Прикладная математика и механика. Том 51, вып. 5, 1987

УДК 62-50

ЗАДАЧА ДЕЦЕНТРАЛИЗОВАННОГО УПРАВЛЕНИЯ С РАЗРЫВНОЙ ФУНКЦИЕЙ ВЫИГРЫША

Ермолов А. Н.

Рассматривается задача децентрализованного управления [1] о встрече двух игроков, приходящих в условленное место в случайные моменты времени. Игроки обладают частичной информацией о неконтролируемых факторах, т. е. знают только момент своего прихода в назначенное место. У игроков общая цель — дождаться друг друга, затратив при этом на ожидание как можно меньше времени. Особенностью рассматриваемой задачи является разрывность критерия по управлениям. Исследуются свойства оптимальной стратегии (ОС) коллектива [1—5] — двух встречающихся игроков, и предлагается метод ее построения. Приводится пример нахождения ОС при равномерном распределении вероятности моментов прихода на отрезке.

1. **Формулировка задачи.** Два игрока условились встретиться в назначенном месте, однако моменты времени t_1 и t_2 их прихода заранее неизвестны и игрокам не выбираются. Известно лишь, что эти моменты случайным образом принимают значения из R, их распределение одинаково, независимо и задается при помощи вероятностной меры μ. Предполагается, что мера μ определена на борелевской σ-алгебре пространства R, положительна и регулярна, т. е. для любого измеримого $A \subset R$ и любого $\varepsilon > 0$ существует открытое множество Q и замкнутое E, такие, что $E \subset A \subset Q$ и $\mu(Q \setminus E) < \varepsilon$ [6].

Каждый игрок может оставаться в назначенном месте и ждать партнера сколь угодно долго; обозначим $u_1 \geqslant 0$, $u_2 \geqslant 0$ длительности нахождения игроков в назначенном месте. Выигрыш коллектива из двух игроков (оперирующей стороны) складывается из условной ценности встречи ρ, если встреча состоялась, из которой вычитается время, затраченное обоими игроками на ожидание. Таким образом, если игроки пришли в назначенное место в моменты t_1 и t_2 и ждали в течение времени u_1 и u_2, то их выигрыш определяется по формуле

$$
f(t_1, t_2, u_1, u_2) = \begin{cases} \rho - (t_2 - t_1), & \text{если } t_1 \leqslant t_2 \leqslant t_1 + u_1 \\ \rho - (t_1 - t_2), & \text{если } t_2 \leqslant t_1 \leqslant t_2 + u_2 \\ -u_1 = u_2 & \text{в противном случае} \end{cases} \tag{1.1}
$$

Обозначим U множество всех измеримых по Борелю функций $u: R \to [0, +\infty)$. Стратегиями оперирующей стороны (двух игроков) являются всевозможные пары

