Автоматика и телемеханика, № 12, 2003

Управление в социально-экономических системах

© 2003 г. **Д. А. НОВИКОВ, д-р техн. наук,**
Т. Е. ШОХИНА
(Институт проблем управления им. В.А. Трапезникова РАН, Москва)

МЕХАНИЗМЫ СТИМУЛИРОВАНИЯ В ДИНАМИЧЕСКИХ АКТИВНЫХ СИСТЕМАХ

Исследуется задача стимулирования в динамической многоэлементной активной системе: дается решение задачи синтеза, изучается согласованность планов, назначаемых центром активным элементам.

1. Введение

Формальные (теоретико-игровые) модели активных систем (АС) исследуются в таких разделах теории управления социально-экономическими системами как: теория активных систем [1–5], теория иерархических игр [6, 7], теория контрактов [8–10] и др. Задача управления в базовой (статической) АС формулируется следующим образом. Управляющий орган – центр, обладающий правом первого хода, сообщает выбранное им управление активным элементам (АЭ), которые при известном управлении центра выбирают собственные стратегии с целью максимизации своих целевых функций. Цель центра заключается в том, чтобы, зная реакцию управляемых субъектов – АЭ – на те или иные управления, выбрать такое управление, которое привело бы систему в наиболее предпочтительное с его точки зрения состояние. Частным случаем задачи управления является задача стимулирования [3, 4], в которой центр осуществляет побочные платежи управляемым субъектам, зависящие от выборов (действий) последних.

Предметом исследования в настоящей работе являются механизмы стимулирования в динамических АС (ДАС), где под ДАС понимается АС, в которой последовательность выбора стратегий, характерная для статических АС, повторяется как минимум несколько раз. Интуитивно понятно, что при таком естественном обобщении простейшей базовой (статической) модели, как рассмотрение нескольких несвязанных периодов функционирования, задачу управления удается декомпозировать на набор базовых задач. Трудности появляются при исследовании ДАС со связанными периодами функционирования, в которых параметры некоторого (или каждого) периода зависят от действий участников в предыдущих периодах. Методы и алгоритмы решения задачи синтеза оптимального механизма управления в этом случае характеризуются высокой структурной и вычислительной сложностью. Как правило, универсального подхода к аналитическому решению этого класса задач найти не удается. Однако преодоление трудностей анализа оправданно, так как в ДАС присутствуют новые качественные свойства, отсутствующие в базовой модели (не говоря уже о том, что большинство реальных организационных систем функционируют в течение продолжительного времени и характеризуются относительной повторяемостью условий и самих фактов принятия решений). ДАС, функционирующие в течение длительного времени, существенно отличаются от статических: возможность долговременного сотрудничества, адаптации, пересмотра стратегий – все эти эффекты проявляются при переходе от статических моделей к динамическим.

Как отмечалось в обзоре [2] современных результатов исследования задач управления динамическими и многоэлементными АС, для этого класса задач на сегодняшний день отсутствуют аналитические решения. Поэтому в настоящей работе акцент сделан на получении именно аналитических решений для частного случая – задач стимулирования в детерминированных многоэлементных ДАС. Изложение материала имеет следующую структуру. Во втором разделе приводятся используемые в дальнейшем известные результаты исследования статической задачи стимулирования. В третьем разделе решается задача синтеза оптимальной системы стимулирования в одноэлементной ДАС. Четвертый раздел содержит обобщение полученных результатов на случай многоэлементной ДАС.

2. Задача стимулирования в многоэлементной статической активной системе

Рассмотрим базовую (многоэлементную детерминированную статическую) двухуровневую АС, состоящую из центра и n АЭ. Стратегией АЭ является выбор действий, стратегией центра – выбор функции стимулирования, т.е. зависимости вознаграждения каждого АЭ от его действий и, быть может, действий других АЭ.

Обозначим: $y_i \in A_i$ – действие i-го АЭ (нижний индекс здесь и далее обозначает номер АЭ), $i \in I = \{1, 2, \ldots, n\}$ – множество АЭ, $y = (y_1, y_2, \ldots, y_n) \in A' = \prod_{i=1}^{n} A_i$ – вектор действий АЭ, $y_{-i} = \{y_1, y_2, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n\} \in A_{-i} = \prod_{j \neq i} A_j$ – обстановка игры для i-го АЭ.

