

ЗАДАЧИ СТИМУЛИРОВАНИЯ В ДИНАМИЧЕСКОЙ СЕТЕВОЙ АКТИВНОЙ СИСТЕМЕ

М.В. Белов

Компания ИБС

Россия, 127018, Москва, Складочная ул., 3 стр.1

E-mail: mbelov59@mail.ru

Ключевые слова: многоэлементные динамические активные системы, стимулирование.

Аннотация: Рассмотрена многоэлементная динамическая активная система с ограничением на действия агентов в форме совместной технологии. Доказаны утверждения о существовании оптимальной компенсаторной системы стимулирования, о выигрышах агентов и центра, о том, что такая система декомпозирует задачу по агентам и по периодам и обеспечивает гарантировано минимальные (по всем возможным дальновидностям агентов) затраты центра на реализацию требуемой траектории. Предложен алгоритм оптимального согласованного планирования в такой активной системе.

1. Введение

Рассмотрим многоэлементную динамическую активную систему [1], включающую центр, конечное множество $N = \{1, 2, \dots, n\}$ агентов, $n \geq 2$, и сеть $G = (N, E)$, вершины которой соответствуют агентам (с правильной нумерацией [2]), а множество дуг $E \subseteq N \times N$ отражает «технологические» связи между агентами; через $N_i = \{j \in N \mid (i, j) \in E\}$ обозначаем множество предшественников i -го агента в сети G , $i \in N$.

Считаем, что сеть имеет единственный выход – n -ю вершину. и $M_0 \subseteq N$ множество входов, через M_k обозначаем множество вершин, в которые входят дуги только из вершин, принадлежащих множествам $\{M_j\}$, $j = \overline{0, k-1}$; число $k(i)$ называем рангом вершины i , принадлежащей множеству M_k .

Рассматриваем функционирование данной активной системы в течение $t = \overline{1, T}$ периодов дискретного времени. В каждом периоде t i -й агент характеризуется своими скалярными действием $y_i(t) \geq 0$ и результатом деятельности $z_i(t) \geq 0$. Обозначим через $y_D(t)$ вектор действий агентов из множества $D \subseteq N$, через $z_D(t)$ – вектор результатов деятельности агентов из этого множества. Связь результата деятельности агента с его действием и используемыми им в процессе этой деятельности результатами других агентов определяется «технологической функцией» $Q_i(\cdot, \cdot): \mathbb{R}_+^{(|N_i|+1)} \rightarrow \mathbb{R}_+^1$, т. е. $z_i(t) = Q_i(y_i^t; z_{N_i}^t)$ (с помощью нотации ξ_i^t обозначаем наборы всех значений величины $\xi_i(\tau)$ для промежутка $\tau \in [1; t]$). Для $i \in M_0$ имеет место $N_i = \emptyset$, поэтому положим $z_i(t) = Q_i(y_i^t; z_0^t)$, z_0 – L -мерный вектор «входов» сети.

Рассматриваемая модель агента подразумевает, что выход агента един для всех его последователей (множество последователей i -го агента определим как

$R_i = \{j \in N \mid (i, j) \in E\}$), т. е. эта модель имеет условную структуру «несколько входов, действие, один результат». Более общим (но выходящим за рамки настоящего исследования) является случай структуры «несколько входов, действие, несколько результатов».

Предположим, выполнено **предположение III**: Функции $Q_i(\cdot, \cdot)$ взаимно однозначны относительно действий самого агента и результатов его предшественников в текущем периоде. Т. е. при заданных результатах деятельности предшественников агента в текущем периоде $z_{N_i}(t)$ (и известных $y_i|_i^{t-1}$ и $z_{N_i}|_i^{t-1}$ в предыдущих периодах) текущее действие агента $y_i(t)$ однозначно определяет результат его деятельности $z_i(t)$, и наоборот – зная результат деятельности агента $z_i(t)$ и его предшественников $z_{N_i}(t)$, можно однозначно восстановить его действие в текущем периоде. •

Выбор действий $y(t)$ требует от агентов затрат $c_i(y_i|_i^t)$, где для функции затрат выполняется **предположение P2**: $c_i(\cdot)$ – неотрицательные, выпуклые, гладкие, строго монотонно возрастающие по всем аргументам функции затрат, $i \in N$. •

Таким образом, технология деятельности i -го агента описывается кортежем $(N_i, c_i(\cdot), Q_i(\cdot, \cdot))$.