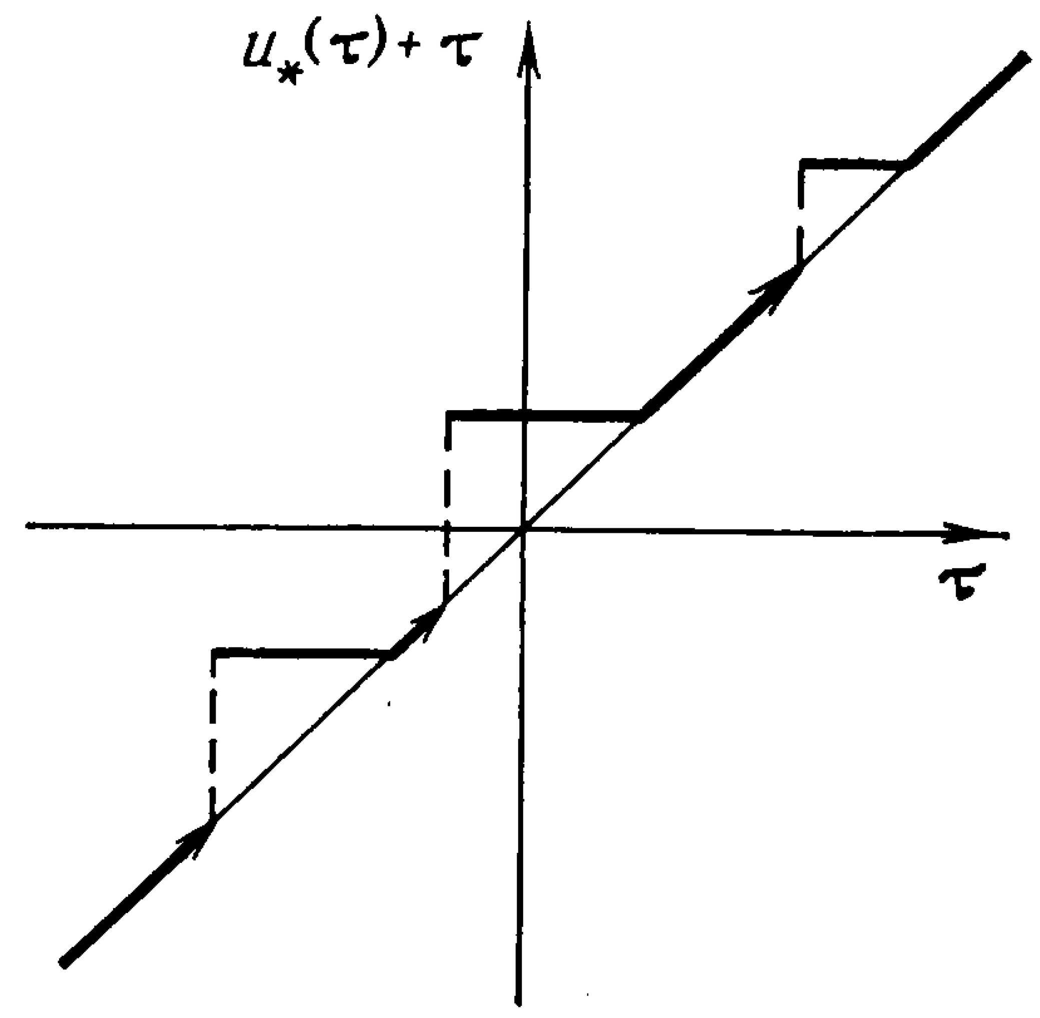
$$
(u_1(\cdot), u_2(\cdot)), \quad u_i(\cdot) \in U \quad (i = 1, 2) \tag{1.2}
$$

где $u_i(t)$ имеет смысл времени ожидания i-го игрока в случае, если он пришел в назначенное место в момент t. Если партнеры выбрали некоторую стратегию (1.2), то математическое ожидание выигрыша

$$
F(u_1(\cdot), u_2(\cdot)) = \int\limits_R \int\limits_R f(t_1, t_2, u_1(t_1), u_2(t_2))\, d\mu(t_1)\, d\mu(t_2) \tag{1.3}
$$

Цель коллектива игроков — максимизация функционала (1.3).

Будет показано, что существует оптимальная стратегия (ОС) оперирующей стороны вопрос о единственности ОС не рассматривается), обладающая свойствами: а) симметричности, т. е. ОС имеет вид $(u_*(\cdot), u_*(\cdot))$ (лемма 1); б) «ступенчатости», т. е. функция $u_*(\tau) + \tau$ имеет вид, изображенный на фигуре (лемма 3). Приводится также способ вычисления ОС при заданных мере μ и ценности встречи ρ (теорема).



В дальнейшем для краткости вместо $d\mu(t)$ будем писать просто $d\mu$.

2. Симметричность и ступенчатость ОС. Покажем, что среди ОС имеются симметричные.

