

УДК 004.896

## СОЦИАЛЬНЫЕ МОДЕЛИ КОМАНДНОГО ПОВЕДЕНИЯ РЕАКТИВНЫХ РОБОТОВ

А.А. Кулинич (kulinich@ipu.ru)  
Институт проблем управления РАН, Москва

**Аннотация:** Исследуются вопросы формирования и функционирования команд роботов с реактивной архитектурой. Предложены принципы поведения каждого робота команды, основанные на анализе ими состояния среды функционирования роботов, оценки состояния среды с помощью эвристик и не предполагают информационного обмена между роботами.<sup>1</sup>

**Ключевые слова:** команда роботов, реактивный робот, принципы поведения, социальная модель.

### Введение

В теории группового управления интерес представляют роевые подходы к управлению группой роботов. Роевые алгоритмы, основанные на локальном взаимодействии множества однородных роботов (роя роботов), обеспечивают их скоординированное движение, обход препятствий и основывается на принципах, формализующих движение стаи птиц [Reynolds, 1987]. В научной литературе много публикаций о реализации роевых алгоритмов для решения разнообразных задач. Так в работе [Бурдун и др., 2010] приводится 40 алгоритмов, реализующих самое разнообразное движение групп роботов.

Кроме задач движения группы роботов строим, в групповой робототехнике существует ряд типовых задач – задачи патрулирования территории, фуражировки, картирование и др. [Карпов и др., 2019]. Для решения таких задач принципов движения роботов строим Рейнольдса оказывается недостаточным. Поиск принципов поведения группы роботов, которые обеспечивали бы решение конкретной общей задачи, в общем, не тривиальная задача. Обычно в качестве таких принципов выбираются принципы поведения насекомых, животных или людей. Однако, прямой перенос принципов поведения насекомых или людей на поведение искусственных роботов связан с ограничениями технологического

---

<sup>1</sup> Работа выполнена при частичной поддержке грантов: проект РФФИ ОФИ\_м № 16-29-04412.

характера. В работе [Кулинич, 2016] были рассмотрены ряд алгоритмов стайного и командного поведения роботов, которые предполагали обмен информацией между роботами о собственных целях и ресурсах. В этой работе сохраняется преемственность, результатов полученных в работе [Кулинич, 2016], однако акцент сделан на разработку принципов поведения так называемых стихийных команд роботов, не предполагающих обмен информацией. Поведение роботов сформулировано в форме принципов, основанных на анализе состояния среды функционирования и собственных целей и ресурсов роботов.

## 1 Среда функционирования формальных роботов

Рассматривается множество роботов  $A = \{R_i\}$ , обладающих свойствами (параметрами)  $F = \{f_i\}$ . Для каждого свойства каждого робота определено упорядоченное множество их возможных значений,  $Z = \{Z_i\}$ , где  $Z_i = \{z_{i1}, \dots, z_{iq}\}$ ,  $z_{iq+1} \succ z_{iq}$ ,  $q = 0 \dots n-1$ . Кроме этого, определено множество объектов  $B = \{b_j\}$ , каждый из которых имеет свойства из множества  $Z$  свойств роботов. Среда функционирования роботов определяется как прямое произведение множеств значений всех свойств роботов,  $SF = \times_i Z_i$ .

Вектор значений всех свойств роботов и свойств всех объектов  $Y(t) = (Y_1(t), \dots, Y_n(t), \dots, Y_{b1}(t), \dots, Y_{bm}(t))$  определяет состояние среды функционирования, где  $Y_1(t), Y_n(t)$  – свойства  $i$ -о робота;  $Y_{b1}(t), Y_{bm}(t)$  – свойства объектов;  $Y_i(t) = (z_{i1}, \dots, z_{inb})$ ,  $z_{ij} \in Z_i, \forall i$ .

Динамика изменения состояния среды функционирования происходит в случаях изменения роботами своих свойств или свойств объектов и представляется как отображение:

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

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

Каждый робот характеризуется следующим кортежем:

$$\langle g_q, r_q, \mu_q(Y_q, g_q), O(r_q) \rangle, \quad (3)$$

где

1)  $g_q = (z_{1j}^g, \dots, z_{nb}^g)$  – вектор целевых значений робота  $q$ , где  $g_q \in SF$ ;

2)  $r_q = (z_{1j}^r, \dots, z_{nb}^r)$  – стратегия достижения цели робота  $q$ , где  $r_q \in U_q$ ,

$U_q = \times_i Z_i^r, Z_i^r \subseteq Z_i$  – ресурсы робота  $q$ .

