

УДК 004.896

МОДЕЛИ СТАЙНОГО ПОВЕДЕНИЯ РОБОТОВ¹

А.А. Кулинич (kulinich@ipu.ru)

Институт проблем управления РАН, Москва

Аннотация. Предложены модели стайного поведения агентов (роботов) на основе критериев возможности достижения цели и их взаимной полезности при выполнении совместных действий. Рассмотрены стайные алгоритмы «ленивых» и «эгоистичных» агентов.

Ключевые слова: Роботы, агенты, стайный алгоритм, кооперация.

Введение

Идеология командной работы агентов позволяет с помощью множества несложных технических устройств (роботов), выполнять сложные задания в результате кооперации их возможностей. В настоящее время решение этой задачи исследуется в рамках теории группового управления роботами [Каляев и др. 2009], а также в теории мультиагентных систем и искусственного интеллекта. В теории мультиагентных систем и искусственного интеллекта исследуются команды интеллектуальных роботов, с так называемой ментальной архитектурой. В этой работе такие подходы не рассматриваются.

В теории группового управления интерес представляют роевые и стайные подходы к управлению группой роботов. Роевые алгоритмы, основанные на локальном взаимодействии множества однородных роботов (роя роботов), обеспечивают их скоординированное движение к цели, обход препятствий и т.д. [Павловский и др., 2015]. Стайные алгоритмы предполагают взаимодействие функционально разнородных роботов на основе их общих знаний о состоянии среды функционирования и индивидуальных правил поведения агентов в стае.

В живой природе существует множество примеров стайного поведения насекомых и животных [Карпов, 2016]. Считается, что такое поведение является оптимальным. Поэтому, в качестве одной из возможностей повышения качества командной работы агентов (роботов) является

¹ Работа выполнена при поддержке ОФИ_м (проект № 16-29-04412) и РФФИ (проект № 15-01-07900)

моделирование поведения стайных насекомых или животных в алгоритмах командной работы искусственных агентов. В этой работе, на основе исследований социальных психологов предложены модели стайного поведения агентов в процессах образования и функционирования команд агентов.

1. Стайное и коллективное поведение агентов (роботов)

Теоретические вопросы создания команд технических устройств (роботов) с системных позиций были рассмотрены в работе [Каляев, и др., 2009]. Здесь группа роботов $R_i, i=1, \dots, N$, каждый из которых характеризуется вектором состояний $R=(r_1, \dots, r_n)$ (вектором значений его характеристик) помещен в среду E , которая также характеризуется вектором состояния, $E=(e_1, \dots, e_n)$. Пара векторов $S=(R, E)$ характеризует состояние системы «группа роботов – среда». Каждый робот может выполнять действия A , изменяющие состояние системы «группа роботов – среда». Непредвиденные изменения среды характеризуются вектором $g(t)$. Действия роботов направлены на то, чтобы перевести текущее состояние системы S_c в некоторое желаемое состояние S_c^f . Изменение состояния системы «группа роботов – среда», как результата совместных действий роботов описывается системой дифференциальных уравнений. Ставится задача определения совместных действий роботов, позволяющих достичь поставленной цели S_c^f . Оценка действий всех роботов для достижения цели осуществляется с помощью оценочного функционала (1).

$$Y_c = \Phi(S_c^f, t_f) + \int_{t_c}^{t_f} F(S_c(t), A_c(t), g(t), t) dt \quad (1)$$

В выражении (1) первое слагаемое характеризует полезность целевого состояния, а второе качество группового управления. Задача взаимодействия группы роботов как команды здесь определена как задача группового управления (управления группой роботов), для решения которой находятся векторы действий всех роботов $A_c(t)$ на промежутке времени $[t_c, t_f]$, которые обеспечивали бы экстремум функционала (1).

