УДК 330.115

ИТЕРАЦИОННЫЕ МЕТОДЫ СОГЛАСОВАННОЙ ОПТИМИЗАЦИИ С ИДЕНТИФИКАЦИЕЙ ЭЛЕМЕНТОВ МОДЕЛИ ВЫБОРА

Виноградов Г.П.*, Кузнецов В.Н.*

**Тверской государственный технический университет, г. Тверь, Россия*

WGP272NG@mail.ru

Исследуется сходимость интерактивной процедуры согласованной оптимизации в задаче управления эволюцией организационно-технологической системы при различных вариантах информированности центра о возможностях агентов. Приведены условия, при которых возможно построение процедуры интерактивного обмена информацией между центром и агентами, гарантирующими сходимость решения задачи согласованной оптимизации.

Ключевые слова: активная система; интеллектуальный агент; согласованная оптимизация; принятие решений.

Введение

Статья посвящена итерационным методам решения задачи оптимального согласованного планирования, когда центр, не обладая полной информацией о возможностях активных, целеустремленных агентов, вынужден осуществлять одновременно с определением плановых заданий идентификацию модели своих представлений по отдельным ее составляющим: исходным данным, ограничениям, взаимосвязи между входами и выходами, о технологических множествах агентов, о целевых функциях агентов и т.п. Такие методы, с одной стороны, расширяют вычислительные возможности теории согласованного планирования и управления, а с другой – могут использоваться как средство моделирования и управления эволюционизирующими системами. Это соответствует ситуации, когда объект управления для центра плохо поддается адекватному моделированию из-за его сложности, проблемы получения точной информации об его параметрах, неудовлетворительной структурированности, плохой формализуемости, распределения знания о предельных возможностях подсистем среди агентов. Причем в последнем случае информация центром может быть получена только в результате коммуникативного процесса с агентами.

Поэтому расчет согласованного плана в этих условиях превращается в сложный вычислительный процесс итерационного типа с интерактивным взаимодействием его участников, цель которого формирование у всех участников согласованных представлений о ситуации выбора. Центр для формирования представлений о возможностях агентов использует результаты анализа возникающих в системе ситуаций от управляющих воздействий, возмущений.

Представления центра в этом случае можно формализовать в виде нестационарной модели, свойства и структура которой определяются информированностью центра. Тогда используемый центром итерационный метод $\varphi \in \{R^n \to R^n\}$ должен порождать сходящуюся последовательность представлений, которая характеризуется помехоустойчивостью, независимостью логической структуры оператора φ от конкретных исходных данных. Условием, обеспечивающим реализацию перечисленных свойств, является необходимость введение в итерационный оператор φ параметров, характеризующих объект управления: среду интересы участников процесса, – например, такого типа $\{x_{h+1} = \varphi[\Omega_h]x_h\}$, $h = 1,2,\dots$, где $\Omega_h = \{y_h, \omega_h, \dots\}$ – перечисленная система параметров (числовых, символьных, лингвистических, предикатных) со значениями на момент h.

Особенность этой схемы состоит в необходимости использования алгоритмов, позволяющих получить информацию, доопределяющую и уточняющую параметры итерационного оператора. Основным источником такой информации являются взаимодействующие агенты, а средством ее получения – интерактивные алгоритмы.

1. Постановка задачи

Рассмотрим задачу планирования в двухуровневой организационно-технологической системе (ОТС), состоящей из центра и m агентов, обладающих свойствами активности, автономности, креативности. Каждый агент управляет каким-либо одним технологическим узлом. План для k-го агента описывается n_k-мерным вектором $x_k = \{x_{kj}, k = \overline{1,n_k}\}$, положительные компоненты которого указывают на выпускаемые продукты, а отрицательные – на затраты для выпуска заданной продукции. Возможности k-го агента по выпуску продукции описываются технологическим множеством – областью X_k в пространстве из n_k измерений: любой допустимый план x_k k-го агента должен принадлежать области X_k: $x_k \in X_k, k = \overline{1,m}$.

План ОТС будет описываться вектором $x = \{x_k, k = \overline{1,m}\}$ с размерностью $\sum_{k=1}^{m} n_k = N$. Очевидно, что $x \in X = \prod_{k=1}^{m} X_k$. Допустимый план x должен удовлетворять также ряду глобальных ограничений вида $G(x) \geq b$, где $G(x)=\{g_i(x), i = \overline{1,n}\}, b = \{b_i, i = \overline{1,n}\}$, n – количество ограничений. Будем считать, что функции $g_i(x), i = \overline{1,n}$ – вогнутые, дифференцируемые функции, а множество X – выпуклое множество. Тогда задача, решаемая центром, может рассматриваться как задача вогнутого программирования

$$f(x) \to max$$

$$G(x) \geq b,\ x \in X = \prod_{k=1}^{m} X_k \qquad (1)$$

Задача (1) является задачей оптимального планирования (ОП) [Полтерович, 1969]. Ее особенность заключается в том, что центр не знает все потенциальные возможности агентов. Поэтому ее решение зависит от информации о своих возможностях, которую агенты сообщают центру.

Пусть интересы агента описываются вектором $o_k \in O_k^{o}$, где O_k^{o} – множество состояний, имеющих различную ценность для агента. Значения вектора o_k определяются значениями вектора y_k = $y_{k1}, y_{k2}, ..., y_{km_k}$ фактического выпуска продукции, полученными при реализации плана x_k агентом. При управлении технологическим узлом свои предпочтения на множестве O_k^{o} агент реализует путем выбора значений режимных параметров z_k, которые в свою очередь определяют значения вектора y_k, а значит и o_k. Справедливо условие $z_k \in Z_k$, где Z_k – множество допустимых значений режимных параметров, определяемое технологическим регламентом. На технологический узел воздействуют различного рода возмущающие воздействия $\xi_k \in \Omega_k$, из которых часть $\zeta_k \in \Xi_k$ агент принимает во внимание. Поэтому $\Xi_k \subseteq \Omega_k$ и $\Xi_k \cap \Omega_k \neq \varnothing$.