Считается, что робот  $q$  применяет стратегию  $r_q$  для достижения своей цели  $g_q$ , предполагая, что другие роботы никаких действий не совершали.

3)  $\mu_q(Y_q(n), g_q)$  – возможность достижения роботом  $q$  целевого состояния  $g_q$  за счет собственных ресурсов. Пусть в пространстве состояний ( $\times Z_i$ ) определена метрика  $\rho(a, b)$ ,  $a, b \in \times Z_i$ . Тогда возможность достижения цели роботом определяется как близость прогнозной  $Y_q(n)$  и его целевой ситуации  $g_q$ :

$$\mu_q(Y_q(n), g_q) = \rho(Y_q(n), g_q)^{-1}.$$

По сути, этот показатель определяет потенциальную «силу» каждого робота команды без поддержки других членов команды.

4)  $O(r_q)$  – полезность целевой ситуации для робота  $q$ . Под полезностью целевой ситуации для робота здесь понимается параметр, характеризующий поведение робота, который используется в процессах формирования команд роботов.

**Задача заключается** в переводе состояния среды функционирования из текущего положения  $Y(t)$  в целевое состояние  $Y^*(n)$  путем изменения собственных свойств роботов  $a_i \in A$  и свойств объектов  $b_i \in B$ . Иначе говоря, задача заключается в изменении свойств объектов или роботов, например, это может быть перемещение объекта или робота на плоскости в некоторую заданную точку. Считается, что зная состояние среды функционирования  $Y(t)$ , робот, используя ресурсы  $r_q$ , пытается достичь своей цели  $g_q$ , осознавая при этом возможность достижения цели самостоятельно  $\mu_q(Y_q(n), g_q)$  и полезность ее достижения  $O(r_q)$ .

Отметим, что в литературе рассматриваются два способа организации совместной работы роботов. Это централизованный метод, когда их совместная работа планируется заранее, и каждый робот получает задание на выполнение своей части общей работы, и децентрализованный, когда роботы на основе принципов самоорганизации самостоятельно решают общую задачу. Далее будет рассмотрен децентрализованный метод взаимодействия роботов, основанный на принципах социальной организации в малых социальных группах.

## 2 Принципы формирования и функционирования стихийных команд реактивных роботов

Минимальное определение интеллектуальных роботов – это перечисление их свойств: автономность – способность действовать самостоятельно; реактивность – способность реагировать на изменение среды; проактивность – способность проявлять активность согласно своим целям; социальность – способность общаться с другими роботами.

Далее мы будем рассматривать реактивных роботов, у которых присутствуют две способности – автономность и реактивность.

Под принципом формирования и функционирования команды роботов мы будем понимать поведение каждого робота, основанное на анализе среды функционирования и имеющейся у робота информации о собственном состоянии.

Основываясь на работах социальных психологов, исследовавших вопросы самоорганизации и сплоченности в малых социальных группах мы будем рассматривать следующие принципы поведения роботов при выполнении совместной работы:

- принцип самостоятельного достижения цели, который означает, что робот пытается самостоятельно достичь цели, независимо от возможности ее достижения;
- принцип взаимной полезности – означает, что робот не способный достичь цели взаимодействует с роботом, который имеет лучшие возможности достижения цели и которому он сам будет полезен;
- принцип «ленивости» - означает, что робот не присоединяется к работе в группе роботов, если эта группа достигает цели этого робота;
- принцип «эгоистичности» - работает в случаях, когда роботы получают «вознаграждение» за достижение цели. Эгоистичные роботы пытаются достичь цели и получить вознаграждение, даже в случаях, если их возможности достижения цели невелики.

### 3 Стихийные команды реактивных роботов

Под стихийной командой реактивных роботов будем понимать множество роботов имеющих общую цель, возможность наблюдать за состоянием среды функционирования и не имеющие возможности общаться и координировать свои действия - строить общий план.

Итак, есть группа реактивных роботов  $R_i$ , каждый из которых определен кортежем (3). Считаем, что роботы не могут общаться, но они знают состояние среды функционирования  $Y(t)$ , и могут изменить это состояния, используя для этого собственные ресурсы.

#### 3.1 Принцип самостоятельного достижения цели

Рассмотрим формирование и функционирование группы роботов на основе принципа самостоятельного достижения цели.

$$Y^*(t+1) = W^o Y(t) \oplus \left( \bigoplus_i r_i \right)$$

В этом случае роботы не знают о существовании других роботов их целей и возможностей их достижения. Все роботы применяют одновременно собственную стратегию  $r_i$  с целью достичь собственного целевого состояния  $g_i$ . Поскольку цели у роботов могут быть разными то,

агрегация стратегий  $\bigoplus_i r_i$  не гарантирует, что каждый робот достигнет цели.