Интересы и предпочтения участников АС – центра и АЭ – выражены их целевыми функциями. Целевая функция центра является функционалом $\Phi(\sigma, y)$ и представляет собой разность между его доходом $H(y)$ и суммарным вознаграждением $\upsilon(y)$, выплачиваемым АЭ: $\upsilon(y) = \sum_{i=1}^{n} \sigma_i(y)$, где $\sigma_i(y)$ – стимулирование i-го АЭ, $\sigma(y) = (\sigma_1(y), \sigma_2(y), \ldots, \sigma_n(y))$, т.е.

$$\Phi(\sigma(\cdot), y) = H(y) - \sum_{i=1}^{n} \sigma_i(y). \tag{1}$$

Целевая функция i-го АЭ является функционалом $f_i(\sigma_i, y)$ и представляет собой разность между стимулированием, получаемым им от центра, и затратами $c_i(y)$, т.е.

$$f_i(\sigma_i(\cdot), y) = \sigma_i(y) - c_i(y), \quad i \in I. \tag{2}$$

Отметим, что индивидуальное вознаграждение и индивидуальные затраты i-го АЭ по выбору действия y_i в общем случае явным или неявным образом зависят от действий всех АЭ (случай сильно связанных АЭ с несепарабельными затратами [4]).

Примем следующий порядок функционирования АС. Центру и АЭ на момент принятия решения о выбираемых стратегиях (соответственно – функциях стимулирования и действиях) известны целевые функции и допустимые множества всех участников АС. Центр, обладая правом первого хода, выбирает функции стимулирования и сообщает их АЭ, после чего АЭ при известных функциях стимулирования выбирают действия, максимизирующие их целевые функции.

Относительно параметров АС введем следующие предположения.

А.1. $\forall i \in I \quad A_i$ – отрезок $\Re^1_+$ с левым концом в нуле.

А.2. $\forall i \in I \quad$ 1) функция $c_i(\cdot)$ непрерывна по всем переменным; 2) $\forall y \in A' \quad c_i(y)$ не убывает по y_i; 3) $\forall y \in A'$, $c_i(y) \geq 0$; 4) $\forall y_{-i} \in A_{-i}$, $c_i(0, y_{-i}) = 0$.

А.3. Функции стимулирования кусочно-непрерывны и принимают неотрицательные значения.

А.4. Функция дохода центра непрерывна и достигает максимума при ненулевом векторе действий АЭ.

Обозначим через $P(\sigma)$ множество равновесных по Нэшу [11] при системе стимулирования σ действий АЭ – множество реализуемых действий (т.е. будем считать, что АЭ выбирают свои стратегии одновременно и независимо друг от друга, не имея возможности обмениваться дополнительной информацией и полезностью). Минимальными затратами центра на стимулирование по реализации вектора действий АЭ $y' \in A'$ будем называть минимальное значение суммарных выплат элементам, при которых данный вектор действий является равновесием Нэша их игры, т.е. решение следующей задачи:

$$\sum_{i \in I} \sigma_i(y') \to \min_{\sigma(\cdot) \in \Xi(y')},$$

где $\Xi(y') = \{\sigma(\cdot) | y' \in P(\sigma)\}$. Как и в одноэлементной АС [3], гарантированной эффективностью (далее просто "эффективностью") стимулирования является минимальное значение целевой функции центра на соответствующем множестве решений игры (всюду, где встречаются минимумы и максимумы, будем предполагать, что они достигаются):

$$K(\sigma(\cdot)) = \min_{y \in P(\sigma(\cdot))} \Phi(\sigma(\cdot), y). \tag{3}$$

Задача синтеза оптимальной функции стимулирования заключается в поиске допустимой системы стимулирования σ^*, имеющей максимальную эффективность (3):

$$\sigma^* = \arg \max_{\sigma(\cdot)} K(\sigma(\cdot)). \tag{4}$$

В [4, 12] доказано, что в рамках предположений А.1–А.4 оптимальной в смысле (4) (точнее – δ-оптимальной, где $\delta = \sum_{i=1}^{n} \delta_i$) является компенсаторная система стимулирования σ_K:

$$\sigma_{iK}(y_i^*, y) = \begin{cases} c_i(y_i^*, y_{-i}) + \delta_i, & y_i = y_i^* \\ 0, & y_i \neq y_i^* \end{cases}, \quad i \in I, \tag{5}$$

где δ_i – сколь угодно малые строго положительные константы, а оптимальный вектор действий y^*, реализуемый системой стимулирования (5) как единственное равновесие в доминантных стратегиях [11] игры АЭ, является решением следующей задачи оптимального согласованного планирования:

$$y^* = \arg \max_{y \in A'} \left\{ H(y) - \sum_{i=1}^{n} c_i(y) \right\}. \tag{6}$$

Приведенный результат об оптимальности компенсаторных систем стимулирования (5)–(6), получивший в [4] название "принципа компенсации затрат", позволяет сразу определить минимальные затраты центра на стимулирование по реализации вектора действий $y \in A'$:

$$(7) \qquad \upsilon(y) = \sum_{i=1}^{n} c_i(y_i),$$

и сконцентрировать все внимание на решении задач выбора оптимальных для центра реализуемых действий – см. (6).