Поведение человека как интеллектуального агента, зависящее от его субъективных представлений о ситуации выбора рассматривалось в работах [Виноградов, 2011, Vinogradov, 2009, Чхартишвили, 2004]. В них было показано, что принимаемое решение агентом о способе действия определяется его оценками компонент ситуации целеустремленного состояния, зависящими от структуры его информированности I_t, которая определяется его знаниями, убеждениями, ценностями, нормами, опытом. В этих же работах была предложена модель принятия решений агентом, позволяющая учитывать его индивидуальные оценки компонент ситуации целеустремленного состояния, которая имеет следующий вид

$$P_k(z_k) = \text{Arg max } E\varphi_k(o_k, u_k)$$

$$\zeta_k \in \Xi_k, z_k \in Z_k(I_t^i), u_k \in U_k,$$

$$o_k = o_k(y_k) \in O_k^o, y_k = f_{I_t}(z_k, \zeta_k) \in X_k, \qquad (2)$$

$$O_k^o = \{o_{kj}^o,\ j = \overline{1, N_k}\}, I_t \subseteq M,$$

$$\chi_1^k(E\varphi_i) \geq \chi_1^{ko},\ \chi_2^k(EE_i) \geq \chi_2^{ko},$$

где $E\varphi_k$ и EE_k – оценки агента удельной ценности ситуации целеустремленного состояния по результату и эффективности; χ_l^{k0}, $l = \overline{1,2}$ – оценки, отражающие эмоциональное отношение агента к ситуации выбора, u_k – управляющее воздействие центра.

Оптимальное решение задачи (1), учитывающей субъективное поведение агента в форме (2), практически невозможно. Проблема состоит в построении моделей выбора и интерактивного процесса обмена информацией между центром и агентами так, чтобы получить оптимальное решение задачи (1, 2) и создать для каждого агента условия для наиболее полного использования своих возможностей в интересах системы, не используя детальную математическую модель их производственных возможностей. На возможность построения таких процедур указано в работе [Полтерович, 1969].

2. Задача согласованной оптимизации с доопределением модели выбора

Математическая постановка задачи оптимального согласованного планирования имеет вид: $f(x)\to \max; G(x)\le b; x\in X; x\in S$. Здесь $S=\prod_{k=1}^{m} s_k$ множество согласованных планов, таких что $s_k=\{x_k \mid \max E\varphi_k(x_k^*,x_k)\le E\varphi_k(x_k,x_k)\}$, где

$$E\varphi_k\left(x_k^*,x_k\right)=E\varphi_k\left(x_k^*\right)+\begin{cases} c_k \text{ } при \text{ } x_k^*=x_k \\ 0 \text{ } при \text{ } x_k^*\ne x_k \end{cases}$$ –

целевая функция k-го агента с учетом фонда материального поощрения c_k, планируемого центром за выполнение предлагаемого им плана x_k; x_k^* – план производственной программы, выбираемый k-м агентом с учетом своих интересов и технологических возможностей.

Допустим, что суммарный фонд материального поощрения зависит от результатов работы системы в целом, и центр может назначать c_k исходя из очевидного ограничения $\sum_{k=1}^{m} c_k\le c(x)$. Предположим, что выполняется условие благожелательности агента по отношению к центру. Оно заключается в том, что если максимальный выигрыш k-го агента достигается как при плане, назначаемом центром, так и при некоторых других возможных значениях x_k^*, то агент предпочтет план центра x_k.

Необходимым и достаточным условием принадлежности плана x_k к согласованным ($x\in S$) является соотношение

$$\forall k: E\varphi_k(x_k)\ge \max_{x_k^*\in X_k} E\varphi_k(x_k^*).$$

Следовательно, условия согласованного планирования будут иметь вид: $c(x)-\sum_{k=1}^{m} c_k\ge 0$,

$$E\varphi_k(x_k)+c_k\ge \max_{x_k^*\in X_k} E\varphi_k(x_k^*).$$

Если функции $E\varphi_k(x)$, $g(x)$, $c(x)$ являются вогнутыми функциями, а множество X – выпуклым множеством, то задача оптимального согласованного планирования (ОСП) является задачей вогнутого программирования.

Пусть в задаче согласованного планирования такие элементы, как целевые функции агентов $E\varphi_k(x_k)\in\{R^{m_k}\to R\}$, вектор правых частей системы ограничений $g\in\{R^{m_k}\to R^n\}$, условия включения $x\in X$ и $x\in S$ – плохо формализуемы. Это означает, что способа конструктивного задания элементов $E\varphi_k$, g, X и S нет. Центр их определяет по результатам наблюдений, экспериментов, экспертных оценок и т.п. Это означает , что ему на момент h известны пары множеств X' и X'', S' и S'', G' и G'' таких, что $g(x)\le b\ \forall x\in G'$ и $g(x)>b\ \forall x\in G''$, $x\in X'$, $X''\subset R^{m_k}\backslash X'$, $x\in S'$, $S''\subset R^{m_k}\backslash S'$.

Относительно целевых функций агентов $E\varphi_k$ представления центра состоят в том, что имеется множество $O^o\subset R^n$ (как правило, оно является конечным) такое, что для любого его элемента $o\in O^o$ известно либо значение $E\varphi_k$, либо некоторые характеристики этого значения, например, $\operatorname{sgn}(E\varphi_k)$, либо подмножество согласованных планов для k-го агента, когда $x_k\in S_k'$, либо принадлежность $E\varphi_k$ некоторому интервалу и т. п. Это означает, что центр обладает правилом $x(o)\to\Delta(x(o))\subset R$ для всех $o\in O^o$, то есть центр изучает информацию о потребностях, мотивирующих поведение агентов. Поскольку O конечно, то и множество $\Delta(x(o))$ для всех $o\in O^o$ конечно. Центр предполагает, что существует дополнение $X\backslash X'$, которое он рассматривает как резервы агентов, которые он должен выявить. Тогда доопредение модели согласованного планирования будет состоять в выявлении $X'\subseteq X, S'\subseteq S$ и $S'\bigcap X'$ таких, что $E\varphi_k(x_k\in X')\le E\varphi_k(x_k\in X)$. Тогда задача оптимального согласованного планирования становится конструктивным описанием символической модели в виде

$$f(x)\to\max; G(x)\le b; x\in X'; x\in S' \qquad (3)$$

и допускает аналитические исследования и численные расчеты.

