

УДК 519.86

ББК 65с

ДИНАМИЧЕСКИЕ ИЕРАРХИЧЕСКИЕ ИГРЫ ДВУХ ЛИЦ В ПРОГРАММНЫХ СТРАТЕГИЯХ И ИХ ПРИЛОЖЕНИЯ

ГЕННАДИЙ А. УГОЛЬНИЦКИЙ*

АНАТОЛИЙ Б. УСОВ

Южный федеральный университет

344090, Ростов-на-Дону, ул. Мильчакова, 8А

e-mail: ougoln@mail.ru, usov@math.rsu.ru.

В статье рассмотрены частные классы дифференциальных иерархических игр двух лиц в программных стратегиях. Дана их интерпретация как моделей коррупции и управления устойчивым развитием в иерархических системах. Сформулированы принципы оптимальности для этих классов игр как модификации равновесия по Штакельбергу. Предложены алгоритмы построения равновесий указанных игр, апробированные на числовых примерах.

Ключевые слова: динамические игры, иерархические системы, равновесие по Штакельбергу, коррупция, устойчивое развитие.

1. Введение

Анализ динамических моделей иерархического управления позволяет заключить, что основные результаты здесь получены в рамках теории динамических игр Штакельберга [8] и ее специфической разновидности – теории динамических игр Гермейера [1]. По сравнению со статическими постановками, в динамике возникает ряд новых моментов, а именно:

1. *Динамика принятия решений и динамика управляемой системы.* Принятие решений может неявно рассматриваться как процесс в любой иерархической игре, поскольку Ведущий и Ведомый игроки делают ходы последовательно. В многошаговых играх динамика принятия решений описывается явно и подробно, приводя к специфическим концепциям решения по Штакельбергу и способам его нахождения [8]. Описание динамики управляемой системы полностью отсутствует в статическом случае и может проводиться в дискретном или непрерывном времени. В первом случае принято говорить о разностных играх, во втором – о дифференциальных.

2. *Информационная структура.* В динамическом случае она гораздо богаче и разнообразнее, нежели в статическом. На наш взгляд, принципиальное значение здесь имеют два «водораздела». Во-первых, наличие обратной связи по состоянию управляемой системы: определяется ли стратегия сразу на весь период рассмотрения (программная стратегия) или она зависит от состояния управляемой системы (позиционная стратегия). Возможные разновидности внутри класса позиционных стратегий (зависит ли стратегия только от текущего состояния, от всей предыстории или ее части и т.п.) представляются не столь существенными, хотя могут приводить к различиям в математических деталях. Во-вторых, наличие обратной связи (зависимости) по управлению: принимает ли Ведущий решения независимо от поведения Ведомого или в зависимости от него. Наиболее рельефно это различие проводится в динамических играх Гермейера (табл.1). Решения этих игр естественно называть равновесиями по Штакельбергу в соответствующих играх Гермейера.

3. *Динамическая устойчивость.* Это важнейшее понятие теории динамических игр содержательно означает, что никому из игроков не выгодно отклоняться от первоначально достигнутого соглашения (решения) на протяжении всего периода развития игры [2,3]. Известно, что равновесие по Штакельбергу в программных стратегиях (open-loop) не является динамически устойчивым, в то время как равновесие по Штакельбергу в позиционных стратегиях (closed-loop) является сильно динамически устойчивым.

В настоящей статье развивается описанная в [1] идея построения решения игры путем разработки взаимовыгодной для игроков про-

граммы и наказания при отклонении от нее.

Таблица 1. Классификация динамических игр Гермейера

Зависимость по состоянию Зависимость по управлению	Отсутствует (программные стратегии)	Присутствует (позиционные стратегии)
Отсутствует	Γ_{1t}	Γ_{1x}
Присутствует	Γ_{2t}	Γ_{2x}

2. Частные классы динамических иерархических игр

Для простоты и определенности в качестве исходной «родовой» модели ограничимся случаем дифференциальных игр в программных стратегиях двух игроков (Ведущий и Ведомый) следующего вида:

$$J_L(p, q, u, x) = \int_0^\infty e^{-\alpha t} g_L(p(t), q(t), u(t), x(t)) dt \rightarrow \max \quad (2.1)$$

$$p(t) \in P, \quad q(t) \in Q, \quad p(t) \in P, \quad q(t) \in Q \quad (2.2)$$

$$J_F(p, u, x) = \int_0^\infty e^{-\alpha t} g_F(p(t), u(t), x(t)) dt \rightarrow \max \quad (2.3)$$

$$u(t) \in U(q), \quad (2.4)$$

$$\dot{x} = f(x(t), u(t)), \quad x(0) = x_0, \quad (2.5)$$

где J_L, J_F – соответственно функционалы выигрыша Ведущего и Ведомого, $p(t)$ или $q(t)$ и $u(t)$ – их программные стратегии, P, Q, U – множества допустимых стратегий, α – коэффициент дисконтирования; $x(t)$ – вектор состояния управляемой динамической системы, описываемый уравнением (2.5), причем $x(t) = x(u(t), t)$; $x(t) \in X$; X – множество возможных состояний системы.

Предполагается, что здесь P, Q, U, X, X^* – компактные множества в соответствующих пространствах; J_L непрерывен на $P \times Q \times U \times X$; J_F непрерывный и строго выпуклый на $P \times U \times X$; p, q, u кусочно-непрерывны по t ; f непрерывно дифференцируема по x и u и удовлетворяет условию Липшица по x . Рассматриваются несобственные интегралы, поскольку это более адекватно задаче управления устойчивым развитием [4].

Предполагается, что модель (2.1)-(2.5) имеет две особенности. Во-первых, динамика (2.5) зависит только от воздействия Ведомого. Ведущий контролирует состояние системы (2.5) опосредованно, через воздействие на Ведомого. Во-вторых, интересы игроков либо сонаправлены по u (т.е. $\sup_{u \in U(q)} J_L(p, q, u, x) = \sup_{u \in U(q)} J_F(p, u, x)$; $\forall p(t) \in P$; $q(t) \in Q$), либо противонаправлены по u ($\sup_{u \in U(q)} J_L(p, q, u, x) = \inf_{u \in U(q)} J_F(p, u, x)$; $\forall p(t) \in P$; $q(t) \in Q$).