Кроме того, результаты анализа статической модели позволяют сделать еще один вывод. Система стимулирования (5) побуждает АЭ выбирать соответствующие действия как доминантные стратегии, т.е. осуществляет декомпозицию игры АЭ (в [4] также доказано, что при использовании центром управлений (5)–(6) вектор действий y^* является сильным равновесием Нэша [11] игры АЭ), что позволяет не рассматривать взаимодействие АЭ, а решать задачи их стимулирования "независимо". Использование управлений, декомпозирующих игру АЭ, в [4] получило название "принципа декомпозиции игры АЭ".

Как будет видно из последующего изложения, использование принципов декомпозиции и компенсации затрат эффективно и при исследовании ДАС. Следует отметить, что в рамках рассматриваемой модели принцип компенсации затрат и принцип декомпозиции применимы при условии, что центру точно известны функции затрат АЭ и выбираемые ими действия. Приведенное в [4, 13] обобщение этих принципов на случай, когда функции затрат АЭ известны центру неточно (имеет место асимметричная информированность) или когда он наблюдает некоторый агрегат от действий АЭ, также может быть использовано в ДАС.

Перейдем к описанию задач стимулирования в ДАС.

3. Задача стимулирования в одноэлементной динамической активной системе

Рассмотрим модель одноэлементной ДАС, взаимодействие участников (центра и единственного АЭ) в которой является совокупностью T повторений их взаимодействия в одноэлементной статической модели (см. второй раздел), т.е. центр в каждом периоде $t = \overline{1,T}$ сообщает АЭ управление $\sigma^t(\cdot)$ на этот период (верхний индекс здесь и далее обозначает номер периода), после чего АЭ выбирает действие y^t, причем все параметры модели АС текущего периода в общем случае зависят от всех параметров прошлых периодов.

Относительно информированности участников ДАС будем предполагать, что центру в первом периоде известны все параметры всех периодов функционирования (т.е. он полностью информирован и принимает решения по выбору программной траектории ДАС), АЭ в каждом периоде обладает информацией только о параметрах этого периода.

Пусть $y^t \in A^t(y^{1,t-1})$ – действие АЭ в периоде t, где $y^{1,t-1} = (u^1, y^2, \ldots, y^{t-1})$, $\sigma^t(\cdot) : A^{1,t} \to \Re^+_1$ – используемая центром в этом периоде система стимулирования, $A^{1,t} = \prod_{\tau=1}^{t} A^\tau$, $t = \overline{1,T}$. Относительно параметров ДАС будем предполагать, что параметры каждого периода удовлетворяют предположениям А.1–А.4.

Выигрыш АЭ в периоде t равен (ср. с (2))

$$(8) \qquad f^t(\sigma^t, y^t) = \sigma^t(y^t) - c^t(y^t), \quad t = \overline{1,T},$$

где $c^t(\cdot)$ – функция затрат АЭ в этом периоде.

Выигрыш центра в периоде t равен (ср. с (1))

$$\Phi^t(\sigma^t, y^t) = H^t(y^t) - \sigma^t(y^t), \tag{9}$$

где $H^t(\cdot)$ – функция дохода центра в этом периоде, $t = \overline{1,T}$.

Будем считать, что центр должен выбирать такие управления, чтобы в каждом периоде значение целевой функции АЭ было неотрицательно, т.е. $f^t(\sigma^t, y^t) \geq 0$, $t = \overline{1,T}$ (условие участия или условие индивидуальной рациональности – Individual Rationality [2]).

Если в каждом периоде целевые функции и допустимые множества удовлетворяют предположениям А.1–А.4 (т.е. периоды несвязанны), то в соответствии с принципом компенсации затрат задача центра заключается в последовательном определении и реализации плановой траектории $x^{1,T} = (x^1, x^2, \ldots, x^T)$ как результата решения следующей совокупности независимых задач оптимального согласованного планирования:

$$x^t = \arg\max_{y^t \in A^t} \left\{ H^t(y^t) - c^t(y^t) \right\}, \quad t = \overline{1,T}. \tag{10}$$

Если целевая функция центра определяется суммой (по всем периодам) значений его выигрышей (9), то задача оптимального согласованного планирования имеет вид:

$$x^{1,T} = \arg\max_{y^{1,T} \in A_0^{1,T}} \sum_{t=1}^{T} \left\{ H^t(y^t) - c^t(y^t) \right\}, \tag{11}$$

где $A_0^{1,\tau} = \left\{ y^{1,\tau} \in A^{1,\tau} | y^t \in A^t(y^{1,t-1}),\ t = \overline{1,\tau} \right\}$, $\tau = \overline{1,T}$.