Построение модели (3) предполагает адаптацию представлений центра о технологических множествах агентов путем активного и пассивного

экспериментирования для организации процесса последовательного пополнения данных о множествах
$X', X'', S', S'', G', G''$ и $\{E\varphi_k(x(o)), o \in O^o\}$.

Это дает возможность построить последовательность представлений об этих множествах, например, такую, что $\lim_{h \to \infty} (X_k \setminus X_k^h) = \varnothing$. В основе такого процесса лежит предположение о том, что агенты обладают свойствами активности и креативности. При выполнении гипотезы рационального выбора агентами центр может формировать управляющее воздействие $u^{(h)} = \{u_k^{(h)}, k = \overline{1,m}\}$ такое, что агенты будут мотивированы выбирать напряженные планы.

Обозначим через $\Sigma^{(h)}$ материал результатов наблюдений на момент h, тогда итерационный оператор для модели (3) примет вид $\varphi[\Sigma^{(h)}](x)$, а через $\{E\varphi_k(x(o)), o \in O^o\}$, $g^{(h)}(x), X'_h, S'_h$ – представления центра о модели выбора (3). На каждом шаге h центр решает задачу согласованной оптимизации

$$\sup\{f(x) : g^{(h)}(x) \le b^{(h)}, x \in X^{(h)}, x \in S'\} \qquad (4)$$

и организует процесс по итерационной схеме $x^{(h+1)} = \varphi[\Sigma^{(h)}](x^{(h)})_{\forall h}$, где $x^{(h+1)}$ является одним из компонентов $u^{(h+1)}$.

В качестве конструктива можно использовать любой алгоритм адаптации, например, стохастической аппроксимации или методы деформируемого многогранника. Пополнение материала наблюдений может осуществляться следующим образом. Пусть центр построил вариант плана $X^{(h)}$ путем решения задачи (4). В пространстве плановых показателей определяется точка $\bar{x}^{(h)}$ более предпочтительного плана, при котором $f(\bar{x}^{(h)}) > f(x^{(h)})$, но не выполняется некоторые ограничения задачи (4). В этом направлении центр определяет величину шага и рассчитывает $x^{(h+1)} = x^{(h)} + \Delta^{(h)}$ и остальные компоненты вектора $u^{(h+1)}$. Результаты расчета сообщаются агентам. Агенты рассчитывают свои варианты планов и формируют предложения по модернизации и интенсификации технологического процесса. Эти данные диагностируются центром и используются им для пополнения множеств $X', X'', S', S'', G', G''$, а множество O_h^o переходит в $O_{h+1}^o \cup o(x^{(h+1)})$.

Скорректированное множество O_{h+1}^o служит основанием для организации институционального, стимулирующего и информационного управления центром.

3. Поведение центра

Итеративный процесс планирования может быть представлен следующей схемой:

1. центр по исходной и дополнительной информации, сообщенной на каждом h-м шаге k-м агентом, рассчитывает вариант производственной программы $x^{(h)} = \{x_k^{(h)}\}$ и планируемый размер фонда материального поощрения $c^{(h)} = \{c_k^{(h)}\}$ за принятие k-м агентом данного варианта в качестве плана. Центр сообщает агентам значения $\{x_k^{(h)}\}$ и $\{c_k^{(h)}\}$;
2. агент выбирает и сообщает центру наилучший для себя план $x_k^{(h)*}$. Примем, что на каждом h-м шаге выполняется гипотеза о локально-оптимальном поведении агента;
3. если $x_k^{(h)*} \ne x_k^{(h)}$, то агент по запросу центра сообщает дополнительную информацию, т.е. осуществляется переход к первому шагу данной схемы. В случае совпадения $x_k^{(h)*} = x_k^{(h)}$, процедура планирования заканчивается.

Алгоритм определения плана будет зависеть от степени информированности центра. Рассмотрим три его варианта.

Вариант 1. Полная информированность центра о локальных ограничениях агентов и их целевых функциях. Таким образом, центр имеет исходную информацию:

$$f(x), G(x), b, X, X_k, \ \max_{x_k^* \in X_k} E\varphi_k(x_k^*).$$

Этому варианту можно сопоставить следующий алгоритм определения плана.

1. На основании исходной информации центр может рассчитать план либо путем решения задачи ОП ($x^{(h)} = x^o$), либо задачи ОСП ($x^{(h)} = x^{oc}$). Т.к. рассматриваемая процедура является итеративной, то центр на h-м шаге процедуры планирования имеет вектор значений планов, выбранных агентами за предыдущие (h-1) шагов: $\{x^{*(1)}, x^{*(2)}, \dots, x^{*(h-1)}\}$, где $x^{*(j)} = \{x_k^{*(j)}\}$.
2. На каждом шаге центр по дополнительной информации корректирует

предлагаемое агентом значение плана производственной программы и размера поощрения:

$$x_k^{(h)} = x_k^{(h-1)} + \beta\left(x_k^{(h-1)} - x_k^{*((h-1))}\right),$$

$$c_k^{(h)} = c_k^{(h-1)} + \alpha[\max_{x_k^* \in X_k} E\varphi_k(x_k^{(h)*}) - - E\varphi_k(x_k^{(h-1)})],$$

где α и β – коэффициенты акселерации, выбираемые из соображений, имеющих, как правило, неформальный характер, хотя можно определить правило их изменения.

