

М. В. КРЫЖАНОВСКИЙ, Б. М. МАГОМЕДОВ
Институт оптико-нейронных технологий РАН, Москва
iont@iont.ru, shico@inbox.ru

РЕШЕНИЕ ЗАДАЧИ УПРАВЛЕНИЯ ГРУППОЙ АГЕНТОВ ПРИ ВЫБОРЕ ЦЕЛЕЙ

Аннотация

В работе рассматривается задача планирования действий группы агентов при выборе целей. Представлен новый алгоритм, работающий заметно быстрее, чем имеющиеся в настоящий момент алгоритмы [1].

1. Введение

Основные принципы децентрализованного коллективного управления группой агентов заключаются в следующем [2]:

1. Каждый объект группы должен самостоятельно решать задачу планирования своих действий в текущей ситуации.
2. Выбор управления агентом осуществляется только на основе информации о текущей ситуации в среде, текущих состояниях и действиях других агентов группы на небольшой интервал времени.
3. Под оптимальным управлением понимается такой выбор действия каждого из объектов, которое дает максимально возможное приращение в целевой функционал.

На основе этих принципов в [2] было продемонстрировано решение модельной задачи по управлению группой агентов при выборе целей, алгоритм решения которой был основан на частичном переборе. В данной работе продемонстрирован другой подход, упрощающий решение задачи.

2. Постановка задачи и ее решение.

Пусть имеется m целей и коллектив из N агентов R_j ($j = 1..n$), которые должны эти цели поразить в условиях заранее неизвестных противодействий противника, т.е. часть агентов может быть уничтожена (задача с переменным количеством участников). Ущерб от поражения i -й цели определяется выражением

$$\Phi_i = \Phi_i(n_i),$$

где n_i – число агентов коллектива, поразивших i -ю цель.

Нелинейная функция Φ_i имеет следующий вид, представленный на рис.1, где n_i^{max} – число агентов, попадание которых в данную цель

гарантирует ее уничтожение, $\Phi_i^{\max} = 1$ – максимальное значение ущерба, получаемое в результате уничтожения i -й цели.

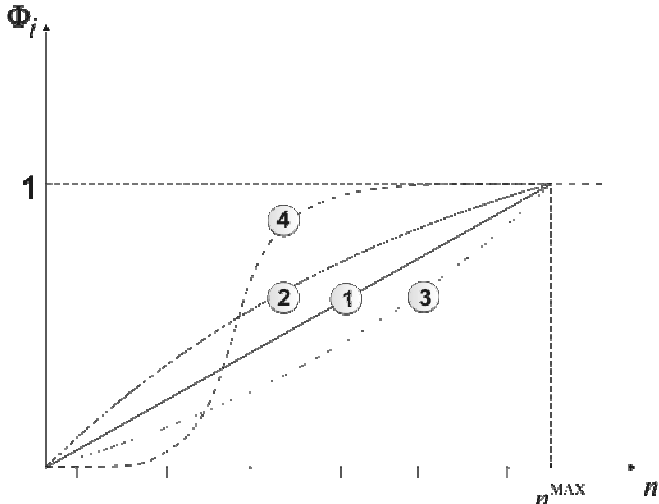


Рис.1. Возможные виды функции $\Phi_i(n)$

В качестве функционала для оптимизации необходимо выбрать величину суммарного ущерба наносимого противнику в зависимости от степени поражения всех целей т.е. доставить максимум функции

$$\Phi = \sum \Phi_i(n_i), \quad (1)$$

где суммирование ведется по индексу i ($i=1..m$), при условии что

$$\sum n_i = N \quad (2)$$

Задача определения максимума функции Φ при условии (2) сводится к задаче определения максимума функционала $W = \Phi - \mu \cdot (\sum n_i - N)^2$, где μ – множитель Лагранжа. Рассмотрим представление функционала $W = W(n_1, n_2, \dots, n_m) = W(\mathbf{n})$, в окрестности точки (вектора) \mathbf{n}_0 т.е. $\mathbf{n} = \mathbf{n}_0 + \mathbf{u}$, а \mathbf{u} – вектор, компоненты которого принимают значения: $-1, 0, +1$. Тогда его представление имеет вид

$$W(\mathbf{u}) = \sum_{i=1}^m p_i u_i - \mu/2 \cdot \left(\sum_{i=1}^m u_i - N + N_0 \right)^2 \quad (3)$$