Предположим, что i -й агент в периоде t получает за результаты $z_i|_i^t$ своей деятельности вознаграждение (стимулирование) $\sigma_i(z_i|_i^t) \geq 0$, причем функции $\sigma_i(\cdot)$ неотрицательны и кусочно-непрерывны по каждому из аргументов. Также предположим, что каждый агент рационален, т. е., выбирая свое действие $y_i(t) \geq 0$, i -й агент, стремится максимизировать свою целевую функцию $f_i(\cdot)$ – разность между вознаграждением и затратами:

$$(1) \quad f_i(y_i|_i^t) = \sigma_i(z_i|_i^t) - c_i(y_i|_i^t), \quad i \in N.$$

Описанная модель – совокупность множества агентов N (быть может, с их интересами и предпочтениями), сети G и набора $\{Q_i(\cdot, \cdot)\}$ технологических функций назовем *динамической сетевой активной системой* (ДСАС). ДСАС является обобщением модели «производственных цепочек» [3] на случай сетевой структуры, а также частным случаем многоэлементной динамической активной системы [1].

Сформулируем задачу стимулирования в ДСАС. Обозначим через σ_D вектор-функцию стимулирования агентов из множества $D \subseteq N$ в различные периоды времени $t \in [1; T]$. Предположим, что центр полностью дальновиден (степень его дальновидности равна T) и осуществляет программное управление. Предположим, что агенты недальновидны¹ и в каждом периоде выбирают действия однократно, одновременно и независимо, причем сеть $G = (N, E)$, технологии деятельности всех агентов, их целевые функции (включая функции стимулирования) и результаты деятельности в предыдущие периоды $z_{N_i}|_i^{t-1}$ являются на момент выбора действий *общим знанием* [4]. Обозначим через $E_M(\sigma_N)$ множество равновесий Нэша в рассматриваемой игре в нормальной форме (данное множество в теории управления организационными системами [4] называется множеством действий, реализуемых системой стимулирования σ_N):

$$(2) \quad E_M(\sigma_N) = \{y_N^*|_i^T \in \mathfrak{R}_+^{n^T} \mid \forall i \in N, \forall t \in [1; T], \\ \forall s \geq 0: f_i(y_i^*(t), y_i|_i^{t-1}, y_{-i}^*|_i^t) \geq f_i(s, y_i|_i^{t-1}, y_{-i}^*|_i^t)\}$$

Предположим, что конечной (например, рыночной) ценностью обладает только результат деятельности выхода сети – результат ДСАС в целом, обозначим его $z_n(y_N|_i^t)$, который в периоде t приносит центру доход $h(z_n(t)) \geq 0$. Тогда целевая функция ДСАС

¹ Или у центра нет информации о дальновидности агентов, и поэтому центр вынужден считать агентов недальновидными и обеспечивать им компенсацию затрат в каждом периоде.

(критерий эффективности управления ДСАС) представляет собой разность между доходом и суммарными затратами на стимулирование всех агентов:

$$(3) \quad \Phi(\sigma_N) = \min_{y_N^* |_1^t \in E_N(\sigma_N)} \left\{ \sum_{t=1}^T \delta(t_0; t) \left[h(z_n^*(y_N^* |_1^t)) - \sum_{i=1}^n \sigma_i(z_i^*(y_N^* |_1^t)) \right] \right\},$$

где $\delta(t_0; t)$ – распределение дальновидностей центра.

