

М. В. КРЫЖАНОВСКИЙ, Б. М. МАГОМЕДОВ
Институт оптико-нейронных технологий РАН, Москва
iont@iont.ru , bashir.magomedov@gmail.com

НЕЙРОСЕТЕВОЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ УПРАВЛЕНИЯ ГРУППОЙ АГЕНТОВ ПРИ ВЫБОРЕ ЦЕЛЕЙ

Аннотация

В работе рассматривается задача планирования действий группы агентов при выборе целей. Предложен нейросетевой алгоритм, работающий заметно быстрее, чем имеющиеся в настоящий момент алгоритмы. Такой подход существенно упрощает решение задачи по децентрализованному управлению. Предложена модификация оптимизируемого функционала, позволяющая обойти ряд трудностей, связанных с выбором коэффициента Лагранжа.

1. Введение

На основе принципов децентрализованного управления в [1,2] было продемонстрировано решение модельной задачи по управлению группой агентов при выборе целей. В данной работе предлагается иной подход, упрощающий решение этой задачи, позволивший добиться двух преимуществ. Во-первых, в развитом в настоящей работе нейросетевом подходе действия каждого из агентов направлены на улучшение состояния всей мультиагентной системы в целом, а не своего личного состояния. Во-вторых, предложена модификация оптимизируемого функционала, позволяющая ограничить вхождение системы в запретную область штрафов и, тем самым, позволяющая обойти ряд стандартных трудностей, связанных с подбором коэффициента Лагранжа.

2. Постановка задачи при централизованном управлении

Пусть имеется m целей и коллектив из N агентов R_k , которые должны эти цели поразить. Предполагается, что действия агентов происходят в условиях заранее неизвестного противодействия противника. В этом случае необходимо рассматривать задачу с переменным количеством участников, поскольку часть агентов может быть уничтожена. Пусть ущерб от поражения i -й цели определяется выражением $\Phi_i = \Phi_i(n_i)$, где n_i – число агентов, поразивших i -ю цель, а $\Phi_i(n_i)$ – нелинейные монотонно нарастающие функции. Возможный вид

этих функций представлен на рис.1, где n_{max} – число агентов, попадание которых в данную цель гарантирует $\Phi(n_{max})=1$.

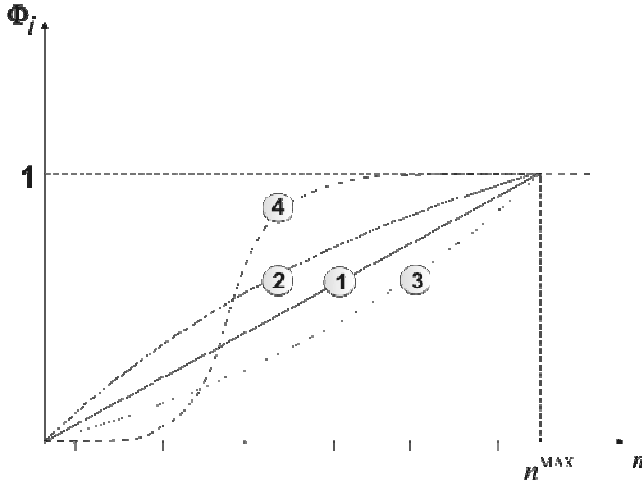


Рис.1. Возможные виды функции $\Phi_i(n)$

В качестве функционала для оптимизации выбрана величина суммарного ущерба, наносимого противнику в зависимости от степени поражения всех целей. Иными словами, задача оптимизации действий мультиагентной системы сводится к отысканию максимума функции

$$\Phi = \sum_{i=1}^m \Phi_i(n_i), \text{ при дополнительном условии } \sum_{i=1}^m n_i = N$$

Задача определения максимума функции Φ при таком условии сводится к задаче определения максимума функционала

$$W = \Phi - \mu \left(N - \sum n_i \right)^2 \quad (1)$$

где μ - множитель Лагранжа. Поставленная задача решается методами градиентного спуска [3]. Однако в рассматриваемом нами случае такой подход не применим, поскольку поиск экстремума W ведется в конфигурационном пространстве переменных. Поэтому их приращения могут быть только целочисленными, т.е их следует описывать некоторым вектором \mathbf{u} , компоненты которого принимают значения $u_i = -1, 0, +1$. Поэтому, для нахождения оптимума функционала W будем использовать следующий итерационный процесс. Зададим начальное значение числа