Очевидно, что при несвязанных периодах функционирования ($c^t = c^t(y^t)$, $H^t = H^t(y^t)$, $A^t(y^{1,t-1}) = A^t$) решение задачи (11) разбивается на решение T несвязанных одноопериодных задач оптимального согласованного планирования (10), а решения задач (11) и (10) совпадают, что объясняется независимостью периодов.

Если периоды слабо связаны (т.е. существует единственное ограничение, связывающее действия, или множества допустимых действий, или затраты, или доходы, или вознаграждения и т.д. – см. аналогии в задачах стимулирования в многоэлементных АС со слабо связанными АЭ [4]), то задача (11) превращается в задачу условной оптимизации (изменяется множество действий, по которому ищется максимум) [2]. Основная идея решения задачи стимулирования в этом классе моделей заключается в том, чтобы "перенести" все ограничения на множество допустимых траекторий, а затем решать задачу выбора оптимальной (по критерию суммарного выигрыша центра) допустимой (с учетом всех ограничений) траектории в расширенном пространстве состояний [5]. Например, если наложено ограничение R на суммарные выплаты АЭ, то, вводя множество $P(R)$ реализуемых при данном ограничении действий АЭ:

$$P(R) = \left\{ y^{1,T} \in A^{1,T} \middle| \sum_{t=1}^{T} c^t(y^t) \leq R \right\},$$

получаем, что оптимальной будет плановая траектория, являющаяся решением следующей задачи:

$$x^{1,T} = \arg\max_{y^{1,T} \in P(R)} \sum_{t=1}^{T} \left\{ H^t(y^t) - c^t(y^t) \right\}.$$

При этом, очевидно, это решение в общем случае не будет являться совокупностью T решений задач (10).

Если стимулирование АЭ в каждом периоде зависит как от его действия в этом периоде, так и от его действий во всех предыдущих периодах, т.е. $\sigma^t = \sigma^t(y^{1,t})$ (будем называть такое стимулирование связанным, в отличие от несвязанного стимулирования, при котором вознаграждение АЭ в каждом периоде может зависеть только от действий в данном периоде), то в соответствии с принципом компенсации затрат построим систему стимулирования

$$\sigma_K^t\left(x^{1,T}, y^{1,t}\right) = \begin{cases} c^t(x^{1,t}), & \text{если} \quad y^i = x^i, \ i = \overline{1,t} \\ 0, & \text{в остальных случаях} \end{cases}, \qquad t = \overline{1,T}. \tag{12}$$

Пусть теперь затраты АЭ в периоде t зависят от всей предыстории $y^{1,t-1}$ и действия y^t, выбранного в этом периоде, $t = \overline{1,T}$, то есть $c^t = c^t(y^{1,t})$. Также будем считать, что доход центра в каждом периоде зависит в общем случае от всей истории: $H^t = H^t(y^{1,t})$, $t = \overline{1,T}$. Введем следующее предположение относительно свойств функции затрат АЭ.

А.2′. $\forall\, t = \overline{1,T}$ 1) функция $c^t(\cdot)$ непрерывна по всем переменным; 2) $\forall y^{1,t} \in A^{1,t}$ $c^t(y^{1,t})$ не убывает по y^t, $t = \overline{1,T}$; 3) $\forall y^{1,t} \in A^{1,t}$ $c^t(y^{1,t}) \geq 0$, $t = \overline{1,T}$; 4) $\forall y^{1,t-1} \in A^{1,t-1}$ $c^t(y^{1,t-1}, 0) = 0$, $t = \overline{1,T}$.