Задача управления заключается в построении системы стимулирования σ_N , имеющей максимальную эффективность:

$$(4) \quad \Phi(\sigma_N) \rightarrow \max_{\sigma_N}$$

Значение функционала (3), достижимое в результате решения задачи (4), назовем *эффективностью ДСАС* и обозначим через K . Задача (1) – (4) не тривиальна – «прямой» поиск ее решения не представляется возможным. Воспользуемся свойствами ДСАС, позволяющими применить теоремы о декомпозиции игры агентов [1, 3, 4].

2. Декомпозиция игры агентов в ДСАС

Построим компенсаторную систему стимулирования в ДСАС, оптимальную с точки зрения затрат центра. Зафиксируем некоторую траекторию выхода сети $z_n |_1^t = x |_1^t$, обозначим как $y_N^* |_1^t(x |_1^t)$ – плановые траектории – наборы векторов действий агентов, приводящих к траектории – результату $x |_1^t$. (В силу предположения П1 несложно показать, что для любой осуществимой траектории результатов $z_n |_1^t = x |_1^t$, последовательно по $t = 1, 2, \dots, T$ может быть построена единственная траектория планов по действиям агентов $y_N^* |_1^t(x |_1^t)$, обеспечивающая $z_n |_1^t = x |_1^t$). Будем обозначать $y_i^*(x; t)$ – значение элемента плановой траектории – действие i -го агента в периоде t , обеспечивающее конечный результат $x |_1^t$, а u_i – *резервную полезность* i -го агента (выигрыш агента при отказе от участия в данной системе [4]) в течение одного периода.

Утверждение 1. Если выполнены предположения П1 и П2, то система стимулирования для всех $i \in N$ в каждом периоде $t \in [1; T]$:

$$(5) \quad \sigma_i^*(x; y |_1^t; t) = \begin{cases} c_i(y_N^* |_1^t(x |_1^t)) + u_i, & y_i(\tau) \geq y_i^*(x; \tau) \text{ для всех } \tau \in [1; t], \\ 0, & y_i(\tau) < y_i^*(x; \tau), \text{ хотя бы для одного } \tau \in [1; t]. \end{cases}$$

реализует вектор действий агентов – плановую траекторию $y_N^* |_1^t(x |_1^t)$ – как равновесие в доминантных стратегиях их игры с минимальными взвешенными суммарными затратами центра на стимулирование. Выигрыши всех агентов в каждом периоде в этом равновесии тождественно равны их резервным полезностям.

Доказательство. Для всех периодов времени $t \in [1; T]$ по возрастанию, начиная с 1-го до T -го включительно, фиксируем произвольный номер агента $i \in N$ и произвольную обстановку $y_{-i} |_1^t = (y_1 |_1^t, \dots, y_{i-1} |_1^t, y_{i+1} |_1^t, \dots, y_n |_1^t)$ игры для него. Покажем, что $\forall y_i(t) \geq 0$ значение целевой функции агента не может быть выше, чем при выборе действия из плановой траектории $f_i(y_i^*(x; t), y_i |_1^{t-1}, y_{-i} |_1^t) \geq f_i(y_i(t), y_i |_1^{t-1}, y_{-i} |_1^t)$. Действительно, подставляя систему стимулирования (5) в целевую функцию агента $f_i(\cdot)$, получим при $y_i(t) < y_i^*(x; t)$:

$$(4) \quad c_i(y_i^*(x; t), y_i |_1^{t-1}, y_{-i} |_1^t) + u_i - c_i(y_i(t), y_i |_1^{t-1}, y_{-i} |_1^t) \geq 0 - c_i(y_i(t), y_i |_1^{t-1}, y_{-i} |_1^t)$$

В силу неотрицательности как резервной полезности, так и функции затрат, данное неравенство всегда выполнено. При $y_i(t) = y_i^*(x; t)$ получим $u_i \geq 0$. При $y_i(t) > y_i^*(x; t)$ получим:

$$c_i(y_i^*(x; t), y_i |_1^{t-1}, y_{-i} |_1^t) + u_i - c_i(y_i^*(x; t), y_i |_1^{t-1}, y_{-i} |_1^t) \geq$$

$$\geq c_i(y_i^*(x;t), y_{i-1}^{t-1}, y_{-i}^t) + u_i - c_i(y_i(t), y_{i-1}^{t-1}, y_{-i}^t),$$

т. е. $c_i(y_i^*(x;t), y_{i-1}^{t-1}, y_{-i}^t) \leq c_i(y_i(t), y_{i-1}^{t-1}, y_{-i}^t)$, что выполнено в силу предположения П2.