В первом случае основная (хотя и не единственная) содержательная интерпретация состоит в том, что деятельность Ведомого приносит доход, который делится между игроками. Во втором случае Ведомый, например, экономит свои трудовые усилия, которые Ведущий хотел бы увеличить. Будем в дальнейшем для определенности считать, что реализуется предположение о сонаправленности интересов игроков (второе предположение рассматривается аналогично).

Рассмотрим с учетом сделанных предположений некоторые частные классы динамических иерархических игр, которые, на наш взгляд, представляют интерес с точки зрения дальнейшего развития теории и приложений моделей иерархически управляемых динамических систем.

Динамические иерархические игры с зависимостью (обратной связью) по управлению. В этом случае стратегия Ведущего игрока зависит от стратегии Ведомого: $p(t) = p(u(t))$ или $q(t) = q(u(t))$. Этот класс игр соответствует динамическим играм Гермейера типа Γ_{2t} . Содержательно данный класс моделей описывает внутреннюю адаптацию в системе; одной из наиболее ярких интерпретаций является коррупция [5,7, 9-14,16]. Отметим, что моделями внешней адаптации служат игры в позиционных стратегиях.

Динамические иерархические игры принуждения и побуждения. Под принуждением будем понимать воздействие Ведущего только на множество допустимых стратегий Ведомого (как правило, без зависимости по управлению), а под побуждением – воздействие Ведущего только на функцию выигрыша Ведомого (как правило, с зависимостью по управлению). Таким образом, в играх принуждения $U = U(q)$, где q – стратегия Ведущего (функция $p = p(t)$ задана и фиксирована), а в играх побуждения $J_F = J_F(p, u, x)$ (функция $q = q(t)$

задана и фиксирована). Содержательно данный класс моделей описывает принуждение и побуждение как базовые методы иерархического управления (административно-законодательный и экономический) [4].

Таблица 2. Теоретико-игровые модели иерархического управления и их интерпретация

№	Модель	Интерпретация
1	Игра принуждения по Штакельбергу с обратной связью по управлению	Административная коррупция
2	Игра побуждения по Штакельбергу с дополнительной обратной связью по управлению	Экономическая коррупция
3	Игра принуждения по Штакельбергу с учетом требований к управляемой динамической системе	Управление устойчивым развитием путем принуждения
4	Игра побуждения по Штакельбергу с учетом требований к управляемой динамической системе	Управление устойчивым развитием путем побуждения
5	Игра принуждения по Штакельбергу с обратной связью по управлению и учетом требований к управляемой динамической системе	Управление устойчивым развитием путем принуждения при административной коррупции
6	Игра побуждения по Штакельбергу с дополнительной обратной связью по управлению и учетом требований к управляемой динамической системе	Управление устойчивым развитием путем побуждения при экономической коррупции

Динамические игры с требованиями к состоянию управляемой системы. В этом случае предполагается, что Ведущий должен обеспечить выполнение определенных требований к состоянию управляемой динамической системы (УДС), которые можно записать в виде $\forall t \in [0, T] \ x(t) \in X^*$ (устойчивость по Лагранжу), где X^* – множество допустимых состояний управляемой системы. Удобно записывать эти требования с помощью штрафного слагаемого в функции выигрыша Ведущего, а именно $g_L(p(t), q(t), u(t), x(t)) = \bar{g}_L(p(t), q(t), u(t), x(t)) - M\rho(x(t), X^*)$, где $\rho(x(t), X^*) = \begin{cases} 0, & x(t) \in X^*, \\ 1, & \text{иначе} \end{cases}$ – индикаторная функция, $M \gg 1$ – штрафная константа. Содержательно данный класс моделей используется применительно к описанию задачи управления устойчивым развитием [4,15]. Комбинируя указанные условия, можно получить содержательные частные классы моделей, наименования и интерпретация которых приведены в табл. 2.

3. Теоретико-игровые модели, интерпретация, принципы оптимальности

Опишем более детально названные выше частные классы динамических теоретико-игровых моделей иерархического управления.

3.1. Модель 1 – игра принуждения по Штакельбергу с обратной связью по управлению

В этом случае модель (2.1) - (2.4) примет вид:

$$J_L(q, u, b, x) = \int_0^\infty e^{-\alpha t} g_L(q(t), u(t), b(t), x(t)) dt \rightarrow \max \quad (3.1)$$

$$q(t) \in Q^{B(t)}; \quad (3.2)$$

$$J_F(u, b, x) = \int_0^\infty e^{-\alpha t} g_F(u(t), b(t), x(t)) dt \rightarrow \max \quad (3.3)$$

$$u(t) \in U(q(t)); \quad b(t) \in B(t); \quad (3.4)$$

Уравнение динамики (2.5) останется без изменений. Модель (2.5)–(3.4) интерпретируется как модель административной коррупции, где $q(t)$ – управление принуждения Ведущего («квота»); $u(t)$ – котируемое действие Ведомого; $b(t)$ – взятка, в обмен на которую квота может быть увеличена; $Q(t)$, $U(q(t))$, $B(t)$ – множества допустимых

управлений субъектов; $Q^{B(t)}$ – множество всех отображений из $B(t)$ в $Q(t)$; $x(t)$ – состояние управляемой системы (зависящее только от действий Ведомого).

Определение 3.1. *Тройка (q^*, u^*, b^*) называется равновесием принуждения по Штакельбергу в программных стратегиях с обратной связью по управлению, если*

$$J_L(q^*, u^*, b^*, x^*) = \sup_{q \in Q^B} \sup_{u \in U(q)} \inf_{b \in R_u(q)} J_L(q, u, b, x),$$

где $x^*, x \in X$ – соответствующие траектории УДС, принадлежащие множеству возможных состояний системы; $R_u(q) = \{b \in B(t) : \forall c \in B(t) J_F(u, b, x) \geq J_F(u, c, x)\}$ – множество оптимальных ответов Ведомого по переменной b на стратегию принуждения Ведущего q при фиксированной функции $u(t) = u(q(t), t) \in U(q)$.