агентов N_0 и некоторое случайное начальное распределение по целям $\mathbf{n}=\mathbf{n}_0$. На первом шаге итерационного процесса рассмотрим представление функционала $W=W(\mathbf{n})$ в окрестности точки \mathbf{n}_0 , положив $\mathbf{n}=\mathbf{n}_0+\mathbf{u}$, где \mathbf{u} – искомый вектор приращений. Тогда, проводя в (5) разложение в окрестности точки \mathbf{n}_0 и заменяя малые приращения переменных $d\mathbf{n}$ компонентами вектора \mathbf{u} , искомое представление с точностью до несущественной константы представим в виде

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

или в виде

$$W(\mathbf{u}) = -\frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m T_{ij} u_i u_j + \sum_{i=1}^m h_i u_i \quad (3)$$

где введены следующие обозначения

$$h_i = p_i + \mu(N - N_0), \quad T_{ij} = \mu, \quad p_i = d\Phi_i / dn_i \Big|_{\mathbf{n}=\mathbf{n}_0} \quad (4)$$

Как нетрудно заметить, выражение (3) совпадает с определением понятия энергии в нейросетевой модели Хопфилда [4], если принять, что T_{ij} - это матрица межсвязей, а h_i - это пороги нейронов. Соответственно этому, поиск максимума квадратичного функционала (2) будем вести следуя обычной [5,6] нейросетевой процедуре оптимизации: нейронная сеть инициализируется в начальное состояние $u_i = 0$ и далее запускается процедура подъема по энергетической поверхности (3). Определив вектор \mathbf{u} , доставляющий функционалу $W=W(\mathbf{n})$ максимум в окрестности точки \mathbf{n}_0 , мы переходим в окрестность новой точки $\mathbf{n}_j = \mathbf{n}_0 + \mathbf{u}$: в этой точке находятся новые значения величин величины p_i (производные функций Φ_i в точке \mathbf{n}_j) и снова проводится нейросетевая оптимизация функционала (3). Этот итеративный процесс продолжается до получения решения с необходимой точностью.

Модификация алгоритма. Однако прямое использование итерационного процесса определения вектора \mathbf{u} затруднительно в силу трудности подбора множителя Лагранжа в функции штрафа. Поэтому выбор вектора \mathbf{u} был модифицирован, так чтобы влияние штрафа после одной итерации было минимальным. Для этого мы изменили искомый функционал $W(\mathbf{u})$ к виду:

$$W(\mathbf{u}) = \sum_{i=1}^m (p_i - \bar{p}) u_i - \frac{\mu}{2} \left(N_0 - N + \sum_{i=1}^m u_i \right)^2 \quad (5)$$

где
$$\bar{p} = m^{-1} \sum_{i=1}^m p_i \quad (6)$$

а также условие инициализации сети Хопфилда, задавая начальные условия в виде $u_i = \text{sgn}(p_i - \bar{p})$.

3. Результаты численных экспериментов

Анализ результатов экспериментов показывает, что поиск максимума функционала $W(\mathbf{n})$ протекает успешно: достигается глобальный экстремум в случае применения функций Φ вида 1÷3 (рис.1) и локальный- при использовании функций 4-го вида.

Видоизменение алгоритма. Чтобы избежать попадания в локальный экстремум и, по возможности, создать алгоритм, доставляющий глобальный оптимум, расчет величин p_i был изменен. Для этого изменялся сам вид функции участвующей в вычислении производной. По функции $\Phi_i(n)$ определялась производная $p_i = d\Phi/dn_i$ для всего диапазона $0 < n < N$ и находились ее максимальное значение p_i^{max} и соответствующая координата n_i^{max} . В диапазоне $0 < n_i < n_i^{max}$ полагалось $p_i = p_i^{max}$, а в остальной части диапазона производная p_i определялась прежним выражением. Такое видоизменение расчета позволило находить глобальный оптимум для функции Φ .

Таким образом, в условиях централизованного управления мы решили задачу определения количества агентов направляемых на каждый объект-цель. При этом, определение цели для каждого из агентов определяется просто. Агенты, имеющие направление на цель, для которой количество объектов не изменилось или увеличилось, свое направление не изменяют. Для агентов, принадлежащих другим целям, переопределение направления определяется их порядковыми номерами и "расстояниями" к ближайшим целям.

