запаздыванием можно применить все известные подходы и результаты теории дифференциальных игр без запаздывания (см., например, [1-3]). Так можно составить уравнение Беллмана — Айзекса для игры (2.1), (1.3), (2.5), построить его характеристики и исследовать решение, как это делается в книге [1]. Предварительно нужно лишь вычислить функции h_0, F_0 из (2.5) по формулам (2.2) — (2.4). Отметим, что если уравнения движения системы E автономны, т. е. функция g не зависит от t, то уравнения движения (2.1) совпадут с (1.1).

Для игр преследования можно применять методику [3], связанную с определением момента первого поглощения. В § 3 приводится решение конкретных примеров на основе очевидной модификации этой методики для игр с запаздыванием информации.

Примечание 2.1. Решив дифференциальную игру (2.1), (1.3), (2.5) без запаздывания информации, определим синтезирующие управления $u = u(x, \eta, t)$ и $v = v(x, \eta, t)$ обеих сторон. Функция $u (x, \eta, t)$ дает синтез гарантирующего управления стороны P для исходной задачи с запаздыванием через текущие результаты измерений $x(t)$ и $\eta(t) = y(t - \tau)$. Функция $v(x, \eta, t)$ характеризует наихудшее для P управление стороны E, однако она не может быть использована стороной E, так как определяет управление v в момент $t—\tau$ через будущее значение фазового вектора $x(t)$. Для определения оптимального управления стороны E требуется задать характер информированности игрока E и снова решить задачу для стороны E.

Примечание 2.2. Доказанный выше факт эквивалентности игр с запаздыванием информации играм без запаздывания информации справедлив для более широкого класса задач, чем (1.1) — (1.5). Например, он имеет место для игр с интегральными ограничениями на управления, для игр с ограничениями на фазовые координаты. Существенно лишь, чтобы фазовые координаты и управляющие функции одной стороны не входили в уравнения движения и в ограничения для другой стороны.

§ 3. Примеры игр преследования с запаздыванием информации. Пусть движение двух объектов (преследователя и преследуемого) описывается системами дифференциальных уравнений с ограничениями

$$\begin{aligned} &\mathbf{r}_1^{\cdot} = \mathbf{w}_1, \quad \mathbf{w}_1^{\cdot} = \mathbf{a}_1, \quad |\mathbf{a}_1| \leqslant a_1 \\ &\mathbf{r}_2^{\cdot} = \mathbf{w}_2, \quad \mathbf{w}_2^{\cdot} = \mathbf{a}_2, \quad |\mathbf{a}_2| \leqslant a_2, \quad a_1 > a_2 \end{aligned} \tag{3.1}$$

Здесь $\mathbf{r}_1$, $\mathbf{r}_2$ — радиусы-векторы объектов, $\mathbf{w}_1$, $\mathbf{w}_2$ — их скорости, $\mathbf{a}_1$, $\mathbf{a}_2$ — ускорения, которые являются управляющими функциями, ограниченными по модулю постоянными a_1, a_2. Индекс 1 относится к преследователю, индекс 2 — к преследуемому. Все векторы в (3.1) имеют произвольную одинаковую размерность N. Игра считается оконченной, если расстояние между объектами станет равным заданному числу l (происходит захват), и условие (1.2) окончания игры имеет вид

$$h(\mathbf{r}_1(T_1), \ \mathbf{r}_2(T), \ T) = |\mathbf{r}_1(T) - \mathbf{r}_2(T)| - l = 0 \tag{3.2}$$

В качестве платы берется время, прошедшее от начала игры до момента захвата, причем преследователь стремится его уменьшить, а преследуемый — увеличить ($F = T$ в формуле (1.4)). Предполагается, что в каждый момент $t \geqslant 0$ преследователю известны свои фазовые координаты $\mathbf{r}_1$, $\mathbf{w}_1$ в этот момент и фазовые координаты преследуемого $\mathbf{r}_2$, $\mathbf{w}_2$ для момента $t - \tau$.