В работе рассматривается два подхода к выработке группового управления: централизованный и децентрализованный. При централизованном управлении векторы действий всех агентов вычисляются в едином центре до решения ими задачи. При децентрализованном управлении роботы обладают автономностью, способны обмениваться информацией и координировать свои действия для достижения индивидуальных целей и общей цели. Это позволяет получить свойства независимости команды от центра, что повышает ее устойчивость к его отказам и отказам отдельных агентов.

Возможны два вида децентрализованного управления: стайное и

коллективное. Как первый, так и второй вид децентрализованного управления предполагает командное поведение роботов для достижения общей цели. Однако, при стайном управлении агенты (роботы) не обмениваются между собой информацией, а строят свою стратегию достижения цели на основе анализа текущего состояния среды функционирования. При коллективном управлении, агенты по каналу связи договариваются о совместном достижении поставленной цели.

В классической постановке стайного поведения каждый робот пытается достичь экстремума функционала (1) самостоятельно, анализируя состояние среды функционирования [Каляев, и др., 2009].

Алгоритм действий роботов заключается в следующем. Роботы анализируют состояние среды функционирования, и выполняют действие для достижения цели, изменяя состояние среды. При этом они не знают, какие действия предприняли другие роботы. Поэтому, на следующем шаге роботы вновь анализируют состояние среды функционирования, которое уже отражает результат действия всех роботов, и каждый из них принимает новые действия для достижения цели. Команда из роботов образуется, если у них есть общая цель. Это необходимое, но не достаточное условие является критерием образования команды в стайном алгоритме. В стайном алгоритме роботы не знают, что они команда, сколько роботов в команде и какие у каждого из роботов возможности для достижения общей цели. В этом случае говорить об оптимальности действий всей команды не приходится, поскольку нет возможности координации их совместных действий. Поведение такой команды будет напоминать суетливое поведение толпы. Действительно, если в команде из множества роботов, есть один или два робота, которые самостоятельно или, объединив усилия способны достичь цели, то остальные члены команды могут «отдохнуть», не растрчивая свои ресурсы.

Повысить качество командной работы стаи роботов можно, если кроме информации о состоянии среды функционирования станут общедоступными сведения о целях роботов и имеющихся у них ресурсах для достижения цели. Далее рассмотрим несколько стайных алгоритмов, позволяющих повысить качество их работы.

2. Стайное поведение агентов (постановка задачи)

Общая постановка стайной работы команды роботов следующая. Рассматривается множество агентов (роботов) $A = \{a_i\}$, обладающих свойствами (параметрами) $F = \{f_i\}$. Для каждого свойства каждого из объектов (агентов) этой среды определено упорядоченное множество их возможных значений, $Z = \{Z_i\}$, где $Z_i = \{z_{i1}, \dots, z_{iq}\}$, $z_{iq+1} \succ z_{iq}$, $q = 0 \dots n-1$.

Среда функционирования агентов определяется как прямые

произведение множеств значений всех свойств агентов, $SF = \times_i Z_i$.

Вектор значений всех агентов $Y(t) = (z_{1j}, \dots, z_{nb})$, $z_{ij} \in Z_i, \forall i$, определяет состояние среды функционирования.

Динамика изменения состояния среды функционирования происходит в случаях изменения агентами своих свойств и представляется как отображение:

$$W: Y(t) \rightarrow Y(t+1), \quad (2)$$

где W – система правил, заданных на множестве всех возможных состояний среды $W: \times_i Z_i \rightarrow \times_i Z_i$; $Y(t), Y(t+1)$ – состояния среды в последовательные моменты времени.

Отметим, что отображение W определяет закономерности среды функционирования, на основе которых агенты принимают решения о своих действиях по достижению цели.

Каждый агент характеризуется следующим кортежем:

$$\langle g_q, r_q, \mu_q(Y_q, g_q) \rangle,$$

где

- 1) $g_q = (z_{1j}^g, \dots, z_{nb}^g)$ – вектор целевых значений агента q , где $g_q \in SF$;
- 2) $r_q = (z_{1j}^r, \dots, z_{nb}^r)$ – стратегия достижения цели агента q , где $r_q \in U_q$,
 $U_q = \times_i Z_i^r, Z_i^r \subseteq Z_i$ – ресурсы агента q .