Алгоритм построения равновесия модели (2.5)–(3.4) состоит в следующем:

1) Вводится стратегия наказания Ведомого Ведущим

$$q^P(t) : J_F(u(q^P(t), t), b(t), x(t)) = \inf_{q(t) \in Q} J_F(u(q(t), t), b(t), x(t))$$

Находится величина гарантированного выигрыша Ведомого, если он отказывается сотрудничать с Ведущим:

$$L_F = \sup_{b(t) \in B(t), u(t) \in U(q^P(t))} J_F(u(t), b(t), x(t));$$

2) Решается задача оптимального управления (2.5)–(3.2), (3.4) с дополнительным условием $L_F < J_F(u(t), b(t), x(t))$.

Максимум (3.1) с учетом предположения о сонаправленности интересов игроков ищется по трем функциям $q(t), u(t), b(t)$. Решение указанной задачи оптимального управления обозначим $q^R(t), u^R(t), b^R(t)$, где $q^R(t)$ – стратегия поощрения Ведомого Ведущим при принуждении.

3) Ведущий предъявляет Ведомому стратегию (коррупционный механизм управления) с обратной связью

$$\hat{q}(b(t), t) = \begin{cases} q^R(t), & \text{если } b(t) = b^R(t) \text{ для } \forall t \in [0, \infty) \\ q^P(t), & \text{иначе} \end{cases}$$

4) При экономически разумном Ведомом равновесие в модели (2.5)–(3.4) имеет вид

$$(q^R(t), u^R(t), b^R(t)).$$

Введем оператор $A_q : U(q) \rightarrow X$ согласно (2.5). Обозначим

$$A_q^{-1}(X) = \{u \in U(q) : A_q u \in X\}.$$

Теорема 3.1. *Равновесие модели (2.5)–(3.4) согласно определению 3.1 существует, если $S_1 = \{q(t) \in Q(t); b(t) \in B(t); u(t) \in A_q^{-1}(X) : L_F < J_F(u(t), b(t), x(u(t)))\} \neq \emptyset$ при заданной функции $p(t) \in P$.*

Доказательство. В силу свойств множеств $B(t), Q(t), U(q(t)), A_q^{-1}(X), X$ и функционала J_F получим, что если $S_1 \neq \emptyset$, то

$$\begin{aligned} & \sup_{(u(t), b(t), q(t)) \in S_1} J_F(u(q(t), t), b(t), x(t)) = \\ & = \max_{(u(t), b(t), q(t)) \in S_1} J_F(u(q(t), t), b(t), x(t)) = \\ & = J_F(u_1(q_1(t), t), b_1(t), x_1(t)); \\ & \quad (u_1(t), b_1(t), q_1(t)) \in S_1 \end{aligned}$$

Тройка функций $(u_1(t), b_1(t), q_1(t))$ есть равновесие модели (2.5)–(3.4) согласно определению 3.1. $\square$

3.2. Модель 2 – игра побуждения по Штакельбергу с дополнительной обратной связью по управлению

В этом случае модель (2.1)–(2.4) примет вид:

$$J_L(p, u, b, x) = \int_0^\infty e^{-\alpha t} g_L(p(t), u(t), b(t), x(t)) dt \rightarrow \max \quad (3.5)$$

$$p(t) \in P^{U(t)*B(t)}; \quad (3.6)$$

$$J_F(p, u, b, x) = \int_0^\infty e^{-\alpha t} g_F(p(t), u(t), b(t), x(t)) dt \rightarrow \max \quad (3.7)$$

$$u(t) \in U(t); \quad b(t) \in B(t); \quad (3.8)$$

Модель (3.5)–(3.8), (2.5) интерпретируется как модель экономической коррупции, где $p(t)$ – управление побуждения Ведущего («налог»); $u(t)$ – действие Ведомого, облагаемое налогом; $b(t)$ – взятка, в обмен на которую налог может быть уменьшен; $P^{U(t)*B(t)}$ – множество всех отображений из декартова произведения $U(t) \times B(t)$ в $P(t)$.

Определение 3.2. *Тройка (p^*, u^*, b^*) называется равновесием побуждения по Штакельбергу в программных стратегиях с дополнительной обратной связью по управлению, если*

$$J_L(p^*, u^*, b^*, x^*) = \sup_{p \in P^{U \times B}} \sup_{u \in U(t)} \inf_{b \in R_u(p)} J_L(p, u, b, x),$$

где $R_u(p) = \{b \in B(t) : \forall c \in B(t) \ J_F(p, u, b, x) \geq J_F(p, u, c, x)\}$ – множество оптимальных ответов Ведомого на стратегию побуждения Ведущего p при фиксированной функции $u(t) = u(t) \in U(t)$.

Алгоритм построения равновесия модели (3.5)–(3.8), (2.5) состоит в следующем.

1) Вводится стратегия наказания Ведомого Ведущим

$$p^P(t) : J_F(p^P(t), u(t), b(t), x(t)) = \inf_{p(t) \in P} J_F(p(t), u(t), b(t), x(t)); u(t) \in U(t).$$

Находится величина гарантированного выигрыша Ведомого, если он отказывается сотрудничать с Ведущим:

$$L_F = \sup_{b(t) \in B(t); u(t) \in U(t)} J_F(u(t), b(t), x(t)).$$

2) Решается задача оптимального управления (3.5), (3.6), (3.8), (2.5) с дополнительным условием $L_F < J_F(p(t), u(t), b(t), x(t))$. Максимум (3.5) ищется по трем функциям $p(t), u(t), b(t)$. Решение указанной задачи оптимального управления обозначим $p^R(t), u^R(t), b^R(t)$, где $p^R(t)$ – стратегия поощрения Ведомого Ведущим при побуждении.

3) Ведущий предъявляет Ведомому стратегию (механизм управления) с обратной связью

$$\hat{p}(u(t), b(t), t) = \begin{cases} p^R(t), & \text{если } b(t) = b^R(t); u(t) = u^R(t) \forall t \in [0, \infty) \\ p^P(t), & \text{иначе} \end{cases}$$

4) При экономически разумном Ведомом равновесие в модели (3.5)–(3.8), (2.5) имеет вид $(p^R(t), u^R(t), b^R(t))$.