3. На каждом шаге агент выбирает наилучшую стратегию путем $\max\left\{\max_{x_k^* \in X_k} E\varphi_k(x_k^{(h)*}), E\varphi_k(x_k^{(h)})\right\}$ и принимает один из вариантов плана: $x_k^{(h)*}$ или $x_k^{(h)}$.

4. Если $x_k^{(h)*} \neq x_k^{(h)}$, то процедура повторяется с п. 2, в случае $x_k^{(h)*} = x_k^{(h)}$ процедура заканчивается утверждением плана.

Сходимость предложенной схемы зависит от вида решаемой задачи.

1. Решается задача ОП. Пусть $x_k^{(h)} \notin S_k$, а также $X \cap S \neq \varnothing$, тогда

$$E\varphi_k(x_k^{(h)}) < \max_{x_k^* \in X_k} E\varphi_k(x_k^{(h)*}).$$

Так как согласно гипотезе благожелательности агента к центру выполняется условие

$$\max\left\{\max_{x_k^* \in X_k} E\varphi_k(x_k^{(h)*}), E\varphi_k(x_k^{(h)})\right\} = = \max_{x_k^* \in X_k} E\varphi_k(x_k^{(h)*}),$$

то $x_k^{(h)*} \neq x_k^{(h)}$ и требуемое число шагов процедуры планирования $h > 1$.

Сходимость процедуры будет обеспечена путем корректировки величины $c^{(h)}$ по правилу: $c^{(h)}(x) = c^{(h-1)}(x) + \gamma(c^{(h-1)}(x^{(h)*})) - - c^{(h-1)}\left(x^{(h)}\right)$, где $\gamma > 0$, что означает уступку выигрыша центра в пользу агента для обеспечения $x \in S$.

2. Пусть $x_k^{(h)} \in S_k$ и $X \cap S = \varnothing$ (центр не учитывает интересы агентов). Рассматриваемая процедура планирования не сходится к $x_k^{(h)}$. Доказательство очевидно.

3. Если $x_k^{(h)} \in S_i$ и $X \cap S \neq \varnothing$, то $E\varphi_k(x_k^{(h)}) \geq \max_{x_k^* \in X_i} E\varphi_k\left(x_k^{(h)*}\right)$, тогда из условия благожелательности агента к центру следует, что

$$\max\left\{\max_{x_k^* \in X_k} E\varphi_k(x_k^{(h)*}), E\varphi_k(x_k^{(h)})\right\} = = \max_{x_k^* \in X_k} E\varphi_k(x_k^{(h)*}).$$

Следовательно, $x_k^{(h)*} = x_k^{(h)}$ и процедура планирования сходится за один шаг, т.е. $h = 1$.

Вариант 2. Полная информированность центра о целевых функциях агента и неполной информированности о их локальных ограничениях, т.е. о множествах X_k. Т.о. центр имеет исходную информацию: $f(x), G(x), b, \max_{x_k^* \in X_k} E\varphi_k\left(x_k^*\right)$.

В этом случае можно рассмотреть следующую итеративную процедуру, использующую первую схему блочного программирования Полтеровича.

1. На h-м шаге центр получает дополнительную информацию от агентов о предлагаемых ими напряженных планах $x^{(h)*} = \{x_k^{(h)*}\}$. Центр, используя набор $\{x^{(1)*}, x^{(2)*}, \ldots, x^{(h)*}\}$ аппроксимирует множество X, используя гипотезу о его выпуклости:

$$x^{(h)} = \left\{x \mid x = \sum_{j=1}^{h} \alpha_j x^{(j)*}, \alpha_j \geq 0, \sum_{j=1}^{h} \alpha_j = 1\right\}$$

2. Центр решает на каждом шаге либо задачу ОП: $f(x) \to \max, G(x) \leq b, x \in X^{(h)}$, либо задачу ОСП, т.е. задачу ОП при дополнительных ограничениях, причем переменными являются как x_k, так и c_k, $k = \overline{1, m}$

$$E\varphi_k\left(x_k^{(h)}\right) + c_k^{(h)} \geq \max_{x_k^* \in X_k} E\varphi_k\left(x_k^{(h)*}\right), \quad c(x) - \sum_{k=1}^{m} c_i - \sum_{j=2}^{h} \sum_{k=1}^{n} \delta_k^{(j)} \geq 0, \quad (5)$$

где $\delta_k^{(j)}$ – планируемый размер дополнительного фонда материального поощрения, назначаемый центром агентам на h-м шаге вместе с планом $x_k^{(j)}$. Его величина определяется как $\delta_k^{(j)} = \langle p_k^{(j-1)}, x_k^{(j-1)} \rangle$, где $p_k^{(j-1)} = \alpha_k^{(j-1)} \mathrm{grad}_x F(x_k^{(j-1)}, y_k^{(j-1)}), k = \overline{1,m}$, $\mathrm{grad}_x F(\bullet)$ – градиент функции Лагранжа, которая для задачи ОСП имеет вид:

$$F(x,y) = f(x) - \sum_{l=1}^{n} y_l \left[g_l(x) - b_l \right] +$$

$$+ \sum_{k=1}^{m} y_{n+k} \left[c_i - \max_{x_k^* \in X_k} E\varphi_k\left(x_k^*\right) \right] + \quad .$$

$$+ y_{n+m+1} \left[c(x) - \sum_{k=1}^{m} c_i - \sum_{j=2}^{h} \sum_{k=1}^{m} \delta_k^{(j)} \right]$$

Для задачи ОП функция Лагранжа имеет тот же вид, но $y_{m+n+1} = 0$.Путем решения этой задачи центр получает значения $x^{(h)}, c^{(h)}, \delta^{(h)}$ и $p^{(h)} = \mathrm{grad}_x F(x^{(h)}, y^{(h)})$, где ($x^{(h)}$, $y^{(h)}$) – седловая точка функции Лагранжа на множестве $X^{(h)}$. $\alpha^{(h)} > 0$ подбирается из условия обеспечения расходов центра на стимулирование агентов за счет роста величины $c(x)$. Причем предполагается, что

$$\langle \mathrm{grad}_x c(x^{(h)}), x^{(h)} - x^{(h-1)} \rangle - (c^{(h)} - c^{(h-1)}) > 0,$$