Поставленная задача решена по описанной выше схеме. Была вычислена функция h_0, и затем полученная игра без запаздывания информации исследовалась по методике [1].

Составлены и проинтегрированы уравнения характеристик, построены оптимальные траектории и управления, определена поверхность барьера в пространстве фазовых координат. Все исследование проводится аналогично решению задачи «изотропные ракеты» в [1].

Приведем здесь решение при помощи методики [3], которое оказывается более коротким из-за простоты областей достижимости. Пусть в начальный момент $t = 0$ преследователю известны свои фазовые координаты $\mathbf{r}_1^\circ$, $\mathbf{w}_1^\circ$ и фазовые координаты преследуемого $\mathbf{r}_2^\circ$, $\mathbf{w}_2^\circ$ для момента $t = -\tau$. Тогда областями достижимости преследователя и преследуемого в момент t в пространстве радиус-векторов будут сферы с центрами в точках

$$\mathbf{r}_1^\circ + \mathbf{w}_1^\circ t, \qquad \mathbf{r}_2^\circ + \mathbf{w}_2^\circ (t + \tau)$$

и с радиусами

$$a_1 t^2 / 2, \qquad a_2 (t + \tau)^2 / 2$$

соответственно. Составим уравнение, определяющее момент первого поглощения области достижимости преследуемого l-окрестностью области достижимости преследователя. Обозначим через $R\ (t)$ вектор, соединяющий центры областей достижимости, а через $Q\ (t)$ разность их радиусов (с учетом l-окрестности). Имеем, очевидно

$$R\ (t) = \mathbf{r}_2^\circ + \mathbf{w}_2^\circ (t + \tau) - \mathbf{r}_1^\circ - \mathbf{w}_1^\circ t \qquad (3.3)$$
$$Q\ (t) = l + a_1 t^2 / 2 - a_2 (t + \tau)^2 / 2$$

Момент поглощения T определится как первый положительный корень уравнения

$$|\, R\ (T)\, | = Q\ (T) \qquad (3.4)$$

Уравнение преследователя, гарантирующее захват в момент T, должно осуществлять прицеливание в ту точку своей области достижимости, которая отстоит на расстоянии l от точки касания области достижимости преследуемого и l-окрестности области достижимости преследователя. Следовательно, управление преследователя должно браться в виде

$$\mathbf{a}_1 = a_1 R\ (T) \,/\, |\, R\ (T)\, | \qquad (3.5)$$

где T есть первый положительный корень уравнения (3.4). Если преследователь придерживается управления (3.5), то преследуемый может избежать захвата до момента T, прицеливаясь в ту же точку касания, т. е. при управлении

$$\mathbf{a}_2 = a_2 R\ (T) \,/\, |\, R\ (T)\, | \qquad (3.6)$$

Пусть выполнено неравенство

$$Q\ (0) = l - a_2 \tau^2 / 2 > 0 \qquad (3.7)$$

представляющее собой необходимое условие возможности окончания игры (если оно нарушено, то преследуемый всегда сможет за время запаздывания τ выйти из зоны захвата радиуса l). Кроме того, предположим, что $|\, R\ (0)\, | > Q\ (0)$, т. е. в начальном положении захват еще не наступил. Из формул (3.3) следует, что при достаточно больших t и при $a_1 > a_2$, имеем $|R(t)| < Q(t)$. Отсюда вытекает, что уравнение (3.4) имеет нечетное число положительных корней. Так как с учетом равенств (3.3) это уравнение сводится к алгебраическому уравнению четвертой степени относительно T, то число положительных корней равно единице или трем.

Пусть при всех положительных t имеем $Q(t) > 0$. Тогда радиус l-окрестности области достижимости преследователя всегда больше радиуса области достижимости преследуемого, и эти области могут касаться не более чем в одной точке. В этом случае [3, 5] первый положительный корень T уравнения (3.4) есть гарантированное время окончания игры, и захват может быть осуществлен за конечное время при любых начальных условиях.