Теорема 3.2. *Равновесие модели (3.5)–(3.8), (2.5) согласно определению 3.2 существует, если $S_2 = \{p(t) \in P(t); b(t) \in B(t); u(t) \in A_q^{-1}(X) : L_F < J_F(p(t), u(t), b(t), x(u(t)))\} \neq \emptyset$ при заданной функции $q(t) \in Q$.*

Доказательство этого утверждения аналогично доказательству утверждения 3.1.

3.3. Модель 3 – игра принуждения по Штакельбергу с учетом требований к управляемой динамической системе

Модель (2.1)–(2.4) принимает вид:

$$J_L(q, u, x) = \int_0^\infty e^{-\alpha t} g_L(q(t), u(t), x(t)) dt \rightarrow \max \quad (3.9)$$

$$q(t) \in Q(t); \quad (3.10)$$

$$J_F(u, x) = \int_0^\infty e^{-\alpha t} g_F(u(t), x(t)) dt \rightarrow \max \quad (3.11)$$

$$u(t) \in U(q(t)); \quad (3.12)$$

$$x(t) = x(u(t)) \in X^* \quad (3.13)$$

Модель (3.9)–(3.13), (2.5) интерпретируется как модель управления устойчивым развитием путем принуждения. Для описания требований устойчивого развития управляемой динамической системы вводится условие (3.13).

Определение 3.3. . Пара (q^*, u^*) называется равновесием принуждения по Штакельбергу в программных стратегиях с учетом требований к УДС, если

$$J_L(q^*, u^*, x^*) = \sup_{q \in Q_X} \inf_{u \in R_1(q)} J_L(q, u, x),$$

где $R_1(q) = \{u \in U(q(t)) : \forall w \in U(q(t)) J_F(u, x) \geq J_F(w, x)\}$ – множество оптимальных ответов Ведомого на стратегию принуждения Ведущего q ; $Q_X = \{q(t) \in Q(t) : \forall u(t) \in R_1(q(t)), x(u(t)) \in X^*, t \geq 0\}$ – множество допустимых управлений принуждения Ведущего с учетом требований к УДС. Если $Q_X = \emptyset$, то равновесие не существует.

Алгоритм построения равновесия модели (3.9)–(3.13), (2.5):

1) В результате оптимизации функционала (3.11) с ограничениями (3.12) определяются оптимальные стратегии Ведомого в зависимости от управлений Ведущего, т.е. функция $u^*(t) = u^*(q(t), t)$.

2) Найденная на первом шаге алгоритма функция $u^*(t) = u^*(q(t), t)$ подставляется в (3.9), (2.5). Решается задача оптимального управления (3.9), (2.5) с дополнительными условиями (3.10) и фазовыми ограничениями (3.13). Оптимальной для Ведущего является функция $q^*(t)$, которая доставляет максимум функционалу (3.9) при выполненных условиях (3.10), (3.13) с учетом (2.5).

3) Равновесие в модели 3 имеет вид $(q^*(t), u^*(q^*(t), t))$.

При наличии нескольких точек максимума у функции выигрыша Ведомого Ведущий проявляет осторожность в соответствии с принципом гарантированного результата Ю.Б. Гермейера.

Теорема 3.3. *Равновесие модели (3.9), (3.10), (3.13), (2.5) согласно определению 3.3 существует, если*

$$S_3 = \{q(t) \in Q(t) : u^*(p(t), q(t), t) \in A_q^{-1}(X^*)\} \neq \emptyset$$

при заданной функции $p(t) \in P$.

Доказательство. В силу свойств множеств $Q(t), U(q(t)), A_q^{-1}(X^*), X^*$ и функционала J_F получим, что если $S_3 \neq \emptyset$, то

$$\begin{aligned} \sup_{q(t) \in S_3} J_F(u(q(t), t), x(t)) &= \max_{q(t) \in S_3} J_F(u(q(t), t), x(t)) = \\ &= J_F(u_1(q_1(t), t), x_1(t)); q_1(t) \in S_3 \end{aligned}$$

Пара функций $(u_1(q_1(t), t), q_1(t))$ есть равновесие модели (3.9), (3.10), (3.13), (2.5) согласно определению 3.3. $\square$

3.4. Модель 4 – игра побуждения по Штакельбергу с учетом требований к управляемой динамической системе

Модель (2.1)–(2.4) принимает вид:

$$J_L(p, u, x) = \int_0^\infty e^{-\alpha t} g_L(p(t), u(t), x(t)) dt \rightarrow \max \quad (3.14)$$

$$p(t) \in P^{U(t)}; \quad (3.15)$$

$$J_F(p, u, x) = \int_0^\infty e^{-\alpha t} g_F(p(t), u(t), x(t)) dt \rightarrow \max \quad (3.16)$$

$$u(t) \in U(t) \quad (3.17)$$

Уравнение динамики (2.5) и условие (3.13) остаются без изменений. Модель (3.14)–(3.17), (2.5), (3.13) интерпретируется как модель управления устойчивым развитием путем побуждения. Здесь $P^{U(t)}$ – множество всех отображений из $U(t)$ в $P(t)$.

Определение 3.4. Пару (p^*, u^*) назовем равновесием побуждения по Штакельбергу в программных стратегиях с учетом требований к УДС, если

$$J_L(p^*, u^*, x^*) = \sup_{p \in P_X} \inf_{u \in R_1(p)} J_L(p, u, x),$$

где $R_1(p) = \{u \in U(t) : \forall w \in U(t) \ J_F(p, u, x) \geq J_F(p, w, x)\}$ – множество оптимальных ответов Ведомого на стратегию побуждения Ведущего p ;

$P_X = \{p(t) \in P^{U(t)} : \forall u(t) \in R_1(p(t)) \ x(u(t)) \in X^*; t \geq 0\}$ – множество допустимых управлений побуждения Ведущего с учетом требований к УДС. Если $P_X = \emptyset$, то равновесие не существует.

Алгоритм построения равновесия модели (3.13)–(3.17), (2.5):

1) В результате оптимизации функционала (3.16) с ограничениями (3.17) определяются оптимальные стратегии Ведомого в зависимости от управлений Ведущего, т.е. функция $u^*(t) = u^*(p(t), t)$.