Лемма 1. Для любой стратегии (1.2) существует функция $v(\cdot) \in U$, такая, что

(2.1) $$F(v(\cdot), v(\cdot)) \geqslant F(u_1(\cdot), u_2(\cdot))$$

Доказательство. Функция (1.1) симметрична, т. е. не меняется при перестановке индексов 1 и 2, поэтому функционал (1.3) можно представить следующим образом:

$$F(u_1(\cdot), u_2(\cdot)) = F'(u_1(\cdot), u_2(\cdot)) + F'(u_2(\cdot), u_1(\cdot))$$

$$F'(u_1(\cdot), u_2(\cdot)) = \int_R \Big(\int_{[\tau, \infty)} f(\tau, t, u_1(\tau), u_2(t))\, d\mu \Big) d\mu(\tau)$$

Положим

$$v_*(t) = \min(u_1(t), u_2(t)), \quad v^*(t) = \max(u_1(t), u_2(t))$$

и докажем неравенство

(2.2) $$\tfrac{1}{2}(F(v^*(\cdot), v^*(\cdot)) + F(v_*(\cdot), v_*(\cdot))) \geqslant F(u_1(\cdot), u_2(\cdot))$$

Его левую часть можно представить в виде

$$\int_R \Big[\Big(\int_{[\tau, \tau+v^*(\tau)]} (\rho - (t-\tau))\, d\mu + \int_{[\tau, \tau+v_*(\tau)]} (\rho - (t-\tau))\, d\mu \Big) -$$
$$- (v^*(\tau)\, \mu((\tau + v^*(\tau), \infty)) + v_*(\tau)\, \mu((\tau + v_*(\tau), \infty))) - I_v \Big] d\mu(\tau)$$

$$I_v = \int_{(\tau+v^*(\tau), \infty)} v^*(t)\, d\mu + \int_{(\tau+v_*(\tau), \infty)} v_*(t)\, d\mu$$

Поскольку для каждого фиксированного $\tau \in R$ либо $v^*(\tau) = u_1(\tau)$ и $v_*(\tau) = u_2(\tau)$, либо $v^*(\tau) = u_2(\tau)$ и $v_*(\tau) = u_1(\tau)$, то правая часть (2.2) отличается лишь тем, что вместо величины I_v будет стоять

$$I_u = \int_{(\tau+u_1(\tau), \infty)} u_2(t)\, d\mu + \int_{(\tau+u_2(\tau), \infty)} u_1(t)\, d\mu$$

Значит, для доказательства (2.2) достаточно показать, что

(2.3) $$I_v \leqslant I_u, \quad \forall \tau \in R$$

Пусть, например, $u_1(\tau) \geqslant u_2(\tau)$. Тогда неравенство (2.3) следует из того, что $\forall t$ $u_1(t) + u_2(t) = v^*(t) + v_*(t)$, а по определению $v_*(t)$

$$\int_{(\tau+u_2(\tau), \tau+u_1(\tau)]} (v_*(t) - u_2(t))\, d\mu \leqslant 0$$

Значит, справедливо и неравенство (2.2), из которого в свою очередь следует, что условие (2.1) выполняется либо для $v(\cdot) = v^*(\cdot)$, либо для $v(\cdot) = v_*(\cdot)$, что и требовалось доказать.

Фиксируем некоторую функцию $v(\cdot) \in U$ и рассмотрим функцию

$$\varphi(t_1, u_1; v(\cdot)) = \int_R f(t_1, t_2, u_1, v(t_2))\, d\mu(t_2)$$

Лемма 2. Функция $\varphi(t, u; v(\cdot))$ полунепрерывна сверху и непрерывна справа по u при всех $t \in R$ и $v(\cdot) \in U$.

Доказательство следует из того, что функция (1.1) полунепрерывна сверху и непрерывна справа по u_1 при всех $t_1, t_2, u_2 \geqslant 0$, а также из регулярности меры μ.

Лемма 3. Среди ОС существует симметричная стратегия $(v(\cdot), v(\cdot))$ вида

$$v(t) = \begin{cases} v_i - (t - \theta_i) & \text{при } \tau \in [\theta_i, \theta_i + v_i], \quad i \in I \\ 0 & \text{в противном случае} \end{cases}$$

причем интервалы $[\theta_i, \theta_i + v_i]$, $i \in I$ попарно не пересекаются.