Пусть теперь функция $Q(t)$ обращается в нуль, и t_1 есть ее первый положительный корень. Если первый положительный корень T уравнения (3.4) удовлетворяет условию $T < t_1$, то касание областей достижимости по-прежнему возможно лишь в одной точке, и T снова есть гарантированное время окончания игры. Но при $T > t_1$ будут существовать такие моменты $t < T$, для которых радиус области достижимости преследуемого может быть больше или равен радиусу l-окрестности области достижимости преследователя. В этом случае, вообще говоря, нельзя гарантировать захват за время T при тех начальных данных, для которых $T > t_1$.

Управляя в соответствии с правилами (3.4) и (3.5), преследователь обеспечит не позже момента T позицию, при которой область достижимости преследуемого за время τ лежит целиком в l-окрестности точки местоположения преследователя.

Эта ситуация удовлетворяет условию гарантированного захвата для игр с запаздыванием информации (см. (2.3)).

При решении задачи по методике [1] приходим к тем же равенствам (3.4), (3.5) для времени T и оптимального уравнения преследователя.

Вдоль каждой оптимальной траектории величина T и управления (3.5), (3.6) постоянны, поэтому оптимальные траектории являются параболами. Формулы (3.3) — (3.5) определяют синтез оптимального управления преследователя, если под $\mathbf{r}_1^\circ$, $\mathbf{r}_2^\circ$, $\mathbf{w}_1^\circ$, $\mathbf{w}_2^\circ$ понимать текущие данные измерений. Время T как функция начальных фазовых координат терпит разрыв в тех точках, где происходит скачкообразное изменение первого корня уравнения (3.4). Это соответствует слиянию двух первых положительных корней уравнения (3.4), т. е. наличию кратного корня уравнения (3.4). Условие кратного корня имеет вид

$$d\,|R(T)|\,/\,dT = dQ(T)\,/\,dT \tag{3.8}$$

Исключая T из уравнений (3.4), (3.8), получим уравнение поверхности в пространстве фазовых координат, на которой T испытывает разрыв и которая называется барьером [1]. Скачок T на этой поверхности связан с необходимостью для преследователя выполнять маневр разворота [1].

Приведенное выше условие $Q(t) > 0$, как можно доказать, эквивалентно условию непересечения барьера [1]. Если же функция $Q(t)$ обращается в нуль при $t > 0$, то две ветви, из которых состоит поверхность барьера, пересекаются и выделяют в фазовом пространстве область, в которой $T < t_1$ и в которой игра оканчивается за конечное время. Вне этой области преследуемый может избежать захвата (обсуждение этих вопросов на примере игры изотропные ракеты см. в [1]).

Таким образом, захват за конечное время T при любых начальных условиях гарантирован, если $Q(t) > 0$ при всех $t \geqslant 0$. С учетом соотношений (3.3), (3.7) условие положительной определенности функции $Q(t)$ можно записать в виде

$$(a_1 - a_2)(l - a_2\tau^2 / 2) > a_2\tau^2 / 2$$

Отсюда получим условие гарантированной возможности захвата в окончательном виде

$$(a_1 - a_2)a_1^{-1} > a_2\tau^2 l^{-1} / 2 \tag{3.9}$$

Отметим, что неравенства $a_1 > a_2$ и (3.7) следуют из (3.9). Условие (3.9) показывает, что чем больше время запаздывания τ, тем большее превосходство в ускорении требуется преследователю для осуществления захвата.

Приведем еще некоторые примеры игр с запаздыванием информации, которые решаются аналогично задаче, рассмотренной выше.