2) Найденная на первом шаге алгоритма функция $u^*(t) = u^*(p(t), t)$ подставляется в (3.14), (2.5). Решается задача оптимального управления (3.14), (2.5) с дополнительными условиями (3.15) и фазовыми ограничениями (3.13). Оптимальной для Ведущего является функция $p^*(t)$, которая доставляет максимум (3.14) при выполненных условиях (3.13), (3.15) с учетом (2.5).

3) Равновесие в модели 4 имеет вид $(p^*(t), u^*(p^*(t), t))$.

При наличии нескольких точек максимума у функции выигрыша Ведомого Ведущий, как и в модели 3, проявляет осторожность в соответствии с принципом гарантированного результата Ю.Б. Гермейера.

Теорема 3.4. Равновесие модели (3.13)–(3.15), (2.5) согласно определению 3.4 существует, если

$$S_4 = \{p(t) \in P(t) : u^*(p(t), q(t), t) \in A_q^{-1}(X^*)\} \neq \emptyset$$

при заданной функции $q(t) \in Q$.

Доказательство этого утверждения аналогично доказательству утверждения 3.3.

3.5. Модель 5 – игра принуждения по Штакельбергу с обратной связью по управлению и учетом требований к управляемой динамической системе

Модель 5 описывается системой функционалов и ограничений на управления (2.5)–(3.4) с учетом дополнительного условия (3.13), отражающего требования устойчивого развития системы. Эта модель интерпретируется как модель управления устойчивым развитием путем принуждения в условиях административной коррупции.

Определение 3.5. *Тройка (q^*, u^*, b^*) называется равновесием принуждения по Штакельбергу в программных стратегиях с обратной связью по управлению и учетом требований к УДС, если*

$$J_L(q^*, u^*, b^*, x^*) = \sup_{(q,u) \in Q_X} \inf_{b \in R_u(q)} J_L(q, u, b, x),$$

где $Q_X = \{q(t) \in Q^{B(t)}; u(t) \in U(q(t)) \mid x(u(t)) \in X^*, t \geq 0\}$. Если $Q_X = \emptyset$, то равновесия не существует.

Алгоритм построения равновесия модели (2.5)–(3.4), (3.13):

1) Вводится стратегия наказания Ведомого Ведущим

$$q^P(t) : J_F(u(t), b(t), x(t)) = \inf_{q(t) \in Q} J_F(u(t), b(t), x(t)).$$

Находится величина гарантированного выигрыша Ведомого, если он отказывается сотрудничать с Ведущим:

$$L_F = \sup_{b(t) \in B(t) : u(t) \in U(q^P(t))} J_F(u(t), b(t), x(t)).$$

2) Решается задача оптимального управления (2.5)–(3.2), (3.4) с дополнительным условием $L_F < J_F(u(t), b(t), x(t))$ и фазовым ограничением (3.13).

Максимум (3.1) с учетом предположения о сонаправленности интересов игроков ищется сразу по трем функциям $q(t)$, $u(t)$, $b(t)$. Решение указанной задачи оптимального управления обозначим $q^R(t)$, $u^R(t)$, $b^R(t)$, где $q^R(t)$ – стратегия поощрения Ведомого Ведущим при принуждении.

3) Ведущий предъявляет Вedomому стратегию (механизм управления) с обратной связью

$$\hat{q}(b(t), t) = \begin{cases} q^R(t), & \text{если } b(t) = b^R(t) \text{ для } t \in [0, \infty) \\ q^P(t), & \text{иначе} \end{cases}$$

4) При экономически разумном Вedomом равновесие в модели (2.5)–(3.2), (3.4), (3.13) имеет вид $(q^R(t), u^R(t), b^R(t))$, где $u(t) \in R(q^R)$.

Теорема 3.5. *Равновесие модели (3.5), (3.6), (3.8), (2.5), (3.13) согласно определению 3.5 существует, если*

$$S_5 = \{q(t) \in Q(t); b(t) \in B(t); u(t) \in A_q^{-1}(X^*) :$$

$$L_F < J_F(u(t), b(t), x(u(t)))\} \neq \emptyset$$

при заданной функции $p(t) \in P$.

Доказательство этого утверждения аналогично доказательству утверждения 3.1.

3.6. Модель 6 – игра побуждения по Штакельбергу с дополнительной обратной связью по управлению и учетом требований к управляемой динамической системе

Модель 6 описывается системой функционалов и ограничений на управления (3.5)–(3.8), (2.5) с учетом дополнительного условия (3.13), отражающего требования устойчивого развития системы. Эта модель интерпретируется как модель управления устойчивым развитием путем побуждения в условиях экономической коррупции.

Определение 3.6. *Тройка (p^*, u^*, b^*) называется равновесием побуждения по Штакельбергу в программных стратегиях с дополнительной обратной связью по управлению и учетом требований к состоянию УДС, если*

$$J_L(p^*, u^*, b^*, x^*) = \sup_{(p, u) \in P_X} \inf_{b \in R_u(p)} J_L(p, u, b, x),$$

где $P_X = \{p(t) \in P^{U(t) \times B(t)}; u(t) \in U(t) x(u(t)) \in X^*, t \geq 0\}$. Если $P_X = \emptyset$, то равновесие не существует.

Алгоритм построения равновесия модели (3.5)–(3.8), (2.5), (3.13):

1) Вводится стратегия наказания Ведомого Ведущим

$$p^P(t) : J_F(p^P(t), u(t), b(t), x(t)) = \inf_{p(t) \in P(t)} J_F(p(t), u(t), b(t), x(t));$$

$$u(t) \in U(t).$$

Находится величина гарантированного выигрыша Ведомого, если он отказывается сотрудничать с Ведущим:

$$L_F = \sup_{b(t) \in B(t); u(t) \in U(t)} J_F(u(t), b(t), x(t)).$$

2) Решается задача оптимального управления (3.5), (3.6), (3.8), (2.5) с дополнительным условием $L_F < J_F(p(t), u(t), b(t), x(t))$ и фазовым ограничением (3.13). Максимум (3.5) ищется по трем функциям $p(t)$, $u(t)$, $b(t)$. Решение указанной задачи оптимального управления обозначим $p^R(t)$, $u^R(t)$, $b^R(t)$, где $p^R(t)$ – стратегия поощрения Ведомого Ведущим при побуждении.