Однако, здесь возможны частные случаи.

1. В группе роботов существует робот, стратегия которого доминирует агрегированные стратегии всех остальных роботов. Такого робота будем называть диктатором. Если в группе роботов есть робот-диктатор, то он достигает своей цели, а остальные роботы нет.
2. В группе роботов существуют роботы, имеющие близкие цели,  $\forall R_i \in K, K \subseteq A, \rho(g_i, g_q) \leq \varepsilon, \forall R_i, R_q \in K, \varepsilon$  – критерий близости целей. В этом случае образуется команда роботов, и в случае если их агрегированная стратегия доминирует стратегии всех остальных роботов, то команда играет роль робота-диктатора.

### 3.2 Принцип взаимной полезности

Необходимым условием работы команды на этом принципе являются наличие в группе роботов  $A$ , роботов с близкими целями. Эти роботы образуют команду, и могут объединять собственные ресурсы. Считаем, что есть группа роботов  $A$  (3); есть подмножество  $K$  роботов с близкими целями  $\forall R_i \in K, K \subseteq A, \rho(g_i, g_q) \leq \varepsilon, \forall R_i, R_q \in K$ .

Вначале сформулируем формально принцип полезности роботов.

Роботов  $R_i, R_q$  с близкими целями будем называть взаимно полезными, если объединение их ресурсов увеличивает возможность достижения общей цели  $g_\Sigma$ . В случае объединения ресурсов роботов коалиции прогноз развития ситуации определится из выражения (4):

$$Y_{iq}^*(n) = W^\circ Y(t) \bigoplus_{R_i \in K} r_i \bigoplus_{R_j \notin K} r_j, \quad (4)$$

$$\rho(g_\Sigma, Y_{iq}^*(n)) < \rho(g_i, Y_i^*(n)) \quad (5)$$

$$\rho(g_\Sigma, Y_{iq}^*(n)) < \rho(g_q, Y_q^*(n)) \quad (6)$$

Выражения (5) и (6) означают, что возможность достижения цели роботы от объединения ресурсов больше, чем, если бы они действовали самостоятельно. Однако отметим, что получить реальные прогнозы применения своих стратегий роботы могут только в условиях, когда

роботы, не принадлежащие коалиции, бездействуют, т.е.  $\bigoplus_{R_j \notin K} r_j = 0$  или

агрегированные ресурсы коалиции доминируют ресурсы всех остальных

роботов, т.е.  $\bigoplus_{R_i \in K} r_i \gg \bigoplus_{R_j \notin K} r_j$ . Если роботы–противники обладают

достаточным ресурсом и противодействуют роботам команды, то

формальная модель (4)(5)(6) не гарантирует достижения командой общей цели. В этих условиях используем эвристику: агент, имеющий большие возможности достижения цели, имеет большую привлекательность для командной работы. В нашей постановке задачи можно говорить, что возможность достижения цели у робота тем выше, чем ближе он находится к целевому объекту, т.е. к объекту или роботу, параметры которого нужно изменить до целевых значений. Можно задать окрестность близости роботов коалиции к целевому объекту  $\varepsilon_k$ , в которой роботы считаются привлекательными, т.е.  $\rho(Y_i(t), Y_q(t)) < \varepsilon_k$ .

Тогда правило поведения роботов при функционировании команды, заключается в выборе и объединении ресурсов лучших роботов. Формально правило поведения каждого робота команды запишем так:

**Если**  $\rho(Y_i(t), Y_q(t)) < \varepsilon_k, \forall R_q \in K$ .

**То**  $Y_i^*(n) = W^\circ Y(t) \oplus_{R_j \in K}^N r_j$

В результате независимых действий каждого робота будет стихийно сформирована агрегированная стратегия лучших роботов команды, т.е.

$$Y_K^*(n) = W^\circ Y(t) \oplus_{R_i \in K} r_i \oplus_{R_j \in K} r_j,$$

где стратегия команды  $\oplus_{R_i \in K} r_i$  включает стратегии роботов, имеющих

большие возможности достижения цели из-за близости целевого объекта, т.е.  $\rho(Y_i(t), Y_q(t)) < \varepsilon_k$ .

