УДК 51-77

ПРИМЕНЕНИЕ СТРАТЕГИЧЕСКОЙ РЕФЛЕКСИИ В МОДЕЛИ ФОРМИРОВАНИЯ КОМАНДЫ

В.О. Корепанов
Институт проблем управления им. В.А. Трапезникова РАН
Россия, 117997, Москва, Профсоюзная ул., 65
E-mail: vkorepanov@ipu.ru

Ключевые слова: формирование команды, управление, стратегическая рефлексия, теория игр.

Аннотация: Рассматривается модель формирования команды агентов, определенная так, что целью действий команды является решение многомерной задачи оптимизации с ограничением типа равенства. Вводятся модели поведения агентов, использующие стратегическую рефлексию (представления о моделях поведения оппонентов). Исследуются возможности использования таких моделей для достижения цели команды.

1. Введение

Рассмотрим модель формирования команды [1]. Задано множество агентов $N = \{1,2,...,n\}$. Каждый агент может выбирать действие $y_i \geq 0$, неся при этом затраты $c(y_i, r_i)$, где r_i – тип агента, отражает его эффективность (тип функции затрат). Пусть цель совместных действий агентов состоит в достижении суммарного действия

(1) $$\sum_{i \in N} y_i = R$$

При минимальных суммарных затратах

(2) $$\sum_{i \in N} c(y_i, r_i) \to \min_{\{y_i \geq 0\}}$$

Без ограничения общности будем считать $R = 1$. В случае затрат вида Кобба-Дугласа $c(y,r) = y^{\alpha} r^{1-\alpha} / \alpha$, решение задачи (1)-(2) выглядит так:

(3) $$y_i^* = r_i / \sum_j r_j \, .$$

В работе [2] предложена модель динамики формирования взаимных представлений агентов о типах друг друга, т.е. динамики информационной рефлексии (далее ИР). При этом считается, что агенты могут решить задачу (1)-(2) и на основании решения восстанавливают такие представления о типах оппонентов, которые способны объяснить наблюдаемые действия оппонентов. В данной работе совершается попытка использовать подход стратегической рефлексии [3]: агенты знают типы друг друга, но не знают точно модели поведения друг друга. Иными словами, передать числовую информацию о своих типах друг другу они могут, а быть уверенными в том, что каждый решает и/или способен решить оптимизационную задачу (1)-(2) они не могут, в противоположность подходу, использованному в [2].

2. Модели стратегической рефлексии

Рассмотрим простейшую модель стратегической рефлексии (СР). Пусть есть множество N^0 «простых» агентов – т.н. агенты нулевого ранга рефлексии. Будем считать, что агенты 0-ранга знают общее число агентов *n*, пусть они выбирают своим действием $y_i^0 = 1/n$.

Если все агенты 0 ранга, то их действия удовлетворяют (1), но критерий (2) не оптимален. Пусть есть «умные» агенты – агенты ранга 1, которые знают решение (1)-(2), знают, что есть агенты 0 ранга, но им неизвестно точно их число и число агентов 1 ранга. Обозначим n^0 – число агентов 0 ранга, N^1 – множество агентов 1 ранга, n_i^{10} - представления агента *i* ранга 1 о числе агентов ранга 0 и n_i^{11} – представления о числе агентов 1 ранга и N_i^{11} – представления о множестве агентов 1 ранга. Тогда действие агента ранга 1 будет

$$y_i^1 = (1 - n_i^{10}/n) r_i / \sum_{j \in N_i^{11}} r_j .$$

Если процесс выбора действий происходит в форме повторяющейся игры, а агенты наблюдают все действия на прошлом шаге $x(t-1) = (x_1(t-1), ..., x_n(t-1))$, то они могут найти множество N^0 (это агенты, у которых действие равно $1/n$), понять, что остальные – агенты 1 ранга и тем самым за одно наблюдение прийти к истинному представлению о рангах агентов. На следующем шаге действия агентов 1 ранга будут оптимальны с учетом наличия агентов 0 ранга:

$$y_i^1 = (1 - n^0/n) r_i / \sum_{j \in N^1} r_j .$$

В рассмотренном примере, во-первых в общем случае не достигается оптимум решения задачи (1)-(2), во-вторых, в случае если агенты наблюдают только агрегат действий $\sum_{i \in N} x_i(t-1)$, то восстановить точные представления о параметрах оппонентов становится практически невозможно – возникает дискретная переборная задача с возможно не единственным решением.

Рассмотрим динамическую модель стратегической рефлексии, построенную на идее индикаторного поведения [3]. Заданы начальные действия агентов $y_i(0)$. Агент 0 ранга наблюдает сумму действий агентов на прошлом шаге и смещает свое действие в сторону действия, при котором удовлетворялось бы (1) при неизменных действиях остальных:

$$y_i(t+1) = (1-\gamma_i) y_i(t) + \gamma_i (1 - \sum_{j \neq i} y_j(t)) = y_i(t) + \gamma_i (1 - \sum_{j \in N} y_j(t)), \tag{4}$$

где $\gamma_i \in [0,1]$ – величина шага в сторону удовлетворения (1)

Агент 1 ранга считает всех остальных агентами 0 ранга, знает модель (4) и соответственно может предсказать их действия, тогда его действием будет:

$$y_i^1(t+1) = y_i^1(t) + \gamma_i (1 - \sum_{j \neq i} y_j(t+1) - y_i^1(t)) . \tag{5}$$

Модель поведения агентов ранга 2 и выше аналогична (5), за тем исключением, что у таких агентов должны быть представления о числе оппонентов ранга ниже своего, например, пусть N_i^{20} и N_i^{21} - множества агентов ранга 0 и 1 в представлении агента *i* ранга 2. Тогда его действие рассчитывается по выражению:

$$y_i^2(t+1) = y_i^2(t) + \gamma_i (1 - \sum_{j \in N_i^{20}} y_j(t+1) - \sum_{j \in N_i^{21}} y_j^1(t+1) - y_i^2(t)) . \tag{6}$$

Не забудем, что стоит задача (1)-(2), для которой мы пытаемся использовать стратегическую рефлексию. В общем случае у построенной модели динамики возникает две проблемы: условия сходимости динамики (4)-(6) и задача сходимости именно к решению (1)-(2).