Считается, что агент q применяет стратегию r_q для достижения своей цели g_q , предполагая, что другие агенты никаких действий не совершали. Прогноз изменения состояния среды на n шагах определится из соотношения (1) при условии, что $Y_q(0) = r_q$, т.е. $Y_q(1) = W^{\circ} r_q, Y_q(2) = W^{\circ} Y_q(1), \dots, Y_q(n) = W^{\circ} Y_q(n-1)$.

- 3) $\mu_q(Y_q(n), g_q)$ – возможность достижения агентом q целевого состояния за счет собственных ресурсов в условиях противодействия других агентов.

При определении возможности достижения целевого состояния считается, что в пространстве состояний ($\times_i Z_i$) определена метрика $\rho(a, b), a, b \in \times_i Z_i$. Тогда возможность достижения цели агентом определяется как близость прогнозной $Y_q(n)$ и его целевой ситуации g_q :

$$\mu_q(Y_q(n), g_q) = \rho(Y_q(n), g_q).$$

По сути, этот показатель определяет потенциальную «силу» каждого агента команды без поддержки других членов команды при условии противодействия противников [Кулинич, 2014].

В этой модели считается, что состояние среды функционирования $Y(t)$, цель каждого агента g_q , его ресурсы r_q и возможность достижения цели самостоятельно $\mu_q(Y_q(n), g_q)$ известны всем агентам.

В такой общей постановке задачи необходимым условием командной работы агентов будем считать критерий близости целей агентов. В команду K в стайном алгоритме включаются агенты a_i , цели которых близки, т.е. $\forall a_i \in K, K \subseteq A, \rho(g_i, g_q) \leq \varepsilon, \forall i, q \in K, \varepsilon$ – критерий близости целей. При этом остальные агенты $A \setminus K$ считаются противниками.

Этот критерий, также как и аналогичный критерий в классической постановке задает необходимые условия командной работы.

Задача заключается в том, чтобы определить стратегии (действия) r_i всех агентов - a_i , включенных в команду $K, \forall a_i \in K$, которые позволяют им наилучшим образом достичь общей цели g .

В этой постановке задачи мы пока четко не определили критерий, по которому будет определена лучшая совместная стратегия достижения цели множеством агентов. Очевидно, что составляющими этого критерия являются время, общие затраты ресурсов для достижения цели и др.

Для решения этой задачи предлагается ряд алгоритмов поведения агентов в команде. Различные алгоритмы стайной работы агентов будем связывать с количеством и качеством общедоступной информации (уровнем информированности агентов), которой владеют агенты команды.

Вначале рассмотрим классический стайный алгоритм.

2.1. Классический стайный алгоритм.

Классический стайный алгоритм в общей постановке описан в терминах дифференциальных уравнений в работе [Каляев и др. 2008]. Далее не изменяя смысла постановки, будем рассуждать о поведении каждого агента терминах конечно-разностных уравнений (3), опираясь на ранее введенные определения.

$$\begin{aligned} Y_1(t+1) &= W(Y(t) + r_{1q}) \\ &\dots\dots\dots \\ Y_n(t+1) &= W(Y(t) + r_{nq}) \end{aligned} \quad (3)$$

Каждый агент на основе знаний о закономерностях среды функционирования W и о состоянии $Y(t)$ принимает решение о действии r_{iq} из множества доступных ему действий, т.е. $r_{iq} \in U_{iq}, U_{iq} = \times_i Z_i^r, Z_i^r \subseteq Z_i$.