3) Ведущий предъявляет Ведомому стратегию (механизм управления) с обратной связью

$$\hat{p}(u(t), b(t), t) = \begin{cases} p^R(t), & \text{если } b(t) = b^R(t); u(t) = u^R(t) \text{ для } t \in [0, \infty) \\ p^P(t), & \text{иначе} \end{cases}$$

4) При экономически разумном Ведомом равновесие в модели (3.5), (3.6), (3.8), (2.5), (3.13) имеет вид $(p^R(t), u^R(t), b^R(t))$.

Теорема 3.6. *Равновесие модели (3.5), (3.6), (3.8), (2.5), (3.13) согласно определению 3.6 существует, если*

$$S_6 = \{p(t) \in P(t); b(t) \in B(t); u(t) \in A_q^{-1}(X^*) :$$

$$L_F < J_F(p(t), u(t), b(t), x(u(t)))\} \neq \emptyset$$

при заданной функции $q(t) \in Q$.

Доказательство этого утверждения аналогично доказательству утверждения 3.1.

4. Система контроля качества речных вод

В качестве примера рассмотрим систему контроля качества поверхностных вод, включающую в свой состав: Ведущего (центр, ЦН), Ведомого (предприятие, ПП) и управляемую динамическую систему (водоток, УДС). Взаимоотношения между элементами исследуемой системы устроены следующим образом: ЦН воздействует на ПП, а ПП на УДС.

Предполагается, что принята следующая совокупность правил относительно поведения и информированности различных субъектов управления:

1) ЦН выбирает свою стратегию поведения первым (делает ход первым) и сообщает ее ПП; в моделях 3–6 он максимизирует свой критерий на множестве тех стратегий, которые позволяют поддерживать УДС в заданном состоянии.

2) ПП выбирает свою стратегию при известной стратегии ЦН и стремится к максимизации своего критерия, не обращая внимания на состояние УДС.

ПП в процессе своей производственной деятельности сбрасывает в водоток загрязняющие вещества (ЗВ), при этом ПП управляет степенью очистки сточных вод. ЦН определяет размер платы за сброс ПП загрязнений в водоток (метод побуждения) или минимально допустимые степени очистки сточных вод на ПП (метод принуждения). Требование поддержания системы в заданном состоянии состоит в не превышении предельно допустимых концентраций загрязняющих веществ (ЗВ) в речной системе. ПП предлагает ЦН взятку, в обмен на которую ЦН занижает размер платы за сброс загрязнений в водоток (при побуждении) или минимально допустимую степень очистки сточных вод (при принуждении).

Критерии субъектов управления имеют вид:

– ЦН

$$J_L(T, P) = \int_0^{\infty} e^{-\alpha t} (T(t) + b(t)) W(t)(1 - P(t)) dt \rightarrow \max \quad (4.1)$$

– ПП

$$J_F(T, P) = \int_0^{\infty} e^{-\alpha t} W(t) \left((T(t) + b(t))(1 - P(t)) + D \frac{P(t)}{1 - P(t)} \right) dt \rightarrow \min .$$

Здесь α – коэффициент дисконтирования; $P(t)$ – доля ЗВ, удаляемых из сточных вод на ПП в процессе их очистки в момент времени t ; $T(t)$ – размер платы за единицу сброшенных ЗВ; $W(t)$ ($W(t)(1 - P(t))$) – количество ЗВ, сбрасываемых в водоток до (после) очистки сточных вод; функция $W(t)$ считается заданной; $b(t)$ – величина взятки, приходящаяся на единицу сброшенных загрязнений; $D = \text{const}$.

Слагаемое $W(t)T(t)(1 - P(t))$ выражает плату ПП за сброс загрязнений в водоток в количестве $W(t)(1 - P(t))$ в момент времени t ; $D \frac{P(t)}{1-P(t)} W(t)$ – расходы на очистку сточных вод в количестве $W(t)$ в момент времени t со степенью $P(t)$; $W(t)b(t)(1 - P(t))$ – есть взятка, предлагаемая ЦН.

Ограничения на управления имеют вид:

$$0 \leq q(t) \leq 1 - \varepsilon; \quad 0 \leq T(t) \leq T_{\max}; \quad t \geq 0; \quad (4.2)$$

$$q(t) \leq P(t) \leq 1 - \varepsilon; \quad 0 \leq b(t) \leq b_{\max}; \quad t \geq 0; \quad (4.3)$$

где $q(t)$ – минимально допустимая степень очистки сточных вод на ПП; величина ε зависит от технологических возможностей очистки сточных вод на ПП ($0 \leq \varepsilon \leq 1$); $T_{\max}, b_{\max} = \text{const}$.

Условия устойчивого развития системы возьмем в виде

$$0 \leq E(t) \leq E_{\max}; \quad \forall t \geq 0 \quad (4.4)$$

где $E(t)$ – концентрация ЗВ в речной воде в момент времени t ; величина $E_{\max}$ – задана.

Пусть изменение концентрации ЗВ в речной системе описывается обыкновенным дифференциальным уравнением вида

$$\frac{dE}{dt} = F(E(t), P(t), t)$$

Функция $F(E(t), P(t), t)$ задана. Для простоты далее будем считать, что условие (4.4) выполнено при

$$P(t) \geq p_+(t); \quad 0 \leq p_+(t) \leq 1 - \varepsilon; \quad \forall t \geq 0, \quad (4.5)$$

где функция $p_+(t)$ – известна.

Исследование модели (4.1)–(4.5) проведем в случае, когда $\forall t \geq 0$ $D > T(t)$. Данное условие обеспечивает сонаправленность интересов ЦН и ПП.

Модель 1 (принуждение с обратной связью по управлению без учета условия (4.4)). Если выполнено условие $T(t)\varepsilon + D\frac{1-\varepsilon}{\varepsilon} > T(t) + b_{\max}$; $\forall t \geq 0$, то

$$q^P(t) = 1 - \varepsilon; \quad q^R(t) = p^R(t) = 0; \quad b^R(t) = b_{\max}.$$

В противном случае $b^R(t) = T(t)\varepsilon + D\frac{1-\varepsilon}{\varepsilon} - T(t) - \gamma$; $\forall t \geq 0$, $\gamma > 0$; $\gamma \ll 1$.