Доказательство. 1°. Функцию φ при произвольной $v(\cdot) \in U$ можно представить в виде

$$\varphi(\tau, u; v(\cdot)) = \int\limits_{(-\infty, \tau)\setminus L(\tau)} (\rho - (t - \tau))\, d\mu + \int\limits_{L(\tau)} (-u - v(t))\, d\mu + \tag{2.4}$$

$$+ \int\limits_{[\tau, \tau+u]} (\rho - (t - \tau))\, d\mu + \int\limits_{(\tau+u, \infty)} (-u - v(t))\, d\mu$$

$$(L(\tau) = \{t \in R \mid t + v(t) < \tau\})$$

Пусть $\theta \in [\tau, \tau + u]$, тогда, пользуясь представлением (2.4), получим

$$\varphi(\theta, u - (\theta - \tau); v(\cdot)) = \varphi(\tau, u; v(\cdot)) - Mu + K \tag{2.5}$$

где M и K не зависят от u, $M = \mu(L(\theta) - L(\tau)) \geqslant 0$.

Рассмотрим некоторую симметричную оптимальную стратегию $(u(\cdot), u(\cdot))$; она существует согласно лемме 1. Фиксируем $\tau \in R$ и рассмотрим последовательность $\{\tau_j\}$, сходящуюся к τ справа. Предположим, что для каждого $j = 1, 2, \ldots$ существует $u_j \in \arg\max \varphi(\tau_j, u; u(\cdot))$, такое, что $u_j + \tau_j \geqslant a$. (Здесь и далее, если не оговорено противное, max берется по $u \geqslant 0$.) Покажем, что тогда и τ обладает таким же свойством. Для любого $j = 1, 2, \ldots$ и $u' \in [0, u_j]$ $\varphi(\tau_j, u_j; u(\cdot)) \geqslant \varphi(\tau_j, u'; u(\cdot))$. Отсюда, пользуясь соотношением (2.5), имеем

$$\varphi(\tau, u_j + (\tau_j - \tau); u(\cdot)) \geqslant \varphi(\tau, u' + (\tau_j - \tau); u(\cdot))$$

Следовательно, для любого $u \in (0, u_j]$ найдется $v \geqslant 0$, такое, что $\tau + v \geqslant a$ и

$$\varphi(\tau, v; u(\cdot)) \geqslant \varphi(\tau, u; u(\cdot)) \tag{2.6}$$

Так как функция $\varphi(\tau, u; u(\cdot))$ непрерывна справа по u, то (2.6) справедливо и при $u = 0$. Значит, для выбранного τ найдется $v \in \operatorname{argmax} \varphi(\tau, u; u(\cdot))$, удовлетворяющее условию $v + \theta \geqslant \tau$.

Доказанное утверждение означает, что для всякого момента τ существует наименьшее $\theta \leqslant \tau$, такое, что найдется $v \in \operatorname{argmax} \varphi(\theta, u, u(\cdot))$, удовлетворяющее условию $v + \theta \geqslant \tau$. Из неравенства (2.6) следует, что для всякого $t < \theta$ и $v \in \operatorname{argmax} \varphi(t, u; u(\cdot))$ выполнено $v + t < \theta$ (в противном случае θ не было бы наименьшим из таких, что существует $v \in \operatorname{argmax} \varphi(\theta, u, u(\cdot))$, $\theta \leqslant \tau \leqslant v + \theta$). Значит, для стратегии $(u(\cdot), u(\cdot))$ имеется набор $\{\theta_i \in R, i \in I\}$, такой, что для любого $i \in I$ существует

$$v_i \in \operatorname{argmax} \varphi(\theta_i, u; u(\cdot)) \tag{2.7}$$

такое, что

а) $t + v < \theta_i$, если $t < \theta_i$, $v \in \operatorname{argmax} \varphi(t, u; u(\cdot))$;

б) $u(t) + t \leqslant \theta_i + v_i$, $t \in [\theta_i, \theta_i + v_i]$.