Действие должно приблизить прогнозируемое им состояния среды $Y_i(t+1)$ к общей цели $g = (z_{1j}, \dots, z_{nb})$, т.е. критерий выбора действия каждого агента следующий: $\forall a_i \in K, \rho(Y_i(t+1), g) < \rho(Y_i(t), g)$.

Особенность стайной работы агентов заключается в том, прогнозируемое состояние среды функционирования каждого агента ($Y_i(t+1)$) отличается от состояния $Y^*(t+1)$, полученного в результате совместных действий агентов. Это означает, что на следующем шаге

работы алгоритма в систему уравнений (3) вместо состояния среды $Y(t)$ подставляется $Y^*(t+1)$. И так далее на всех n шагах реализации алгоритма.

Работа алгоритма останавливается, если на очередном шаге n , совместными усилиями агентов цель будет достигнута, т.е. $\rho(Y^*(t+n), g) < \varepsilon$.

В приведенном алгоритме все агенты действуют одновременно, не согласовывая свои действия, что приводит к неоптимальному результату.

2.2. Итерационный стайный алгоритм

В работе [Каляев и др., 2008] приводится итерационный алгоритм согласования действий всех роботов (агентов), который предполагает наличие между ними канала обмена информацией. Согласно этому алгоритму агенты упорядочены, например, по номерам и выбирают свои действия последовательно в порядке возрастания номеров. Причем первый, выбрав действие, сообщает о нем второму, который выбирает свое действие, учитывая действие первого, и сообщает об этом третьему агенту и т.д. Такая последовательная передача информации о действиях от роботов с меньшими номерами всем последующим роботам позволяет согласовать их действия некоторым оптимальным образом.

В этом алгоритме, до начала решения некоторой задачи группой роботов на основе последовательного обмена информацией о своих действиях формируется общий план последовательных действий агентов достижения общей цели. Далее, считается, что реализация этого плана позволяет оптимальным образом решить поставленную задачу.

Однако, необходимость скоординированного обмена информацией между роботами не позволяет говорить, что такой алгоритм относится к стайным алгоритмам, в которых агенты выбирают действия на основе анализа состояния среды функционирования и по косвенным признакам.

Такой алгоритм не применим, если среда функционирования динамичная, есть агенты–противники, препятствующие достижению цели команды агентов, и команда агентов будет вынуждена, при каждой смене состояния среды функционирования строить новый план оптимального взаимодействия агентов. Построение общего плана может потребовать много времени, что в динамических ситуациях неприемлемо.

Далее будут рассмотрены стайные алгоритмы, в которых применяется широкоэвентальный канал обмена информацией (доска объявлений), не предполагающий скоординированного диалога агентов.

3. Стайный алгоритм на основе критериев возможности достижения цели и взаимной полезности агентов

Этот алгоритм основан на анализе возможности достижения цели каждым агентом самостоятельно и их взаимной полезности. Вначале

формально определим взаимную полезность агентов [Кулинич, 2014].

Агенты q и i называются взаимно полезными, если объединение их стратегий $r_i \oplus r_q$ увеличивает возможность достижения собственных целей (g_q, g_i) агентами q и i в условиях противодействия их противников. То есть если $\mu(Y_{i+q}, g_i) < \mu(Y_i, g_i)$, $\mu(Y_{i+q}, g_q) < \mu(Y_q, g_q)$. Степень полезности $P(i, q)$ агента q для агента i определяется из соотношения [Кулинич, 2014]:

$$P(i, q) = 1 - \frac{\mu(Y_{i+q}, g_i)}{\mu(Y_i, g_i)} \quad (4)$$

где $\mu(Y_{i+q}, g_i)$, – возможности достижения агентами q и i своих целей при объединении их стратегий; $\mu(Y_q, g_q)$, $\mu(Y_i, g_i)$ – возможности достижения агентами q и i своих целей собственными силами в условиях противодействия противников.