Теорема 1. *Если выполнены предположения А.1, А.2′, А.3 и А.4, то при использовании центром системы стимулирования* (12)*, где оптимальная плановая траектория определяется*

$$x^{1,T} = \arg \max_{y^{1,T} \in A_0^{1,T}} \sum_{t=1}^{T} \left\{ H^t(y^{1,t}) - c^t(y^{1,t}) \right\}, \tag{13}$$

действия АЭ совпадут с планами и эффективность стимулирования будет максимально возможной при связанном стимулировании и равной

$$\max_{y^{1,T} \in A_0^{1,T}} \sum_{t=1}^{T} \left\{ H^t(y^{1,t}) - c^t(y^{1,t}) \right\}.$$

Доказательство теоремы 1 приведено в Приложении. Обсудим результат теоремы 1. Очевидно, что, во-первых, в соответствии с (12) центр может не запоминать, какие действия выбирает АЭ в каждом периоде – ему необходимо лишь знать, отклонялся ли АЭ в прошлом хотя бы раз от планов или нет. Во-вторых, в силу полной дальновидности центра результат теоремы 1 справедлив для любого режима управления АЭ со стороны центра, т.е. центр может как сообщать АЭ всю информацию (12)–(13) до начала первого периода, так и в каждом периоде сообщать только управление для этого периода и/или на любое число будущих периодов (см. аналогии в [5]).

Содержательно в силу (12) центр может в любом (в том числе – в каждом) периоде наказать АЭ за отклонения от плана в прошлом. Рассмотрим случай, в котором такая возможность у центра отсутствует – при несвязанном стимулировании он может в любом периоде поощрять АЭ только за действия, выбранные в этом периоде. Другими словами, рассмотрим случай, когда затраты АЭ несепарабельны по периодам, а вознаграждение АЭ в каждом периоде может зависеть только от его действий в этом периоде, т.е. $\sigma^t = \sigma^t(y^t)$ и центр использует систему стимулирования

$$\sigma_K^t(x^t, y^t) = \begin{cases} c^t(x^t), & \text{если} \quad y^t = x^t \\ 0 & \text{в остальных случаях} \end{cases}, \qquad t = \overline{1,T}. \tag{14}$$

Если вознаграждение АЭ в любом периоде может зависеть от его действий, выбранных только в этом периоде (ср. (12) и (14)), то результат теоремы 1 уже не имеет места – см. теорему 2. Если АЭ недальновиден, или если его затраты не связаны, то оптимальна и реализуема плановая траектория (13). Отличие появляется при использовании центром программного управления (12), т.е. сообщения дальновидному АЭ со связанными затратами до начала первого периода сразу всей (или части) плановой траектории и всех (или части) зависимостей вознаграждения от действий. Оказывается, что при связанных затратах и несвязанном стимулировании множество реализуемых траекторий не шире, а эффективность стимулирования не выше, чем при связанном стимулировании.

Теорема 2. *Если выполнены предположения А.1, А.2′, А.3 и А.4, то при использовании центром системы стимулирования* (14) *и оптимальной плановой траектории* $x^{1,T} = \arg \max_{y^{1,T} \in X^{1,T}} \sum_{t=1}^{T} \left\{ H^t(y^{1,t}) - c^t(y^{1,t}) \right\}$, *где*

$$
(15) \qquad X^{1,T} = \left\{ x^{1,T} \in A_0^{1,T} | \forall \ y^{1,T} \in A_0^{1,T} \sum_{\substack{t=1 \\ x^t \neq y^t}}^{T} c^t(y^{1,t}) \geq \sum_{\substack{t=1 \\ x^t = y^t}}^{T} \left[c^t(x^{1,t}) - c^t(y^{1,t-1}, x^t) \right] \right\},
$$

действия АЭ совпадут с планами и эффективность стимулирования будет максимально возможной при несвязанном стимулировании.

Доказательство теоремы 2 приведено в Приложении. Содержательно множество $X^{1,T}$, определяемое выражением (15), может интерпретироваться как множество согласованных планов.

Отметим, что, если центр может использовать следующую систему стимулирования, являющуюся более “мягкой”, чем (12), но более жесткую, чем (14):

$$
(16) \qquad \sigma_K^t(x^{1,T}, y^{1,t}) = \begin{cases} c^t(y^{1,t-1}, x^t), & \text{если} \quad y^t = x^t \\ 0 & \text{в остальных случаях} \end{cases}, \qquad t = \overline{1,T},
$$

то реализуема любая траектория из $A_0^{1,T}$, но при этом в соответствии с (16) выплаты АЭ в текущем периоде зависят уже от всей предыстории (в отличие от (12)). Это утверждение сформулируем в виде следствия из теорем 1 и 2:

Следствие. *Системы стимулирования* (12) *и* (16) *характеризуются максимальным множеством реализуемых действий и максимальной эффективностью.*

Содержательно, при использовании системы стимулирования (12) центр отслеживает отклонения АЭ от плана в течение всей предыстории (по отношению к рассматриваемому периоду) и выплачивает АЭ ненулевое вознаграждение (компенсирует ему затраты), только если он ни разу не отклонился от плана. В соответствии с (16) центр может не “помнить” отклонения, а компенсировать в каждом периоде затраты АЭ при выполнении им плана в этом периоде с учетом фактически сложившейся истории. Легко видеть, что при этом АЭ не может получить в текущем периоде выигрыша за счет отклонений в предыдущих периодах, т.е. любой план будет согласованным.