2°. Обозначим

$$v(t) = \begin{cases} v_i - (t - \theta_i), & \text{если } t \in [\theta_i, \theta_i + v_i] \\ 0 & \text{в противном случае} \end{cases} \tag{2.8}$$

Покажем, что $F(v(\cdot), v(\cdot)) \geqslant F(u(\cdot), u(\cdot))$. Из определения множества $L(t)$ следует

$$\frac{1}{2} F(u(\cdot), u(\cdot)) = \int\limits_{(-\infty, \infty)} \Bigg[\int\limits_{(\tau, \tau+u(\tau))} (\rho - (t - \tau))\, d\mu(t) -$$

$$- u(\tau)\, \mu((\tau + u(\tau), \infty)) - u(\tau) \mu(L(\tau)) \Bigg] d\mu(\tau)$$

Отсюда, пользуясь определением функционала F функции φ и соотношениями (2.7), (2.8), получим

$$(2.9)\qquad {}^1/_2F(v(\cdot), v(\cdot)) - {}^1/_2F(u(\cdot), u(\cdot)) = \\ = \sum_{i\in I}\int\limits_{[\theta_i,\,\theta_i+v_i]} (\varphi(\theta_i, v(\tau)+\tau-\theta_i; u(\cdot)) - \varphi(\theta_i, u(\tau) + \\ +\tau-\theta_i; u(\cdot)))\,d\mu(\tau)$$

Правая часть равенства (2.9) неотрицательна, так как $v(\tau)+\tau-\theta_i = v_i \in \in \operatorname{argmax}\varphi(\theta_i, u; u(\cdot))$, $\tau \in [\theta_i, \theta_i + v_i]$.

Таким образом, функция $v(t) \in U$, определенная соотношением (2.8), оптимальна и имеет вид, описанный в условии леммы.

3. Построение ОС. Рассмотрим функцию

$$(3.1)\qquad g(\tau, u, u_0) = \int\limits_{[\tau,\tau+u]} (\rho-(t-\tau))\,d\mu - u\mu((-\infty, u)\cup(\tau+u, \infty)) - \\ - \int\limits_{(\tau+u,\tau+u_0]} (u_0-(t-\tau))\,d\mu$$

Здесь множество $(a, b]$ при $a \geqslant b$ считается пустым.

Лемма 4. Пусть $u_0' < u_0''$, тогда для любого τ $g(\tau, u, u_0') - g(\tau, u, u_0'')$ — неотрицательная монотонно возрастающая функция u; функция $g(\tau, u, u_0)$ непрерывна по u_0.

Доказательство. По определению функции $g(\tau, u, u_0)$

$$(3.2)\qquad g(\tau, u, u_0') - g(\tau, u, u_0'') = \int\limits_{(\tau+u,\,\tau+u_0']} (u_0''-u_0')\,d\mu + \\ + \int\limits_{(\tau+\max(u,u_0'),\,\tau+u_0'']} (u_0''-(t-\tau))\,d\mu$$

Выражение в левой части равенства неотрицательно и монотонно не возрастает по u, так как подынтегральные функции неотрицательны. Непрерывность $g(\tau, u, u_0)$ также следует из равенства (3.2).

Функция $g(\tau, u, u_0)$ при любых τ и u_0 полунепрерывна сверху и непрерывна справа по u; доказательство аналогично доказательству этого утверждения для $\varphi(\tau, u; v(\cdot))$.

Обозначим

$$V_0(\tau) = \{u_0 \geqslant 0 \mid \max g(\tau, u, u_0) = g(\tau, u_0, u_0)\}$$

Теорема. Существует функция $u_0(t) \in V_0(t)$, $t \in R$, такая, что стратегия $(v(\cdot), v(\cdot))$ доставляет максимум функционалу (1.3), где

$$(3.3)\qquad v(\tau) = \max_{t\leqslant\tau}(u_0(t)-\tau+t),\quad \tau\in R$$

Доказательство. Согласно лемме 3, существует набор пар $\{(\theta_i, v_i)$, $i \in I\}$, такой, что отрезки $[\theta_i, \theta_i + v_i]$, $i \in I$ попарно не пересекаются и функция

$$v(t) = \begin{cases} v_i+\theta_i-t & \text{при } t\in[\theta_i,\theta_i+v_i],\ i\in I\\ 0 & \text{в противном случае}\end{cases}$$