Таким образом, показано, что любому агенту в любом периоде, независимо от обстановки игры, выгодно выбирать соответствующую компоненту вектора $y_{N|1}^t(x|1^t)$, при этом он получит выигрыш, в точности равный своей резервной полезности. А это и есть определение равновесия в доминантных стратегиях; (любое такое равновесие является и равновесием Нэша) [5].

Следовательно, система стимулирования (5) реализует плановую траекторию - вектор действий агентов $y_{N|1}^t(x)$ - как равновесие в доминантных стратегиях, при этом суммарные взвешенные затраты центра равны

$$(6) \quad C(x|1^t) = \sum_{t=1}^T \delta(t_0;t) \sum_{i=1}^n (c_i(y_{N|1}^t(x|1^t)) + u_i)$$

Пусть существует другая система стимулирования $\hat{c}_N(\cdot)$, реализующая тот же вектор действий агентов $y_{N|1}^t(x|1^t)$ и характеризующаяся строго меньшими суммарными затратами центра на стимулирование. Это означает, что найдется агент $j \in N$, для которого в каком-то периоде времени $t \in [1;T]$ выполняется $\hat{c}_j(y_{N|1}^t(x|1^t)) < c(y_{N|1}^t(x|1^t)) + u_j$, то есть, значение его целевой функции строго меньше резервной полезности. Но тогда для этого агента целесообразным является отказ от участия в данной системе и получение выигрыша, равного резервной полезности. Получили противоречие, утверждение доказано. •

Из утверждения 1 непосредственно следует

Утверждение 2. В ДСАС для любой осуществимой траектории результатов сети $z_n|1^t$ может быть построена компенсаторная система стимулирования (5), которая:

- реализует траекторию действий агентов, приводящих к требуемой траектории результатов $z_n|1^t$;
- декомпозирует задачу по агентам и по периодам;
- обеспечивает гарантировано минимальные (по всем возможным дальновидностям агентов) затраты центра на реализацию траектории $z_n|1^t$, определяемые (6).

Отметим, что оптимальная система стимулирования (5) не зависит от структуры сети G и совокупности технологических функций !

Имеет смысл подчеркнуть, что результат утверждения 1 и оптимальная компенсаторная система стимулирования (5) могут быть корректно расширены на случай, когда множество возможных действий $y_i(t)$ каждого из агентов конечно. Соответственно и в этом случае функции стимулирования агентов могут формироваться независимо друг от друга и не зависят от технологии их деятельности.

Важность такого расширения определяется не сложностью получающейся математической модели, напротив, модель становится практически тривиальной. На современном этапе развития экономики в подавляющем большинстве случаев моделирование «нетривиального» множества возможных действий агентов оказывается не вполне адекватным. При вступлении в отношения с центром (или отказе от этого) у агента в большинстве случаев нет множества альтернативных действий, описываемых полупрямой $y_i > 0$. Если агент вступает в отношения с центром, то от него требуется вполне определенный результат z_i , определяемый узким диапазоном его характеристик, в условиях развитых технологий отвечает этому результату практически единственное действие $y_i(x)$. Это происходит и в случае взаимоотношений «сотрудник – работодатель», и в случае взаимоотношений фирм заказчика и подрядчика; т. е. по сути выбор агента би-

нарен: участвовать в ДСАС (и реализовать требуемое действие) или предпочесть альтернативную деятельность и получить резервную полезность.