Получив решение задачи стимулирования в одноэлементной ДАС, обобщим результаты теорем 1 и 2 на случай многоэлементных ДАС, используя результаты второго раздела.

4. Задача стимулирования в многоэлементной динамической активной системе

Рассмотрим многоэлементную динамическую модель – ДАС с n АЭ, стратегией каждого из которых в каждом периоде является выбор (при известном управлении со стороны центра) действия $y_i^t \in A_i^t$, $i \in I = \{1, 2, \ldots, n\}$, $t = \overline{1,T}$. Обозначим:

$y^t = (y_1^t, y_2^t, \ldots, y_n^t)$ – вектор действий всех АЭ в момент времени t,

$y^{1,T} = (y^1, y^2, \ldots, y^T)$ – вектор действий всех АЭ за все T периодов.

Пусть $\sigma_i^t = \sigma_i^t(y^{1,t})$, $c_i^t = c_i^t(y^{1,t})$, $H^t = H^t(y^{1,t})$, $A_i^t = A_i^t(y^{1,t-1})$, $i \in I$, $t = \overline{1,T}$. Определим

$$A^t = \prod_{i \in I} A_i^t, \quad A_{-i}^t = \prod_{j \neq i} A_j^t, \quad A^{1,\tau} = \prod_{t=1}^{\tau} A^t,$$
$$A_0^{1,\tau} = \left\{ y^{1,\tau} \in A^{1,\tau} | y^t \in A^t(y^{1,t-1}),\ t = \overline{1,\tau} \right\}, \quad \tau = \overline{1,T}.$$

Введем дополнительное предположение относительно свойств функций затрат АЭ (отметим, что данное предположение является “объединением” предположений А.2 и А.2$'$, отражающих свойства функций затрат соответственно в статической многоэлементной АС и в одноэлементной ДАС).

А.2$''$. $\forall\ t = \overline{1,T}$, $i \in I$ 1) функция $c_i^t(\cdot)$ непрерывна по всем переменным; 2) $\forall\ y^{1,t} \in A^{1,t}$ $c_i^t(y^{1,t})$ не убывает по y_i^t; 3) $\forall\ y^{1,t} \in A^{1,t}$ $c_i^t(y^{1,t}) \geq 0$; 4) $\forall\ y^{1,t-1} \in A^{1,t-1}$, $\forall\ y_{-i}^t \in A_{-i}^t$ $c^t(y^{1,t-1}, y_{-i}^t, 0) = 0$.

Теорема 3. *Если выполнены предположения А.1, А.2$''$, А.3 и А.4, то при использовании центром системы стимулирования*

$$\sigma_{iK}^t(x^{1,T}, y^{1,t}) = \begin{cases} c_i^t(x_i^{1,t}, y_{-i}^{1,t}), & \text{если } y_i^k = x_i^k,\ k = \overline{1,t} \\ 0 & \text{в остальных случаях} \end{cases}, \quad t = \overline{1,T},\ i \in I, \tag{17}$$

где

$$x^{1,T} = \arg \max_{y^{1,T} \in A_0^{1,T}} \sum_{t=1}^{T} \left\{ H^t(y^{1,t}) - \sum_{i \in I} c_i^t(y^{1,t}) \right\},$$

действия АЭ совпадут с планами и эффективность стимулирования будет максимально возможной.

Идея доказательства теоремы 3 заключается в следующем. В [4] был введен принцип декомпозиции игры АЭ в задачах стимулирования (см. также второй раздел), заключающийся в том, что при использовании в многоэлементных АС компенсаторных систем стимулирования, в которых АЭ компенсировались затраты в случае выбора им соответствующей плановой компоненты (независимо от действий других АЭ!), выбор действий, совпадающих с планами, является доминантной стратегией каждого АЭ. Если выполнено (17), то, применяя принцип декомпозиции, получаем возможность независимо рассматривать n задач управления несвязанными между собой АЭ. Для каждой из этих задач в отдельности применима теорема 1. Кроме того, для многоэлементных ДАС в предположении, что АЭ в каждом периоде выбирают равновесные по Нэшу стратегии, справедливы аналоги теоремы 2 и следствия 1.

Результаты использования теорем 1–3 в прикладных моделях управления ДАС описаны в [14].