4. Решение задачи для агентов, управляемых автономно

Решение задачи по выбору цели для агентов, управляемых автономно, ведется на основе решения задачи при централизованном управлении, для которого характерно представление о целях как о самостоятельных объектах конкурирующих между собой за увеличение количества агентов направленных на нее. Поэтому выражение (5) для $W(\mathbf{u})$ мы представим как сумму

$$W(\mathbf{u}) = \sum_{i=1}^m W_i(\mathbf{u}) \quad (7)$$

где для каждой цели вводится свой оптимизационный функционал

$$W_i(\mathbf{u}) = (p_i - \bar{p})u_i - \frac{\mu}{2}(N_0 - N + u_i) \left(N_0 - N + \sum_{s=1}^m u_s \right) \quad (8)$$

Для того, чтобы агент мог определять свое движение он должен иметь информацию о целях, информацию о текущих состояниях всех агентов и об их намерениях изменить направления. Поэтому каждый агент имеет таблицу, которая описывает состояние целей: функция каждой i -ой цели Φ_i , количество агентов направленных на нее и их номера, производную от функции цели и требование об изменении числа агентов (компоненты вектора \mathbf{u}). Содержание таблиц до начала движения задается извне.

Обмениваясь информацией с другими членами коллектива в процессе движения агент R_k , получает производные p_i и вычисляет ее среднее значение \bar{p} . Это дает возможность определить агентам начальное приближение вектору \mathbf{u} . На этой основе агент R_k , направленный на i -ю цель определяет величину $W_i(\mathbf{u})$ и, соответственно, компоненту u_i . После определения агентами вектора \mathbf{u} и занесения его в таблицу наступает этап принятия решения.

Перераспределение целей между агентами коллектива определяется путем анализа компонент векторов \mathbf{p} и \mathbf{u} . В первую очередь удовлетворяются потребности в изменении пары целей, для которых отклонение от среднего $|p_i - \bar{p}|$ максимально и у которых компоненты \mathbf{u} разных знаков. Агенты, принадлежащие к целям, для которых компоненты \mathbf{u} отрицательны, изменяют направление, а именно: среди агентов, принадлежащих i -й цели, изменению цели подлежит агент, имеющий максимальный порядковый номер. Приняв решение о выборе цели, агент сообщает о нем другим членам коллектива. Таблицы агентов модифицируются, а соответствующие компоненты вектора \mathbf{u} обнуляются. После изменения направления агентов коллектива, процесс, описанный выше повторяется.

5. Заключение

Приведенный выше алгоритм существенно упрощает решение задачи по децентрализованному управлению и на порядки уменьшает

вычислительную сложность алгоритма. Анализ вычислительной сложности и результаты множественных экспериментов показывают, что скорость алгоритма с ростом числа целей m увеличивается в m^2 раз, по сравнению с стандартным подходом.

Предложенная модификация оптимизируемого функционала, заключающаяся в замене градиентов (p_i) их отклонением от среднего ($|p_i - \bar{p}|$), резко сузила возможность вхождения системы в запретную область штрафов. При этом, алгоритм стал значительно менее чувствительным к выбору значения коэффициента Лагранжа, позволяя избегать стандартных проблем, возникающих при оптимизации функционалов с штрафными членами.

Работа выполнялась в рамках программы "Фундаментальные основы информационных технологий и систем" (проект 1.8) и при финансовой поддержке РФФИ (проект 05-07-90049).

Список литературы

1. И.А.Каляев, А.Р.Гайдук, С.Г. Капустян. 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.
3. G.E.Forsythe, C.V.Moler. Computer Solution of Linear Algebraic Systems. Englewood Cliffs, New Jersey, PrenticeHall,1967.
4. J.J.Hopfield. Neural Networks and physical systems with emergent collective computational abilities. Proc.Nat.Acad.Sci.USA. 79, 2554-2558 (1982).
5. J.J.Hopfield, D.W.Tank. Neural computation of decisions in optimization problems. Biological Cybernetics 52, 141-152 (1985).
6. Б. В. Крыжановский, Б. М.Магомедов, А. Л. Микаэлян. Доменная модель нейронной сети. Доклады Академии наук, том. 401, №4, с. 462-466