Модель 2 (побуждение с обратной связью по управлению без учета (4.4)). В этом случае $T^P(t) = T_{\max}$; $p^R(t) = 0$; $T^R(t) = T_{\max} - \gamma$; $b^R(t) = 0$; $\gamma > 0$; $\gamma \ll 1$.

В моделях 1, 2 наличие постоянной γ делает для ПП выгоднее стратегию его поощрения, по сравнению с наказанием.

Модель 3 (принуждение без обратной связи по управлению с учетом (4.4)). В этом случае $q^* = p_+$; $P^*(q^*) = q^*$.

Равновесие в модели 3 существует всегда при выполнении наложенных выше ограничений на входные параметры задачи.

Модель 4 (побуждение без обратной связи по управлению с учетом (4.4)). В этом случае $T^* = T_{\max}$; $P^* = 0$. Выполнить условие (4.4) не удастся.

Модель 5 (принуждение с обратной связью по управлению с учетом (4.4)). Если выполнено условие $T(t)\varepsilon + D\frac{1-\varepsilon}{\varepsilon} > T(t)(1 - p_+) + D\frac{p_+}{1-p_+} + b_{\max}(1 - p_+)$; $\forall t \geq 0$, то

$$q^P(t) = 1 - \varepsilon; \quad q^R(t) = p^R(t) = p_+; \quad b^R(t) = b_{\max}.$$

В противном случае $b^R(t) = T(t)\varepsilon + D\frac{1-\varepsilon}{\varepsilon} - T(t) - \gamma$; $\forall t \geq 0$, $\gamma > 0$; $\gamma \ll 1$.

Модель 6 (побуждение с обратной связью по управлению с учетом (4.4)). Если $T_{\max} - D\frac{p_+}{1-p_+} \geq 0$; $\forall t \geq 0$, то $b^R(t) = \min(b_{\max}; T_{\max} - \frac{T_{\max} - D\frac{p_+}{1-p_+}}{1-p_+} + \gamma)$;

$$T^P(t) = T_{\max}; \quad T^R(t) = \min(T_{\max}; \frac{T_{\max} - D\frac{p_+}{1-p_+}}{1-p_+} - \gamma);$$

$$\gamma > 0; \quad \gamma \ll 1; \quad p^R(t) = p_+.$$

В противном случае условие (4.4) выполнить не удастся, равновесие модели 6 не существует.

Пример 1. Для входных данных (у.е. – условные единицы):

$$D = 500; T_{\max} = 200 \text{ у.е.}; \alpha = 0.1; \gamma = 0.1; \varepsilon = 0.01; p_+(t) = 0.5;$$

$$W = 200 \text{ мг/л}; b_{\max} = 200 \text{ у.е.}$$

$T(t) = 50 \text{ у.е.}$ ($\forall t \geq 0$) (в случае принуждения) получены следующие результаты.

$$\text{Модель 1. } q^P(t) = 0.99; q^R(t) = p^R(t) = 0; b^R(t) = b_{\max} = 200 \text{ у.е.};$$

$$J_L = 4.1 \cdot 10^5 \text{ у.е.}; J_F = 4.1 \cdot 10^5 \text{ у.е.}$$

где J_L, J_F – величины средств, остающихся у ЦН и ПП, после учета всех их расходов.

$$\text{Модель 2. } T^P(t) = T_{\max} = 200 \text{ у.е.}; T^R(t) = T_{\max} - \gamma = 199.9 \text{ у.е.}; \\ p^R(t) = 0; b^R(t) = 0 \text{ у.е.};$$

$$J_L = 4 \cdot 10^5 \text{ у.е.}; J_F = 4 \cdot 10^5 \text{ у.е.}$$

$$\text{Модель 3. } q^* = 0.5; P^*(q^*) = 0.5. b(t) = 0 \text{ у.е.}$$

$$J_L = 5 \cdot 10^4 \text{ у.е.}; J_F = 1.05 \cdot 10^6 \text{ у.е.}$$

Модель 4. $T^* = T_{\max} = 200 \text{ у.е.}, P^*(T^*) = 0; b(t) = 0 \text{ у.е.}$ Условия (4.4) не выполнены, а

$$J_L = 4 \cdot 10^5 \text{ у.е.}; J_F = 4 \cdot 10^5 \text{ у.е.}$$

$$\text{Модель 5. } q^P(t) = 0.99; q^R(t) = p^R(t) = 0.5; b^R(t) = b_{\max} = 200 \text{ у.е.};$$

$$J_L = 2.5 \cdot 10^5 \text{ у.е.}; J_F = 1.25 \cdot 10^6 \text{ у.е.}$$

Модель 6. Равновесие не существует.

Пример 2. В случае входных данных примера 1 и $p_+(t) = 0.1$ (ослаблены условия устойчивого развития системы) равновесия моделей 1, 2, 4 не изменятся. В остальных случаях получим:

$$\text{Модель 3. } q^* = 0.1; P^*(q^*) = 0.1. b(t) = 0 \text{ у.е.}$$

$$J_L = 9 \cdot 10^4 \text{ у.е.}; J_F = 2.05 \cdot 10^5 \text{ у.е.}$$

$$\text{Модель 5. } q^P(t) = 0.99; q^R(t) = p^R(t) = 0.1; b^R(t) = b_{\max} = 200 \text{ у.е.};$$

$$J_L = 4.5 \cdot 10^5 \text{ у.е.}; J_F = 5.6 \cdot 10^5 \text{ у.е.}$$

Модель 6. Равновесие в отличие от примера 1 существует, причем

$$T^P(t) = T_{\max} = 200 \text{ у.е.}; T^R(t) = 160 \text{ у.е.}; p^R(t) = 0.1; b^R(t) = 0 \text{ у.е.};$$

$$J_L = 2.88 \cdot 10^5 \text{ у.е.}; J_F = 3.9 \cdot 10^5 \text{ у.е.}$$

По сравнению с примером 1 в моделях 3, 5 доход ЦН вырос, ПП – несколько снизился.