3. Планирование в ДСАС

Если субъекту, осуществляющему управление ДСАС – *центру* - известны технологии деятельности всех агентов, то он может:

а) по набору функций $\{Q_i(\cdot), i \in N\}$ найти функцию $Q(y_N, z_0)$, определяющую зависимость результата $z_n|_n^T$ деятельности n -го агента от вектора $y_N|_n^T$ действий всех агентов и входа САС $z_0|_n^T$ ($Q(\cdot)$ функцию условно назовем *агрегированной технологией* ДСАС в целом);

б) для всех осуществимых траекторий выхода сети $z_n|_n^T = x|_n^T$ построить траекторию действий агентов $y_N^*|_n^T(x|_n^T)$, обеспечивающую $z_n|_n^T = x|_n^T$, и получить значение функции $C(\cdot)$ - минимальных (по всем возможным системам стимулирования) взвешенных затрат, необходимых для реализации $x|_n^T$

$$C(x|_n^T) = \sum_{t=1}^T \delta(t_0; t) \sum_{i=1}^n (c_i(y_i^*(x; t)) + u_i)$$

с) решить *задачу планирования* – найти оптимальный *план* $x^*|_n^T$ (траекторию выхода сети), максимизирующий разность между выручкой и суммарными затратами всех агентов с учетом распределения дальновидностей центра:

$$(7) \quad x^*|_n^T = \text{Arg max}_{x|_n^T} \left\{ \sum_{t=1}^T \delta(t_0; t) h(x|_n^t) - C(x|_n^T) \right\}$$

Если агенты полностью дальновидны и центр знает распределения их дальновидностей $\delta_i(t_0; t)^2$, он может воспользоваться этим и повысить эффективность своего управления $\Phi(\sigma_N)$, используя аккордную систему стимулирования:

$$(8) \quad \sigma_i^*(x; y|_n^t; t) = \begin{cases} \tilde{C}_i, & \text{при } y_i(\tau) \geq y_i^*(x; \tau) \text{ для всех } \tau \in [1; T] \text{ и } t = T; \\ 0, & \text{при } y_i(\tau) < y_i^*(x; \tau) \text{ хотя бы для одного } \tau \in [1; t] \text{ или } t < T. \end{cases}$$

где $\tilde{C}_i = \sum_{t=1}^T \delta_i^{-1}(t_0; T-t) (c_i(y_N^*|_n^t(x|_n^t)) + u_i)$.

В этом случае траектория оптимального плана будет получена из выражения

$$x^*|_n^T = \text{Arg max}_{x|_n^T} \left\{ \sum_{t=1}^T \delta(t_0; t) h(x|_n^t) - \delta(t_0; T) \sum_{i=1}^n \tilde{C}_i \right\}$$

и может отличаться от траектории, получаемой из (7) в силу возможных различий распределений дальновидностей центра и агентов.

Заметим, что дополнительный выигрыш от использования аккордной системы стимулирования (8) вместо системы (5) возникает у центра не всегда и зависит от соотношения распределений дальновидностей центра и агентов. В частности, в случае совпадающих распределений вида $\delta_i(t_0; t) = \delta(t_0; t) = \exp(-\alpha(t - t_0))$ (где $\alpha \geq 0$) эффективность управления $\Phi(\sigma_N)$ для обоих случаев (5) и (8) совпадает.

² И вся эта информация является общим знанием всех участников игры.

Список литературы

1. Новиков Д.А., Смирнов И.М., Шохина Т.Е. Механизмы управления динамическими активными системами. М.: ИПУ РАН, 2002. 124 с.
2. Бурков В.Н., Горгидзе И.А., Ловецкий С.Е. Прикладные задачи теории графов. Тбилиси: ВЦ АН ГССР, 1974. 232 с.
3. Новиков Д.А., Цветков А.В. Механизмы стимулирования в многоэлементных организационных системах. М.: Апостроф, 2000. 184 с.
4. Novikov D.A. Theory of Control in Organizations. N.Y.: Nova Science Publishers, 2013. 341 p.
5. Myerson R.B. Game Theory: Analysis of Conflict. London: Harvard University Press, 1997. 584 p.