Степень полезности следующего агента, для команды из двух агентов q и i будем определять из соотношения:

$$P((i, q), j) = 1 - \frac{\mu(Y_{((i+q)+j)}, g_i)}{\mu(Y_{((i+q))}, g_i)} \quad (5)$$

Таким образом, для определения полезности вновь добавляемого агента к существующей команде, команда считается единым агентом.

Считается, что каждый агент или команда на очередном такте работы помещает на доску объявлений информацию: о действиях r_{iq} (стратегию достижения цели) и о возможности достижения цели самостоятельно - $\mu_q(Y_q(n), g_q)$. Эти сведения доступны всем агентам.

Эту информацию каждый агент анализирует автономно по одинаковым алгоритмам. Поэтому результаты вычислений у всех агентов будут одинаковыми, что исключает необходимость их согласования.

Алгоритм функционирования агентов следующий.

Каждый агент упорядочивает всех агентов (или команду агентов) в порядке возрастания их возможностей достижения цели.

1. Если в перечне агентов есть агент (команда), способные достичь цели самостоятельно, то агент (команда) и будут решать задачу достижения цели, т.е. $\exists i | \mu(Y_i, g_i) < \varepsilon$. Остальные агенты в этом процессе не участвуют.
2. Если такого агента нет, то в качестве члена команды выбирается агент с большей возможностью достижения цели, для которого по формулам (4,5) подбирается агент с наибольшей взаимной полезностью. Эта команда будет работать на следующем такте работы стайного алгоритма. Остальные члены команды бездействуют.

Особенность стайного алгоритма заключается в том в том, что на каждом шаге работы алгоритма изменяются возможности достижения цели и, соответственно, взаимные полезности агентов.

3. Шаги алгоритма 1,2 и 3 повторяются до достижения цели, т.е. $\rho(Y^*(t+n), g) < \varepsilon$, для новых значений возможности достижения цели агента и их взаимной полезности.

Таким образом, на каждом шаге этого алгоритма состав команды может меняться по количеству агентов или составу. Но важно, то, что на каждом шаге команда всегда будет содержать агента (команду) с наибольшей возможностью достижения цели и самого взаимно полезного агента. Считается, что такой подбор команды позволит повысить качество командной работы агентов и приблизить их поведение к оптимальному.

Далее рассмотрим стайные алгоритмы, в которых алгоритмы командной работы стаи искусственных агентов (роботов) строятся на основе моделирования поведения людей в социуме.

3.1 Стайный алгоритм «ленивых» агентов

Этот алгоритм будем формулировать в терминах ранее введенной постановки задачи. Общее описание алгоритма следующее.

4. Каждый агент упорядочивает всех потенциальных членов команды по возможности достижения ими цели.
5. Если есть агент, способный достичь цели самостоятельно, то он достигает ее. Остальные агенты бездействуют.
6. Если такого агента нет, (здесь начинаются отличия от ранее описанного алгоритма), то команда принимает решение, чтобы лучший агент попытался самостоятельно достичь целевого состояния.
7. Остальные агенты, наблюдая за состоянием среды функционирования, ждут, когда этот агент исчерпает свои возможности достижения цели. Т.е. множество агентов находятся в состоянии «ленивого» ожидания результатов работы лучшего агента.
8. Если, наблюдаемый агент исчерпал свои возможности, то из множества «ленивых» агентов выбирается агент с большей возможностью достижения цели, который пытается ее достичь.
9. Множество «ленивых» агентов ждет, когда и этот агент исчерпает свои возможности и назначает нового исполнителя.
10. Процесс заканчивается, когда цель будет достигнута, т.е. если $\rho(Y^*(t+n), g) < \varepsilon$.

Этот алгоритм похож на итерационный алгоритм [Каляев и др., 2009], но его отличие заключается в том, что выбор очередного агента для

включения его в команду происходит в процессе решения агентами задачи, а не до его начала.