5. Заключение

В настоящей работе получено решение задачи синтеза оптимальной функции стимулирования в одноэлементной ДАС для связанного (теорема 1) и несвязанного (теорема 2) стимулирования, а также для многоэлементных ДАС (теорема 3). При этом предполагалось, что центр реализует программный режим управления. Поэтому перспективным направлением дальнейших исследований представляется изучение сравнительной эффективности различных (текущих, скользящих, программных и др. [2]) режимов управления, а также исследование влияния дальновидности участников ДАС на эффективность управления.

ПРИЛОЖЕНИЕ

Доказательство теоремы 1. Задача стимулирования заключается в выборе центром системы стимулирования $\sigma^{1,T}(\cdot)$, которая максимизировала бы его целевую функцию, учитывающую в силу полной дальновидности центра все T периодов.

$$(\text{П.1}) \qquad \Phi\left(\sigma^{1,T}(\cdot), y_*^{1,T}\right) = \sum_{t=1}^{T} \left\{H^t(y_*^{1,t}) - \sigma^t(y_*^{1,t})\right\}$$

при условии, что действия АЭ $y_*^{1,T}$, выбираемые им при известной системе стимулирования, максимизируют его целевую функцию $f\left(\sigma^{1,T}(\cdot), y^{1,T}\right)$, учитывающую в силу полной дальновидности АЭ все T периодов, т.е.

$$(\text{П.2}) \qquad y_*^{1,T} \in \operatorname{Arg} \max_{y^{1,T} \in A_0^{1,T}} \sum_{t=1}^{T} \left\{\sigma^t(y_*^{1,t}) - c^t(y_*^{1,t})\right\}.$$

Фиксируем произвольную плановую траекторию $z^{1,T} \in A_0^{1,T}$. Пусть некоторая система стимулирования $s^{1,T}(\cdot)$ реализует эту плановую траекторию, т.е.

$$(\text{П.3}) \qquad \sum_{t=1}^{T} \left\{s^t(z^{1,t}) - c^t(z^{1,t})\right\} \geq \sum_{t=1}^{T} \left\{s^t(y^{1,t}) - c^t(y^{1,t})\right\}, \quad \forall y^{1,T} \in A_0^{1,T}.$$

Перейдем от системы стимулирования $s^{1,T}(\cdot)$ к соответствующей квазисистеме стимулирования [1] $qs^{1,T}(\cdot)$ следующим образом:

$$(\text{П.4}) \qquad qs^t(y^{1,t}) = \begin{cases} s^t(z^{1,t}), & y^{1,t} = z^{1,t} \\ 0, & y^{1,t} \neq z^{1,t} \end{cases}, \qquad t = \overline{1,T}.$$

Если заменить в выражении (П.3) $s^{1,T}(\cdot)$ на $qs^{1,T}(\cdot)$, то система неравенств останется в силе, т.е. плановая траектория $z^{1,T}$ будет реализовываться и системой стимулирования $qs^{1,t}(\cdot)$, а фактические выплаты АЭ не изменятся.

Таким образом, доказано, что без потери эффективности можно ограничиться классом систем стимулирования типа (П.4), которому в том числе принадлежит система стимулирования (12).

Фиксируем произвольную плановую траекторию $z^{1,T} \in A_0^{1,T}$. Из (П.1) и (П.3) следует, что при фиксированной плановой траектории центр стремится найти реализующую ее систему стимулирования, которая обладала бы минимальными затратами на стимулирование, т.е. центр решает следующую задачу:

$$(\text{П}.5)\qquad \sum_{t=1}^{T} s^t(z^{1,t}) \to \min,$$

$$(\text{П}.6)\qquad \sum_{t=1}^{T} \left\{ s^t(z^{1,t}) - c^t(z^{1,t}) \right\} \geq - \sum_{t=1}^{T} c^t(y^{1,t}), \quad \forall y^{1,T} \in A_0^{1,T}.$$

Из предположения А.2$'$ следует, что максимум правой части выражения (П.6) достигается в том числе при нулевых действиях АЭ и равен нулю. Кроме того, выше предполагалось, что центр должен в каждом периоде обеспечить АЭ неотрицательную полезность, т.е. каждое из слагаемых в левой части выражения (П.6) неотрицательно. Следовательно, одно из решений задачи (П.5)–(П.6) имеет вид

$$(\text{П}.7)\qquad s^t(z^{1,t}) = c^t(z^{1,t}), \quad t = \overline{1,T}.$$

Значит, минимальная система стимулирования, реализующая плановую траекторию $z^{1,T}$, удовлетворяет одновременно (П.4) и (П.7), что дает выражение (12). При этом значение целевой функции АЭ в каждом периоде неположительно, а при выборе действий, совпадающих с планами, равно нулю.