т.е. на каждом шаге фонда поощрения должно хватать для стимулирования агентов за выбор напряженного плана. Так как

$$\langle \alpha^{(h)} \mathrm{grad}_x F\left(x^{(h)}, y^{(h)}\right), x^{(h)} - x^{(h-1)} \rangle +$$

$$\left(c^{(h)} - c^{(h-1)}\right) \le \langle \mathrm{grad}_x c\left(x^{(h)}\right), x^{(h)} - x^{(h-1)} \rangle ,$$

то

$$\alpha^{(h)} \le \frac{\langle \mathrm{grad}_x F(x^{(h)}, y^{(h)}), \Delta x^{(h)} \rangle}{\langle \mathrm{grad}_x c(x^{(h)}), \Delta x^{(h)} \rangle - (c^{(h)} - c^{(h-1)})},$$

где $\Delta x^{(h)} = x^{(h)} - x^{(h-1)}$.

Значения $x^{(h)}, c^{(h)}, \delta^{(h)}, p^{(h)}$ центр сообщает агентам.

3. По своей целевой функции

$$E\varphi_i\left(x_k^{(h)*}\right) + \begin{cases} c_k^{(h)}\left(x_k^{(h)}\right) + \sum\limits_{j=2}^{h-1} \delta_k^{(j)} + \delta_k^{(h)} > 0 \\ 0 + \delta_k^{(h)} < 0 \end{cases},$$

где для первого слагаемого в фигурных скобках должно выполняться условие $x_k^{(h)*} = x_k^{(h)}$ и $x_k^{(h)*} \in L^k$, а для второго $x_k^{(h)*} \ne x_k^{(h)}$ и $x_k^{(h)*} \in X_k^{(h)}$; агенты в соответствии с гипотезой о локально-оптимальном поведении выбирают наилучшую для себя ситуацию:

$$\max\{ \max_{x_k \in X_k} E\varphi_k\left(x_k^{(h)*}\right); E\varphi_i\left(x_k^{(h)*}\right) + c_k^{(h)} +$$

$$+ \sum_{j=2}^{h} \delta_k^{(j)}; E\varphi_k\left(x_k^{(h)*}\right) + c_k^{(h)} + \quad .$$

$$+ \sum_{j=1}^{h} \delta_k^{(j)} + \langle p_k^{(h)}, x_k^{(h)*} - x_k^{(h)} \rangle \}$$

Для задачи ОП

$$L^{(h)} = \{ x \mid G(x^{(h)}) \ge b,\ f(x^{(h)}) > f(x^{(h)*}) \},$$

а для задачи ОСП

$$L^{(h)} = \{ x \mid G(x^{(h)}) \ge b, E\varphi_k\left(x_k\right) + c_k \ge$$

$$\ge \max_{x_k \in X_k} E\varphi_k(x_k^{(h)*}), (c(x^{(h)}) - \sum_{i=1}^{n} c_k^{(h)} - ,$$

$$- \sum_{j=2}^{h} \sum_{k=1}^{m} \delta_k^{(h)}) \ge 0,\ f(x^{(h)}) > f(x^{(h)*}) \}$$

здесь $x^{(h)*}$ – план, выбираемый агентом на h-м шаге, $x^{(h)*} \in L^{(h)}$ – напряженный план, предлагаемый агентом центру на h шаге;

$$\begin{cases} \delta_k^{(h)} > 0 \text{ при } x_k^{(h)*} \in L^{(h)} \\ \delta_k^{(h)} < 0 \text{ при } x_k^{(h)*} \in X^{(h)} \end{cases}.$$

4. Агенты сообщают центру свои наилучшие стратегии. Если $x_k^{(h)*} \ne x_k^{(h)}, x_k^{(h)} \in X_k^{(h)}$, то центр корректирует свое решение по схеме, описанной в п.2. Если $x_k^{(h)} \in L_k^{(h)}$, то центр обновляет информацию, расширяет множество $X^{(h)} \subset X^{(h+1)}$ и вычисляет $x^{(h+1)}, c^{(h+1)}, \delta^{(h+1)}, p^{(h+1)}$. Если $x_k^{(h)*} = x_k^{(h)}$, то это означает, что

$$E\varphi_k\left(x_k^{(h)}\right) + c_k^{(h)} + \sum_{j=2}^{h} \delta_k^{(h)} \ge E\varphi_k\left(x_k^{(h)*}\right) +$$

$$+ c_i^{(h+1)} + \sum_{j=1}^{h} \delta_k^{(h)} + \langle p_k^{(h)}, x_k^{(h)*} - x_k^{(h)} \rangle \quad .$$

Так как

$$E\varphi_k\left(x_k^{(h)*}\right)+c_k^{(h+1)} \ge E\varphi_k\left(x_k^{(h)*}\right)+c_k^{(h)} \ge$$

,

$$\ge \max_{x_k^* \in X_k} E\varphi_k\left(x_k^{(h)*}\right)$$

то $\langle p_k^{(h)}, x_k^{(h)*}-x_k^{(h)}\rangle \le 0$ для всех k-х агентов,

т.е. $x_k^{(h)} \ne x_k^{(h+1)}, \operatorname{grad}_x c\left(x_k^{(h)}\right) \ne 0,$. Поэтому

$$\operatorname{grad}_x F\left(x_k^{(h)}, y_k^{(h)}\right) \ne 0$$

$\alpha_k^{(h)} \ne 0$ и

$$\langle \operatorname{grad}_x F\left(x_k^{(h)}, y_k^{(h)}\right), x_k^{(h)*}-x_k^{(h)}\rangle \le 0 .$$

Следовательно, точка $(x_k^{(h)}, y_k^{(h)})$ оказывается седловой точкой Лагранжа в задаче ОП или ОСП, и значение $x_k^{(h)*}=x_k^{(h)}$ является соответственно оптимальным планом или оптимальным согласованным планом. Интерактивная процедура заканчивается.