такова, что стратегия $(v(\cdot), v(\cdot))$ доставляет максимум функционалу (1.3). Для такой функции $v(t)$ нужно построить функцию $u_0(t)$, удовлетворяющую условиям теоремы. Так как $(v(\cdot), v(\cdot))$ — ОС, то [1]

$$(3.4)\qquad \varphi(\tau, v(\tau); v(\cdot)) = \max\varphi(\tau, u; v(\cdot)),\ \forall\tau\in R$$

При $\tau = \theta_i$ и $u \in [0, v_i]$

$$\varphi(\tau, u; v(\cdot)) = \int\limits_{[\tau,\tau+u]} (\rho-(t-\tau))\,d\mu - \int\limits_{(\tau+u,\tau+v_i]} (v_i-(t-\tau))\,d\mu - \\ - u\mu((-\infty,\tau)\cup(\tau+u,\infty)) - \int\limits_{(-\infty,\tau)\cup(\tau+v_i,\infty)} v(t)\,d\mu$$

Последнее слагаемое не зависит от u, поэтому из (3.4) следует

(3.5) $$g(\tau, v_i, v_i) \geqslant g(\tau, u, v_i), \quad u \in [0, v_i]$$

Аналогично можно показать, что из (3.4) следует неравенство (3.5) и для $u > v_i$, так как величина

(3.6) $$-\int\limits_{-\infty,\tau)\cup(\tau+u,\infty)} v(t)\, d\mu$$

не убывает по u. Таким образом, $v_i \in V_0(\theta_i)$, $i \in I$.

При $\tau \in [\theta_i, \theta_i + v_i]$ имеем $\varphi(\tau, 0, 0) = \max \varphi(\tau, u; v(\cdot))$, а значит, $g(\tau, 0, 0) = \max g(\tau, u, 0)$, так как величина (3.6) не убывает по u.

Остается показать, что для $\tau \in [\theta_i, \theta_i + v_i]$ существует $u_0 \in V_0(\tau)$, $u_0 \leqslant v_i - (\tau - \theta_i)$.

По определению функций $\varphi(\tau, u; v(\cdot))$ и $g(\tau, u, u_0)$

$$g(\tau, u, v_i - (\tau - \theta_i)) = \varphi(\tau, u; v(\cdot)) - u\mu([\theta_i, \tau]) -$$
$$- \int\limits_{(\theta_i+v_i,\infty)} v(t)\, d\mu, \quad u \in [0, v_i - (\tau - \theta_i)]$$

Так как $\varphi(\tau, u, v(\cdot))$ достигает максимума в точке $u = v_i - (\tau - \theta_i)$, то отсюда следует, что существует точка $\xi_1 \in [0, v_i - (\tau - \theta_i)]$, такая, что

$$g(\tau, \xi_1, v_i - (\tau - \theta_i)) = \max g(\tau, u, v_i - (\tau - \theta_i))$$

Согласно лемме 4, существует $\xi_2 \in [0, \xi_1]$, такая, что $g(\tau, \xi_2, \xi_1) = \max g(\tau, u, \xi_1)$ и т. д.

Последовательность $\{\xi_k\}$ монотонно не возрастает и ограничена, т. е. имеет предел u_0, причем $u_0 \in [0, v_i - (\tau - \theta_i)]$. Для любого $u \geqslant 0$ выполнено неравенство

(3.7) $$g(\tau, \xi_{k+1}, \xi_k) \geqslant g(\tau, u, \xi_k), \quad k = 1, 2, \ldots$$

Так как функция $g(\tau, u, u_0)$ непрерывна по u_0 и непрерывна справа по u, то, переходя к пределу в неравенстве (3.7), получим

$$g(\tau, u_0, u_0) \geqslant g(\tau, u, u_0), \quad u \geqslant 0$$

Последнее означает, что $u_0 \in V_0(\tau)$. Построенная функция $u_0(\tau)$, $\tau \in R$ совпадает с $v(\tau)$ при $\tau \notin [\theta_i, \theta_i + v_i]$ и $u_0(\tau) \leqslant v(\tau)$ при $\tau \in (\theta_i, \theta_i + v_i]$.

Из определения $v(t)$ следует равенство (3.3), что и завершает доказательство теоремы.