Пример 3. В случае входных данных примера 1 и $D = 1000$; $T = 200 \text{ у.е.}$; $T_{\max} = 400 \text{ у.е.}$ выросли затраты на очистку сточных вод) оптимальные стратегии субъектов управления не меняются по сравнению с примером 1. В моделях 4, 6, как и в примере 1, условия (4.4) выполнить не удастся. Доходы субъектов определяются формулами:

$$\text{Модель 1. } J_L = 8.01 \cdot 10^5 \text{ у.е.}; J_F = 8.01 \cdot 10^5 \text{ у.е.}$$

$$\text{Модель 2 } J_L = 7.99 \cdot 10^5 \text{ у.е.}; J_F = 7.99 \cdot 10^5 \text{ у.е.}$$

$$\text{Модель 3. } J_L = 4 \cdot 10^5 \text{ у.е.}; J_F = 2.4 \cdot 10^6 \text{ у.е.}$$

$$\text{Модель 4. } J_L = 1.2 \cdot 10^6 \text{ у.е.}; J_F = 1.2 \cdot 10^6 \text{ у.е.}$$

$$\text{Модель 5. } J_L = 4 \cdot 10^5 \text{ у.е.}; J_F = 2.4 \cdot 10^6 \text{ у.е.}$$

Для большинства моделей доходы ЦН и ПП выросли по сравнению с примером 1.

5. Заключение

В работе предложена модификация определения равновесия по Штакельбергу для некоторых частных классов динамических игр. Дана интерпретация соответствующих теоретико-игровых моделей в терминах приложений (моделирования коррупции и управления устойчивым развитием в иерархических системах управления). Разработаны и апробированы алгоритмы нахождения указанных решений теоретико-игровых моделей на основе дискретизации исходных дифференциальных игр. Приведены примеры построения равновесий сформулированных моделей в задаче контроля качества поверхностных вод для ряда входных данных.

В дальнейшем представляется целесообразным развитие проведенных исследований в следующих направлениях:

– рассмотрение динамических кооперативных теоретико-игровых моделей, формализующих переход от иерархических методов принуждения и побуждения к методу убеждения;

- изучение выделенных классов игр в позиционных стратегиях;
- исследование игр для трех и более игроков (уровней управления) и соответствующих регламентов сочетания различных игр Гермейера и методов иерархического управления;
- учет стохастических факторов;
- анализ различных «комбинированных» вариантов теоретико-игровых моделей.

СПИСОК ЛИТЕРАТУРЫ

1. Горелик В.А., Горелов М.А., Кононенко А.Ф. *Анализ конфликтных ситуаций в системах управления*. М.: Радио и связь, 1991.
2. Зенкевич Н.А., Петросян Л.А., Янг Д.В.К. *Динамические игры и их приложения в менеджменте*. СПб.: Изд-во «Высшая школа менеджмента», 2009.
3. Петросян Л.А., Зенкевич Н.А., Шевкопляс Е.В. *Теория игр*. СПб.: «БХВ-Петербург», 2012.
4. Угольников Г.А. *Иерархическое управление устойчивым развитием*. М.: Изд-во физико-математической литературы, 2010.
5. Угольников Г.А., Усов А.Б. *Устойчивое развитие систем управления с учетом коррупции* // Математическая теория игр и ее приложения. 2010. Т. 2. Вып. 4. С. 106–119.
6. Угольников Г.А., Усов А.Б. *Исследование дифференциальных моделей иерархических систем управления посредством их дискретизации* // Автоматика и телемеханика. 2013. № 2. С. 108–122.
7. Balafoutas L. *Public beliefs and corruption in a repeated psychological game* // Journal of Economic Behavior and Organization. 2011. № 78. P.51–59.
8. Basar T., Olsder G.J. *Dynamic Noncooperative Game Theory*. country-regionSIAM: CityplacePhiladelphia, 1999.

9. Basu K., Bhattacharya S., Mishra A. *Notes on bribery and the control of corruption* // Journal of Public Economics. 1992. № 48. P. 349–359.
10. Bhattacharya S., Hodler R. *Natural resources, bureaucracy and corruption* // European Economic Review. 2010. № 54. P. 608–621.
11. Blackburn K., Forgues-Puccio G.F. *Financial liberalization, bureaucratic corruption and economic development* // Journal of International Money and Finance. 2010. № 29. P. 1321–1339.
12. Cerqueti R., Coppier R. *Economic growth, corruption and tax evasion* // Economic Modeling. 2011. № 28. P. 489–500.
13. Lambert-Mogiliansky A., Majumdar M., Radner R. *Strategic analysis of petty corruption: Entrepreneurs and bureaucrats* // Journal of Development Economics. 2007. № 83. P. 351–367.
14. Olsen T. E., Torsvik G. *Collusion and Renegotiations in Hierarchies: A Case of Beneficial Corruption* // International Economic Review. 1998. Vol. 39. № 2. P. 143–157.
15. Ougolnitsky G. *Game theoretic formalization of the concept of sustainable development in the hierarchical control systems* // Annals of Operations Research. 2012. DOI 10.1007/s10479-012-1090-9.
16. Yang D. *Corruption by monopoly: Bribery in Chinese enterprise licensing as a repeated bargaining game* // China Economic Review. 2005. № 16. P. 171–188.

DYNAMIC HIERARCHICAL PLAY TWO PERSONS IN
PROGRAMME STRATEGY AND THEIR EXHIBITS

Gennady A. Ougolnitsky, South State University, Dr.Sc., prof.
(ougoln@mail.ru).

Anatoliy B. Usov, South State University, Dr.Sc., prof.
(usov@math.rsu.ru).

Abstract: The private classes of the differential hierarchical games of two players in programming strategies are considered in this article. Their interpretation is given as models to corruptions and controls of the sustainable development in hierarchical systems. The optimum principles for these classes of the games as modifications of the Stackelberg balance are formulate. The algorithms of the construction balance of these games are offered. These algorithms are approved on numeric examples.

Keywords: dynamic games, hierarchical systems, Stackelberg balance, corruption, sustainable development.