Пусть есть только агенты 0 ранга и $\gamma_i = 1$. Тогда $y_i(t+1) = 1 - \sum_j y_j(t)$. Из такой динамики следует $y_i = y_j$, динамика будет расходящаяся, единственная неподвижная точка $y = (1/(n+1),...,1/(n+1))$, к которой динамика может дойти (за один шаг), если $\sum_i y_i(0) = n/(n+1)$. В любом случае неподвижная точка не является решением (1)-(2).

Заметим из (3)-(5), что при одинаковой информированности у агентов одного ранга действия всегда будут одинаковы, если у них одинаковы γ_i. Исходя из (3) в общем случае нам это неинтересно. А если $\sum_i \gamma_i = 1$, то действия на (3) удовлетворяют (1):

$$\sum_i y_i(t+1) = \sum_i y_i(t) + \sum_i \gamma_i (1 - \sum_i y_i(t)) = 1.$$

Тогда получается, что динамика (3) останавливается на первом шаге. Зная решение (3), получаем, что эти же действия на 1 шаге будут удовлетворять (2), если

(7) $$\forall i : (y_i(0) \sum_j r_j - r_i) / r_i = \sum_j y_j(0) - 1.$$

В частности, при $\gamma_i = r_i / \Sigma r_j$:

$$y_i(0) / \sum_j y_j(0) = r_i / \sum_j r_j .$$

Интересно далее рассмотреть возможности решения (1)-(2) с введением агентов 1 и более высокого рангов.

Рассмотрим также другую модель стратегической рефлексии, в которой агенты 0 ранга действуют также по (4). А агенты 1 ранга знают решение (3) и пытаются своими действиями привести остальных к нему, т.е. это агенты, которые управляют.

Рассмотрим частные случаи при $\gamma_i = 1$ и $n = 3$. Пусть сначала у нас агент 1 ранга 0 и агенты 2 и 3 ранга 1. Тогда при $y_2(0) = y_2(1) = y_2^*$ и $y_3(0) = y_3(1) = y_3^*$, действие первого игрока на первом шаге $y_1(1) = 1 - y_2^* - y_3^* = y_1^*$. То есть два «умных» агента могут сформировать команду с одним «глупым» агентом.

Рассмотрим далее случай с одним «умным» агентом 3. Действия агентов 1 и 2 0 ранга на первом шаге: $y_i(1) = 1 - y_{3-i}(0) - y_3(0)$, на последующих шагах получается следующее общее выражение:

$$y_i(t) = y_i(t-2) - y_3(t-2) - y_3(t-1).$$

Агент 3 должен выбрать так свои действия, чтобы $y(t) = (y_1(t), y_2(t), y_3(t))$ сошлись к y^* и сошлись как можно быстрее, при этом с некоторого $T \geq 0$: $y_3(T) = y_3^*$.

Для случая $T = 0$: $y_i(2t) = y_i(0)$, $y_i(2t+1) = 1 - y_{3-i}(0) - y_3^*$. Для схождения необходимо $y_1(0) = 1 - y_2(0) - y_3^*$, т.е. $y_1(0) + y_2(0) + y_3^* = 1$. Тогда для схождения к решению необходимо $y_1(0) = y_1^*, y_2(0) = y_2^*$.

Для случая $T = 1$, агент 3 должен определить $y_3(0)$, а далее $\forall t \geq 1 : y_3(t) = y_3^*$. Получаем следующие соотношения: для $t \geq 1$: $y_i(2t) = y_i(0) + y_3(0) - y_3^*$, $y_i(2t+1) = 1 - y_{3-i}(0) - y_3(0)$. Для схождения необходимо: $y_3(0) = (1 - y_2(0) - y_3(0) + y_3^*)/2$. Для схождения к решению (3) необходимо выполнение условия

(8) $$y_1(0) - y_2(0) = y_1^* - y_2^* .$$

Таким образом мы рассмотрели две модели использования стратегической рефлексии для задачи формирования команды. Проанализированы их некоторые частные случаи, получены условия для формирования команды, например (7) и (8).

Интересным кажется анализ предложенных моделей СР в более общих случаях, в т.ч. и их численный анализ. Далее интересно их сравнение, например, с распределенными алгоритмом оптимизации ADMM [], а также варианты применения в нем стратегической рефлексии.

Работа выполнена при поддержке проекта РФФИ № 17-07-01550 А.

Список литературы

1. Burkov V.N. et al. Introduction to theory of control in organizations. CRC Press, 2015.
2. Novikov D.A. Teams: Building, Adaptation and Learning // New Frontiers in Information and Production Systems Modelling and Analysis. Springer, 2016. P. 3-34.
3. Novikov D.A., Korepanov V.O. The reflexive partitions method in models of collective behavior and control // Automation and Remote Control. 2012. Vol. 73, No. 8. P. 1424-1441.
4. Boyd S., Parikh N., Chu E., (2011) Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers // Foundations and Trends in Machine Learning. 2011. Vol. 3, No. 1, P. 1-122.