То, что АЭ при использовании центром управления (12)–(13) выберет действия, совпадающие с планами, следует из подстановки (12) в (П.3) – если в любом из периодов АЭ выбирает действия, отличающиеся от планов, то значение его целевой функции не увеличивается (для того, чтобы планы были единственными точками максимума, достаточно доплачивать АЭ за их выбор, помимо компенсации затрат, сколь угодно малую, но строго положительную величину – см. второй раздел и [3]).

Суммируя (П.7) по всем периодам, получим следующую оценку минимальных затрат $\upsilon(\cdot)$ на реализацию плановой траектории $z^{1,T}$ (ср. с (7)):

$$(\text{П}.8)\qquad \upsilon(z^{1,T}) = \sum_{t=1}^{T} c^t(z^{1,t}).$$

Таким образом, мы показали, что системы стимулирования вида (П.4), (П.7) реализуют плановую траекторию $z^{1,t}$ с минимальными затратами центра на стимулирование, определяемыми (П.8). Вспоминая, что плановая траектория выбиралась произвольной, получаем, что необходимо найти плановую траекторию, которая максимизировала бы разность между $\sum_{t=1}^{T} H^t(z^{1,T})$ и $\upsilon(z^{1,T})$, что и отражено выражением (13).

Доказательство теоремы 2. Отметим, что формулировка теоремы 2 отличается от формулировки теоремы 1 только видом системы стимулирования (ср. (14) и (12)) и тем множеством траекторий, по которому ведется максимизация при определении оптимальной плановой траектории (ср. (13) и (15)).

Невозможность реализации произвольной плановой траектории системой стимулирования (14) обусловлена тем, что, выбирая в некотором периоде действия, отличные от планов, в случае связанных затрат АЭ может в общем случае изменить свои затраты в будущих периодах, а центр не имеет возможности в текущем периоде наказать АЭ за отклонения в прошлых периодах.

Система неравенств (15) отражает невыгодность отклонения АЭ от плана. Действительно, при отклонениях АЭ несет потери, фигурирующие в левой части (суммирование ведется по тем периодам, в которых планы не выполнялись), в правой части стоит выигрыш от отклонений. Если потери превышают выигрыш, то отклонение невыгодно.

Итак, выражение (15) определяет множество плановых траекторий, реализация которых выгодна для АЭ (точнее – невыгодно отклонение от них). В остальном доказательство теоремы 2 повторяет доказательство теоремы 1 и опускается.

СПИСОК ЛИТЕРАТУРЫ

1. *Бурков В.Н., Новиков Д.А.* Теория активных систем: состояние и перспективы. М.: Синтег, 1999.
2. *Новиков Д.А.* Механизмы стимулирования в динамических и многоэлементных социально-экономических системах // АиТ. 1997. № 6. С. 3–26.
3. *Новиков Д.А.* Стимулирование в социально-экономических системах (базовые математические модели). М.: ИПУ РАН, 1998.
4. *Новиков Д.А., Цветков А.В.* Механизмы стимулирования в многоэлементных организационных системах. М.: ИПУ РАН, 2001.
5. *Щепкин А.В.* Динамические активные системы с дальновидными элементами. I, II // АиТ. 1986. № 10. С. 89–94; № 11. С. 82–94.
6. *Гермейер Ю.Б.* Игры с непротивоположными интересами. М.: Наука, 1976.
7. *Кононенко А.Ф., Халезов А.Д., Чумаков В.В.* Принятие решений в условиях неопределенности. М.: ВЦ АН СССР, 1991.
8. *Fudenberg D., Tirole J.* Game theory. Cambridge: MIT Press, 1995.
9. *Myerson R.B.* Game theory: analysis of conflict. London: Harvard Univ. Press, 1991.
10. *Mas-Collel A., Whinston M.D., Green J.R.* Microeconomic theory. N.Y.: Oxford Univ. Press, 1995.
11. *Губко М.В., Новиков Д.А.* Теория игр в управлении организационными системами. М.: Синтег, 2002.
12. *Новиков Д.А., Цветков А.В.* Декомпозиция игры активных элементов в задачах стимулирования // АиТ. 2001. № 2. С. 173–180.
13. *Новиков Д.А., Цветков А.В.* Агрегирование информации в задачах стимулирования // АиТ. 2001. № 4. С. 120–127.
14. *Новиков Д.А., Смирнов И.М., Шохина Т.Е.* Механизмы управления динамическими активными системами. М.: ИПУ РАН, 2002.

Статья представлена к публикации членом редколлегии П.Ю. Чеботаревым.

Поступила в редакцию 17.06.2002