При решении задачи градиентным методом компоненты считаются по формулам:

$$u_i = \alpha \cdot dW(\mathbf{u}_o) / du_i$$

$$u_i = \alpha \cdot [p_i - \mu \cdot (\sum_{i=1}^m u_i - N + N_o)] = \alpha \cdot (p_i - r_o);$$

$$r_o = \mu \cdot (\sum_{i=1}^m u_i - N + N_o).$$

Однако такое использование итерационного процесса определения вектора \mathbf{u} затруднительно в силу трудности выбора множителя Лагранжа. Поэтому необходимо модифицировать выбор вектора \mathbf{u} , так чтобы влияние штрафа было минимальным, а именно

$$u_i = \alpha \cdot (p_i - r_o + b);$$

Для этого определим значение штрафа после одного шага итерации:

$$r_o = \mu \cdot (\sum_{i=1}^m u_i^{(o)} - N + N_o); \quad u_i^{(1)} = u_i^{(o)} + \delta u_i;$$

$$\delta u_i = \alpha \cdot (p_i - r_o + b);$$

$$r_1 = r_o \cdot (1 - m\mu\alpha) + m\mu\alpha \cdot (b - \frac{1}{m} \sum_{i=1}^m p_i);$$

Откуда следует, что если выражение $(b - \frac{1}{m} \sum_{i=1}^m p_i) = 0$, то значение штрафа может уменьшаться. Такая модификация в выборе вектора \mathbf{u} означает, что мы можем модифицировать искомым функционал $\mathbf{W}(\mathbf{u})$ т.е.

$$\mathbf{W}(\mathbf{u}) = \sum_{i=1}^m (p_i - b) \cdot u_i - \mu/2 \cdot (\sum_{i=1}^m u_i - N + N_o)^2; \quad b = \frac{1}{m} \sum_{i=1}^m p_i; \quad (4)$$

Такое представление функционала $\mathbf{W}(\mathbf{u})$ позволяет без трудностей решать задачу оптимизации, а правило нахождения начального решения в этом случае будет формулироваться следующим образом: если $(p_i - b) > 0$ то $u_i = +1$, в противном случае $u_i = -1$;

3. Алгоритм решения.

Каждый агент характеризуется параметрами: порядковым номером L , направлением на цель C , функцией цели Φ . Информация о других агентах содержится в таблицах l_j , c_j , ϕ_j . По этим данным (в момент принятия решения) определяется число агентов n_i нацеленных на i -ю цель, производные p_i , величина b т.е. определяется ситуация в среде. На этой

основе агент, направленный на i -ю цель определяет i -ю компоненту градиента и соответственно компоненту u_i .

Если каждой цели соотносится множество агентов направленных поразить ее, тогда перераспределение целей между агентами коллектива определяется путем анализа компонент вектора \mathbf{u} . Такому перераспределению подлежат агенты, принадлежащие к целям для которых компоненты \mathbf{u} отрицательны, а именно: среди агентов, принадлежащих i -й цели, изменению цели подлежит агент, имеющий максимальный порядковый номер. Выбор же цели для такого агента осуществляется путем коммуникационного обмена информацией с другими агентами коллектива.

4. Заключение

Приведенный алгоритм решения задачи существенно упрощает решение задачи по децентрализованному управлению и на порядок уменьшает вычислительную сложность алгоритма ($\sim m^2$).

Работа выполнена при поддержке РФФИ (проекты №04-07-90038 и №05-07-90049)

Список литературы

1. Каляев И.А., Гайдук А.Р., Капустян С.Г. Распределенные системы планирования действий коллективов роботов. // М.: Янус-К, 2002.
2. Каляев И.А. Использование принципов коллективного принятия решений при управлении группой автоматических лифтов. // Мехатроника, №4, 2001.