Предположим, что условия образования фонда материального стимулирования таковы, что для любого $x \in X$ выполняется неравенство $\left|\operatorname{grad}_x c(x)\right| > 0$, где $|\bullet|$ – длина вектора. Тогда рост фондов материального поощрения больше расходов на расширение множества $X^{(h)}$, поэтому в этом случае допустимое множество

$$\{x \mid G(x^{(h)}) \ge b\,;\, c(x^{(h)}) - \sum_{k=1}^{m} c_k^{(h)} - \sum_{j=2}^{h}\sum_{k=1}^{m} \delta_k^{(j)} {}_i^j \ge 0\,;\, E\varphi_k(x_k^{(h)*})+c_k^{(h)} \ge \max_{x_k^{(h)*}\in X_k^{(h)}} E\varphi_i(x_k^{(h)*})\} \cap X^{(h)} \qquad (6)$$

с ростом множества $X^{(h)}$ монотонно не сужается.

Достаточным условием сходимости описанного итеративного процесса является выполнение условия (3). Поскольку $x_k^{(h)} \in S_k$, то $E\varphi_k(x_k^{(h)})+c_k^{(h)} \ge \max_{x_k^* \in X_k} E\varphi_k(x_k^{(h)*})$. Из соотношения

$$c_k^{(h+1)} = c_k^{(h)} + \langle \operatorname{grad}_x c(x_k^{(h+1)}), x_k^{(h+1)} - x_k^{(h)}\rangle$$

следует неравенство

$$E\varphi_k(x_k^{(h)*})+c_k^{(h+1)} \ge E\varphi_k(x_k^{(h)})+c_k^{(h)} \ge$$

$$\ge \max_{x_k^* \in X_k} f_i(x_k^{(h)*})$$

для $x^{(h)*} \in L^{(h)}$. Согласно условию алгоритма блочного программирования $\langle p_k^{(h)}, x_k^{(h)*}-x_k^{(h)}\rangle > 0$ для всех $x^{(h)*} \in L^{(h)}$. Отсюда следует, что

$$E\varphi_k\left(x_k^{(h)*}\right)+c_k^{(h+1)}+\sum_{j=2}^{h+1}\delta_k^{(j)}+\langle p_k^{(h)}, x_k^{(h)*}-x_k^{(h)}\rangle > E\varphi_k\left(x_k^{(h)}\right)+c_k^{(h)}+ . + \sum_{j=2}^{h}\delta_k^{(j)}$$

Поэтому при решении задачи ОСП агент всегда будет выбирать $x^{(h)*} \in L^{(h)}$

Следовательно, на каждом h-м шаге агенты формируют предложения для $(h+1)$-го шага в виде значения напряженного плана $x^{(h)*} \in L^{(h)}$, который они могли бы реализовать. В силу условий (5) и (6) новое множество $X^{(h+1)}$ будет расширяться по сравнению с множеством $X^{(h)}$. При этом будут выполняться ограничения по фонду материального поощрения и условия совершенного согласования. Поскольку $X^{(1)} \subset \ldots \subset X^{(h)} \subset \ldots \subset X$, то допустимые множества в задаче ОСП не убывают. За счет расширения допустимого множества значений $f(x^{(h)})$ для каждой задачи ОСП будет только возрастать. Поэтому последовательность $f(x^{(1)})$, $f(x^{(2)}), \ldots,$ $f(x^{(h)}), \ldots$ монотонно не убывает и в силу компактности множества X имеет предел f^*, т.е. $\lim\limits_{h\to\infty} f(x^{(h)}) = f^*$ и $f^* = f\left(x^{oc}\right)$.

Можно показать, что если решается задача ОП и нарушено условие совершенного согласования, то $x_k^{(h)*} \ne x_k^{(h)}$ и $x^{(h)*} \notin L^{(h)}$. Следовательно, $x_k^{(h)*} = x_k^{(h+1)}$ и множество $X^{(h)}$ не расширяется в силу отсутствия у агентов стимулов отыскивать резервы для принятия напряженных планов и процедура итеративного планирования закончится не достигнув x^o.

<u>*Вариант 3*</u>. Полная информированность центра об ограничениях и неполной информированности о целевых функциях $f(x), E\varphi_k(x_k)$. Центр имеет исходную информацию: $G(x), b, X, X_k$. То есть центр имеет полную математическую модель системы и ему не надо затрачивать усилия на идентификацию технологического множества системы. Но центру не известны мотивы и

поведение агентов при выборе ими плановых заданий.

В этом случае можно рассмотреть любой итерационный алгоритм, например, использующий релаксационную схему условного градиента.

1. Центр на каждом шаге получает дополнительную информацию:

$$\operatorname{grad}_{x_k} E\varphi_k(x_k^{(h-1)*}), \operatorname{grad}_{x_k} f(x^{(h-1)*}), x^{(h-1)*}$$

и $c^{(h-1)*}$

2. Центр на каждом h-м шаге в точке $x^{(h+1)*}$ выбирает направление поиска на основе решения задачи ОП

$$\langle \operatorname{grad}_x f\left(x^{(h-1)*}\right), x - x^{(h-1)*}\rangle \to \max$$

или

$$G(x) \geq b, x \in X^{(h-1)}$$

задачи ОСП с учетом условий согласования и ограничения по фонду материального поощрения $\sum_{k=1}^{m} c_k \leq c(x)$.

В последнем случае в задачу оптимизации вводится новая переменная $c_k > 0$. После этого центр сообщает агентам значение плана $x_k^{(h)}$ и размер планируемого поощрения $c_k^{(h)}$.