Во всех примерах платой считается время преследования, условие окончания игры имеет прежний вид (3.2), индекс 1 относится к преследователю, индекс 2 — к преследуемому. Задача решается за преследователя, который получает информацию о фазовых координатах преследуемого с запаздыванием τ. Во всех следующих ниже примерах управляющие вектор-функции игроков обозначаются по-прежнему буквами $\mathbf{a}_1$, $\mathbf{a}_2$, но имеют различный смысл (либо скорости, либо ускорения). Момент поглощения T по-прежнему есть первый положительный корень уравнения (3.4), а оптимальные управления игроков определяются формулами (3.5), (3.6). Однако функции $R(T)$, $Q(T)$ в соотношениях (3.4) — (3.6) здесь имеют иной вид, чем в формулах (3.3). Ниже приводятся выражения этих функций для конкретных примеров.

А. Пусть оба игрока управляются по скорости, и их уравнения движения и ограничения имеют вид

$$\mathbf{r}_1^{\cdot} = \mathbf{a}_1, \quad \mathbf{r}_2^{\cdot} = \mathbf{a}_2, \quad |\mathbf{a}_1| \leqslant a_1, \quad |\mathbf{a}_2| \leqslant a_2$$

Здесь, как и всюду, a_1, a_2 — постоянные. В этом случае функции $R(T)$, $Q(T)$ равны

$$R(T) = \mathbf{r}_2^{\circ} - \mathbf{r}_1^{\circ}, \quad Q(T) = l + a_1 T - a_2(T + \tau)$$

Захват при любых начальных ситуациях гарантирован, если выполнены два условия $a_1 > a_2$ и $l > a_2\tau$.

Б. Пусть преследователь управляется по ускорению, а преследуемый — по скорости. Уравнения движения имеют вид

$$\mathbf{r}_1^{\cdot} = \mathbf{w}_1, \quad \mathbf{w}_1^{\cdot} = \mathbf{a}_1, \quad \mathbf{r}_2^{\cdot} = \mathbf{a}_2, \qquad |\,\mathbf{a}_1\,| \leqslant a_1, \qquad |\,\mathbf{a}_2\,| \leqslant a_2$$

Здесь функции R, Q равны

$$R\,(T) = \mathbf{r}_2^{\circ} - \mathbf{r}_1^{\circ} - \mathbf{w}_1^{\circ} T, \quad Q\,(T) = l + a_1 T^2 \,/\, 2 - a_2\,(T + \tau)$$

Захват при любых начальных условиях гарантирован, если

$$l > a_2^2 a_1^{-1} \,/\, 2 + a_2 \tau$$

Отметим, что при $\tau = 0$ это решение переходит в решение задачи изотропные ракеты [1].

В. Если преследователь управляется по скорости, а преследуемый — по ускорению, то уравнения движения и функции R, Q имеют вид

$$\mathbf{r}_1^{\cdot} = \mathbf{a}_1, \quad \mathbf{r}_2^{\cdot} = \mathbf{w}_2, \quad \mathbf{w}_2^{\cdot} = \mathbf{a}_2, \qquad |\,\mathbf{a}_1\,| \leqslant a_1, \qquad |\,\mathbf{a}_2\,| \leqslant a_2$$

$$R\,(T) = \mathbf{r}_2^{\circ} + \mathbf{w}_2^{\circ}\,(T + \tau) - \mathbf{r}_1^{\circ}, \quad Q\,(T) = l + a_1 T - a_2\,(T + \tau)^2 \,/\, 2$$

Здесь при любых параметрах задачи существуют начальные ситуации, в которых преследуемый может избежать захвата.

Поступила 12 V 1969

ЛИТЕРАТУРА

1. Айзекс Р. Дифференциальные игры. М., «Мир», 1967.
2. Понтрягин Л. С. К теории дифференциальных игр. Усп. матем. н., 1966, т. 21, вып. 4.
3. Красовский Н. Н. Теория управления движением. М., «Наука» 1968.
4. Черноусько Ф. Л. О дифференциальных играх с запаздыванием информации. Докл. АН СССР, 1969, т. 188, № 4.
5. Красовский Н. Н., Субботин А. И. Задача о сближении управляемых объектов. ПММ, 1968, т. 32, вып. 4.