Замечание 1. Утверждение теоремы остается в силе в случае, когда $\mu(R) < 1$. Достаточно определить функцию $g(\tau, u, u_0)$ так:

$$g(\tau, u, u_0) = \int\limits_{[\tau, \tau+u]} (\rho - (t - \tau))\, d\mu - u(1 - \mu([\tau, \tau + u]) -$$
$$- \int\limits_{(\tau+u, \tau+u_0]} (u_0 - (t - \tau))\, d\mu$$

Приведем пример построения ОС при помощи теоремы. Пусть

$$\mu(A) = \int\limits_A p(t)\, dt, \qquad p(t) = \begin{cases} 1, & \text{если } t \in [0, 1] \\ 0 & \text{в противном случае} \end{cases}$$

(равномерное распределение вероятности прихода в назначенное место на отрезке [0, 1]). Тогда

$$g(\tau, u, u_0) = \begin{cases} u(\rho - 1 - u_0), & u \in [0, u_0] \\ \rho u + u^2/2 - u, & u \in (u_0, 1 - \tau] \end{cases}$$

Множество решений уравнения

$$g(\tau, u_0, u_0) = \max_{u \in [0, 1-\tau]} g(\tau, u, u_0)$$

имеет вид

$$V_0(\tau) = \begin{cases} \{1-\tau\}, & \tau \in [0, 2\rho - 1) \\ \{1-\tau, 0\}, & \tau \in [2\rho - 1, \rho] \\ \{0\}, & \tau \in (\rho, 1] \end{cases}$$

Значит, если $\rho \geqslant {}^1/_2$, то единственная функция, удовлетворяющая условиям теоремы, имеет вид

$$v(t) = 1 - t, \quad t \in [0, 1]$$

Если же $\rho < {}^1/_2$, то все функции

$$v_a(t) = \begin{cases} 0, & t \in [0, a] \\ 1 - t, & t \in (a, 1] \end{cases}$$

а также функция $v(t) = 0$, $t \in [0, 1]$ удовлетворяют условиям теоремы. Вычислив значение критерия $F(v_a(\cdot), v_a(\cdot))$ и проведя оптимизацию по параметру, получим, что при $\rho > {}^1/_3$ ОС — $v(t) = 1 - t$, $t \in [0, 1]$; при $\rho < {}^1/_3$ ОС — $v(t) = 0$, $t \in [0, 1]$; при $\rho = {}^1/_3$ обе стратегии оптимальны.

Замечание 2. Если распределение вероятности времени прихода в условленное место у игроков неодинаково, то ОС, вообще говоря, не обладает свойствами симметричности и ступенчатости. Например, в случае, если

$$\mu_i(A) = \int\limits_A p_i(t)\,dt, \quad i = 1, 2$$

$$p_1(t) = \pi^{-1/2} \exp(-t^2), \quad t \in R$$

$$p_2(t) = \begin{cases} 1/(2b), & \text{если } t \in [-b, b] \\ 0 & \text{в противном случае} \end{cases}$$

можно показать, что при соответствующем выборе чисел ρ и b любая ОС несимметричная и неступенчатая.

ЛИТЕРАТУРА

1. *Marschak J.*, *Radner R.* Economic Theory of Teams. New Haven: Yale Univ. Press. 1971. 345 p.
2. *Witsenhausen H. S.* Separation of estimation and control for discrate time systems // Pros. IEEE. 1971. V. 59. No. 11. P. 1557—1566.
3. *Arrow K. J.*, *Radner R.* Allocation of resources in large teams // Econometrica. 1979. V. 47. No. 2. P. 361—385.
4. *Tsitsiklis J. N.*, *Athans M.* On complexity of decentralized decision making and detection problems // IEEE Trans. Autom. Contr. 1985. V. 30. No. 5. P. 440—446.
5. *Ермолов А. Н.* О динамической задаче децентрализованного управления // ПММ. 1983. Т. 47. Вып. 1. С. 11—20.
6. *Варга Дж.* Оптимальное управление дифференциальными и функциональными уравнениями. М.: Наука. 1977. 623 с.

Москва

Поступила в редакцию
12.IX.1986