3. Агент по направлению поиска на каждом шаге выбирает наилучший для себя шаг $\beta_k^{(h)}$, решая в соответствии с гипотезой о локально-оптимальном поведении задачу:

$$E\varphi_k\left[x_k^{(h)*} + \beta_k\left(x_k^{(h)*} - x_k^{(h-1)*}\right)\right] + \\ + \left[c_k^{(h)*} + \beta_k\left(c_k^{(h)*} - c_k^{(h-1)*}\right)\right] \to \max_{\beta_k \in [0,1]}.$$

4. По величине шага агент определяет план $x_k^{(h)*}$, который он обязан выполнить за поощрение $c_k^{(h)*}$, т.е.

$$x_k^{(h)*} = x_k^{(h-1)*} + \beta_k^{(h)}\left(x_k^{(h)} - x_k^{(h-1)*}\right), c_i^{*k} = \\ = c_i^{*(k-1)} + \beta_i^k\left(c_k^{(h)} - c_k^{(h-1)*}\right)$$

и сообщает его величину центру.

5. Если $x_k^{(h)*} \approx x_k^{(h)}$, то центр заканчивает процесс планирования, в противном случае центр запрашивает у агентов дополнительную информацию, т.е. осуществляет переход к п.1 данной схемы.

4. Поведение агента

Как было показано выше, при разработке производственной программы центр использует приближенное описание множества X производственных возможностей агентов. Агент, обладая более детализированной информацией, может при создании центром более привлекательных условий u_k путем решения задачи (2) определять для себя такой вектор $x_k^{(h)}$, который позволяет расширить множество $X_k^{(h)}$ представлений центра о его возможностях, h – шаг итерационного процесса. В данном случае вектор $x_k^{(h)}$ можно рассматривать как вектор способов действия.

Обозначим через $\omega_k = \{\omega_k^{(h)}, h = \overline{1,H}\} \in A_k$ – вектор параметров состояния, определяющий значения вектора действий $x_k^{(h)} = \left\{x_{kj}^{(h)}, j \in [\overline{1,m_k}]\right\} \in X_k^{(h)}$. Можно считать, что этот вектор описывает знание агента возможностей контролируемого им объекта управления. Здесь A_k – множество возможных значений вектора состояния. Будем считать, что агент обладает способностями, знаниями, которые гарантируют существование $\Psi_k : A_k \to X_k^{(h)}$.

Конструктивные возможности технологических узлов и доступный агенту уровень знаний делают справедливым предположение о существовании для k-го агента предельного множества параметров состояния. Обозначим через $O_k^* = \{o_k^* \mid o_k(x_k), x_k \in X_k^{(h)}(\omega_k^{(h)}), \omega_k^{(h)} \in A_k\}$ – множество достижимости или множество предельных возможностей.

Будем предполагать, что агент за счет своих креативных способностей, способности к самообучению и поиску новой информации при соответствующем стимулирующем воздействии центра способен определять такие состояния $\omega_k^{(1)} \in A_k$ и $\omega_k^{(2)} \in A_k$, что возможно $\omega_k^{(2)} \succ \omega_k^{(1)}$, где символ $\succ$ означает «более значимо» и при этом $X_k^{(1)}(\omega_k^{(1)}) \subseteq X_k^{(2)}(\omega_k^{(2)})$. Следовательно, существует такая последовательность $\omega_k^{(1)}, \omega_k^{(2)}, \omega_k^{(3)}, \ldots$, что $\lim_{h \to \infty} o_k^{(h)}(x_k^{(h)}(\omega_k^{(h)})) = O_k^*$. То есть агент путем изучения объекта управления способен определить его предельные возможности для достижения желаемого для себя состояния. Последовательность $\omega_k^{(1)}, \omega_k^{(2)}, \omega_k^{(3)}, \ldots$ будем называть последовательностью вскрытия резервов технологического узла за счет лучшего обслуживания и управления им агентом.

Такая способность агента формировать расширяющееся множество способов действия позволяет определить следующие свойства целевой функции агента и областей достижимости:

$$\forall \omega_k^{(1)}, \omega_k^{(2)} \in A_k, \omega_k^{(2)} \succ \omega_k^{(1)}, X_k^{(1)}(\omega_k^{(1)}) \subseteq \\ \subseteq X_k^{(2)}(\omega_k^{(2)}) \mapsto E\varphi_k(x_k^{(2)}) > E\varphi_k(x_k^{(1)}) \quad (7)$$

Условие (7) означает, что поведение агента при выполнении принципа рациональности соответствует закону повышающихся потребностей, который в литературе по психологии поведения определяет мотивированность и целеустремленность агента.

В реальных условиях агент при превышении некоторого порога значимости $\Delta = E\varphi_k(x_k^{(2)}) - E\varphi_k(x_k^{(1)})$ изменения ценности ситуации целеустремленного состояния по результату способен идентифицировать предпочтительные способы действия и видит открывающиеся возможности при изменении структуры своей информированности (знания). Выработка решения при таком подходе заключается в реализации совокупности последовательных процедур, предназначенных для поиска промежуточных решений, на основании которых агент уточняет свои возможности и формирует окончательное решение. Полный цикл его формирования k-м агентом состоит в выполнении следующих шагов на этапе h:

1. Формирование множеств A_k и $X_k^{(h)}$ на основе знаний, опыта, интуиции и располагаемой информацией о параметрах состояния $\omega_k^{(h)}$. Просмотр множества A_k и формирование точки

$$O_k^{(h)*} = \left\{ \begin{array}{l} o_k^{(h)*} \mid o_k^{(h)*}(x_k^{(h)}), \\ x_k^{(h)} \in X_k^{(h)}(\omega_k^{(h)}), \omega_k^{(h)} \in A_k \end{array} \right\}.$$

Проверка, существует ли $x_k^{(h)*}$ такое, что $o_k^{(h)}(x_k^{(h)*}) = o_k^{(h)*}$. Если – да, то $x_k^{(h)*}$ – это компромиссное решение, а $o_k^{(h)*}$ – прогнозируемая ситуация, в противном случае переход к п.2.