3.2. Стайный алгоритм «эгоистичных» агентов

Этот алгоритм предполагает, что устойчивая команда многократно выполняет некоторую работу. Например, в социальных системах, в бригаде рабочих при многократном выполнении проектов и т.д. между членами коллектива могут сложиться отношения (симпатии, неприязни, лидерства и т.д.), которые могут сказаться на работе коллектива. Отношения между агентами, возникающие в коллективе связаны с их индивидуальными отличиями и благами, которые могут получать агенты с большими возможностями. Например, в ранее рассмотренных алгоритмах, основанных на возможности достижения цели и взаимной полезности, основную командную работу будут выполнять только агенты с лучшими возможностями. Остальные будут «отдыхать». Если лучшие агенты за свою работу будут получать большие блага, то остальные («ленивые») агенты могут стремиться получать эти блага, перехватывая работу у лучших агентов, даже в том случае если качество работы всей команды ухудшится. Таких агентов будем называть эгоистичными.

В алгоритме «эгоистичных» агентов каждый агент кроме характеристики возможности достижения цели самостоятельно имеет еще и характеристику его работы O_i в предыдущих циклах работы, величина которой пропорциональна объему выполненной им работы и, соответственно, благам, которые агент за эту работу получил. Уровень эгоистичности агентов может быть определен из соотношения:

$$D_{ij} = \frac{|O_i - O_j|}{O_i + O_j}$$

В этом соотношении эгоистичность агентов $D_{ij}=0$, если они выполнили равную по объему работу, и соответственно, равна 1, если один из них никакой работы не выполнял – «отдыхал» («ленился»).

В алгоритме «эгоистичных» агентов считается, что сведения о выполненной работе каждого агента является общедоступными и, следовательно, агенты могут быть упорядочены по этой характеристике.

Рассмотрим алгоритм «эгоистичных» агентов.

1. Каждый агент упорядочивает всех потенциальных членов команды в два упорядоченных множества: по возможности достижения ими цели и по уровню эгоистичности агента относительно агента с максимальным объемом, выполненной работы.
2. Если есть агент, способный достичь цели самостоятельно, и его эгоистичность максимальна (или эгоистичности всех агентов равны), то он достигает ее. Остальные агенты бездействуют.

3. Если такого агента нет, то команда принимает решение, чтобы агент с максимальным уровнем эгоистичности попытался самостоятельно достичь целевого состояния.
4. Процесс заканчивается, когда цель будет достигнута, т.е. если $\rho(Y^*(t+n), g) < \varepsilon$.

Этот алгоритм позволяет равномерно распределить работу между всеми членами команды и обеспечить им одинаковые вознаграждения за работу в команде.

Заключение

В работе проанализирован классический стайный алгоритм роботов, выявлены его недостатки при формировании общего плана решения задачи в динамических ситуациях. Предложены стайные алгоритмы на основе критериев возможности самостоятельного достижения агентом цели и взаимной полезности агентов. На основе этих же критериев рассмотрены стайные алгоритмы «ленивых» и «эгоистичных» агентов.

Дальнейшие имитационные исследования формальных стайных алгоритмов позволят предложить их для реализации в командах роботов.

Список литературы

- [**Каляев и др., 2009**] Каляев И.А., Гайдук А.Р., Капустян С.Г. Модели и алгоритмы коллективного управления в группах роботов / М.: ФИЗМАТЛИТ, 2009.-279 с.
- [**Павловский, 2015**] Павловский В.Е., Павловский В.В. Математическая модель двумерной гомогенной стаи роботов// Искусственный интеллект и принятие решений. М.: URSS, 2015, №4, с. 62-71.
- [**Карпов, 2016**] Карпов В. Э. Модели социального поведения в групповой робототехнике / Управление большими системами. Выпуск 59. М.: ИПУ РАН, 2016. С.165-232.
- [**Кулинич, 2014**] Кулинич А. А. Модель командного поведения агентов (роботов): когнитивный подход / Управление большими системами. Выпуск 51. М.: ИПУ РАН, 2014. С.174-196.