2. Решение задачи поиска потенциально-предпочтительного набора действий $x_k^{(h)*} \in X_k^{(h)}(\omega_k^{(h)})$, позволяющего сформировать вектор $\bar{o}_k^{(h)}$ предельных значений критериев при использовании имеющегося на данный момент знания о правиле $\Psi_k^{(h)}$ и структуре множества A_k. Так как компоненты $\bar{o}_{ki}^{(h)}$, $i = \overline{1, N}$ порознь достижимы, а вместе – нет, то делается попытка найти компромиссное решение. Если агент не согласен попытаться найти компромиссное решение за счет компенсаторных уступок по каждому критерию, которые несколько хуже решения $\bar{o}_k^{(h)}$, то переход к п. 3, иначе к п. 5.

3. Исследование направлений возможного расширения множества A, организация процедур поиска новой информации (знания) о $\omega_k^{(h)} \in A_k$ и правиле $\Psi_k : A_k \to X_k^{(h)}$.

4. Если расширение множества A_k возможно, то переход к п. 1, иначе фиксация ситуации, что компромиссное решение не может быть найдено при выбранном векторе $o_k^{(h)*}$.

5. Получение сведений от агента достаточных для определения вектора $\underline{o}_k^{(h)} \prec o_k^{(h)*}$, где $\underline{o}_k^{(h)}$ минимальные требования агента к принимаемым им во внимание результатам.

6. Выполнение процедуры поиска минимально-предпочтительной точки в пространстве критериев по направлению предпочтения $\underline{o}_k^{(h)}, o_k^{(h)*}$, определение вектора $\omega_k^{(h)*} \in A_k$ и $\underline{x}_k^{(h)*} \in X_k^{(h)}(\omega_k^{(h)*})$ – минимального значения плановых показателей, соответствующих значениям компонент вектора $\underline{o}_k^{(h)}$.

7. Если полученные значения для $\underline{x}_k^{(h)*}, \underline{o}_k^{(h)*}$ принимаются как компромиссное решение, то процедура останавливается, в противном случае переход к п. 8.

8. Для ограничений на $\underline{o}_k^{(h)}$ определяется приоритетная координата $i \in \overline{[1, N_k]}$, по которой

делается расширение множеств A_k и $X_k^{(h)}$, так чтобы $o_{ki}^{(h)}(x_k^{(h)}) = o_{ki}^{(h)} + \Delta_{ki}^{(h)}$, где $\Delta_{ki}^{(h)}$ минимально возможное улучшение, которое является значимым для агента и определяется по его высказываниям о "гибкости" ограничения на основе выполнения процедур поиска дополнительной информации. Переход к п. 1.

Легко видеть, что в данном случае возможно информационное управление со стороны центра. Оно состоит в предоставлении управляемому субъекту определенной информации (информационная картина), ориентируясь на которую агент, имея возможность доступа к собственным источникам информации, выбирает линию своего поведения при формировании встречной информации.

Заключение

1. В процессе определения оптимального плана возможны две ситуации. Первая связана с тем, что оптимальное решение $x^* \in X_{(h)}$, тогда интерактивное взаимодействие центра и агента должно быть направлено на локализацию решения. Если же x^* находится на границе или вне множества $X_{(h)}$, то это говорит о том, что центр не обладает представлениями о предельных возможностях агентов и должен предпринимать усилия для получения информации по расширению множества $X_{(h)}$, для решения задачи вида

$$f^{(h+1)}(x) \to max \ / \ x \in X_{(h+1)} ,$$

где множество допустимых решений шире, чем на $X_{(h)}$ и выполняется условие

$f^{(h)}(x) < f^{(h+1)}(x)$, что определяет мотивированность центра в расширении представлений о множестве $X_{(h)}$.

Такой прием будем называть расширением представлений центра за счет получения информации от агентов об их возможностях.

2. Для случая полной информированности центра показано, что решение задачи ОСП будет реализовано агентами, если найдено стимулирующее воздействие $c_k^{(h)}$, при котором

$$E\varphi_k\left(x_k^{(h)}\right) + c_k^{(h)} \ge \max_{x_k^* \in X_k} E\varphi_k\left(x_k^{(h)*}\right).$$

3. Для случая неполной информированности центра о множестве X_k сходимость к оптимальному решению будет обеспечена путем дополнительного стимулирования агентов за более полное раскрытие ими своих возможностей. Приведена итеративная процедура, использующая первую схему алгоритма блочного программирования Полтеровича. Показано, что достаточным условием ее сходимости будет условие (3).

Библиографический список

[Полтерович, 1969] Полтерович, В.М. Блочные методы выпуклого программирования и их экономическая интерпретация / В.М. Полтерович // Экономика и математические методы. 1969, т. V, № 6

[Виноградов, 2011] Виноградов, Г.П. Моделирование поведения агента с учетом субъективных представлений о ситуации выбора / Г.П. Виноградов, В.Н. Кузнецов // Искусственный интеллект и принятие решений. № 3. 2011. с. 58-72.

[Vinogradov, 2009] Vinogradov, G.P. Model of Individual Decosion-Making: Behavior of Intellectual Agent / G.P. Vinogradov, B. V. Paluch // Interactive Systems and Technologies: the Problems of Human-Computer Interaction. Volume III. – Collection of scientific papers. – Ulyanovsk: UlSTU, 2009. P. 127-134.

[Чхартишвили, 2004] Чхартишвили, А.Г. Теоретико-игровые модели информционного управления / А.Г. Чхартишвили М.: ЗАО «ПМСОФТ», 2004.

ITERATIVE METHODS OF OPTIMIZATION WITH COORDINATED IDENTIFICATION ELEMENOTOV CHOICE MODELS

Vinogradov, GP Kuznetsov, VN

Tver State Technical University, Tver, Russia

WGP272NG@mail.ru

The convergence of interactive procedures agreed in the optimization problem of controlling the evolution of organizational and technological systems in different variants of the center of awareness about the possibilities of agents. The conditions under which the procedure is possible to build an interactive exchange of information between the center and its agents, ensuring the convergence of the solution of the consistent optimization.

Keywords: active system, an intelligent agent, consistent optimization, decision making