### 3.3 Принцип «ленности» робота

При командной работе на основе принципа взаимной полезности, в условиях, когда неизвестны ресурсы групп роботов, не входящих в коалицию  $K$ , объединение ресурсов лучших роботов может быть избыточной. Рассмотрим принцип «ленности» робота. Правило поведения каждого робота в этом случае будет выглядеть так:

*Условие 1. Если*  $\rho(g_i, Y_i(t)) > \rho(g_i, Y_i(t+1))$

$$r_i^* = 0, Y_i^*(n) = W^\circ Y(t) \oplus_{R_j \in K}^N r_j$$

*То переход на проверку условия 1*

**Иначе**

*Условие 2. Если*  $\rho(Y_i(t), Y_q(t)) < \varepsilon_k, \forall R_q \in K$ .

$$\text{То } Y_i^*(n) = W^{\circ} Y(t) \oplus r_i \bigoplus_{R_j \in K}^N r_j$$

*переход на проверку условия 1*

При выполнении действий, основанных на принципе «ленивости» робота, все роботы контролируют состояние среды (Условие 1). Если состояние среды меняется и целевой объект приближается к их цели, то действия отсутствуют. В противном случае роботы пытаются изменить состояние объекта в направлении цели. При этом свойства объекта изменяет команда из роботов, ресурсов, которых достаточно для изменения свойств объекта. Сам процесс образования такой «ленивой» команды происходит на основе только контроля состояния среды и не предполагает обмен информацией между роботами команды.

В командах роботов на основе полезности и «ленивости» работу по достижению цели будут выполнять лучшие роботы, в то время как роботы, имеющие меньшие возможности достижения цели будут «простаивать». Этот недостаток команды на основе принципа «ленивых» роботов можно устранить, если к принципам взаимной полезности и «ленивости» добавить еще принцип «эгоистичности» роботов.

### 3.4 Принцип «эгоистичности» робота

При разработке принципа «эгоистичности» робота рассматривается эвристика, основанная на том, что за совместную работу по достижению цели роботы получают вознаграждение. Вознаграждение это параметр, позволяющий регулировать функционирование команды. Считается, что все роботы команды, независимо от их возможностей по достижению цели стремятся получить вознаграждение. В командах, построенных на принципах полезности и «ленивости» вознаграждения будут получать только лучшие роботы. Поскольку роботы не способны обмениваться сообщениями, а могут только анализировать состояние среды функционирования, будем считать, что лучшие роботы, которых мы определили как роботов, находящихся ближе к целевому объекту будут получать большие вознаграждения за работу и, следовательно, иметь большую полезность. Т.е.  $\rho(g, Y_i(t)) > \rho(g, Y_j(t)) \Rightarrow O(r_i) > O(r_j)$ .

Этот принцип будем применять совместно с принципом «ленивости» роботов модифицировав условие 2:

*Условие 2. Если  $\rho(Y_i(t), Y_q(t)) > \varepsilon_k, \forall R_q \in K$ .*

В этом случае в работе команды роботов будут принимать участие роботы, имеющие не лучшие возможности достижения цели, но они будут получать вознаграждения в случаях, если достичь цели все-таки удастся.

Второй вариант модификации принципа «ленивых» роботов заключается в исключении условия 2. Т.е.:

**Иначе**

$$Y_i^*(n) = W \circ Y(t) \oplus r_i \bigoplus_{R_j \in K}^N r_j$$

*переход на проверку условия 1*

В этом случае, роботы вне зависимости от их возможности достижения цели будут подключаться к выполнению общей работы на основе принципа «ленивости» робота, т.е. в случае выполнения условия 1. Такой алгоритм дает возможность роботу с любой возможностью достижения цели, достичь ее в команде и получить вознаграждение.

## **4 Требования к полноте информации о состоянии среды функционирования**

Рассмотренные принципы формирования и функционирования стихийных команд роботов не предполагают, что каждый робот обладает полной информацией о состоянии среды функционирования  $Y(t)$ . Для выполнения своей задачи достижения цели роботу достаточно знать текущие  $Y(t)$  и целевые  $Y^*(t)$  значения целевых объектов или роботов. Будем считать, что неопределенность возможна как в целевых значениях параметров робота, так в текущих состояниях целевых объектов.

Пусть параметры робота или объекта  $A$ ,  $Y_A(t) = (z_{1j}, \dots, z_{nb})$ ,  $z_{ij} \in Z_{Ai}$ . Неопределенность в векторе состояния будем обозначать знаком вопроса «?». Например, в векторе параметров  $Y_A(t) = (z_{1j}, ?, z_{3j}, ?, \dots, z_{nb})$  неопределенны второй и четвертый параметры, т.е.  $z_{2A}$  и  $z_{4A}$ . Поведение робота зависит от неопределенности в параметрах целевого объекта и в целевых параметрах и от характера задачи, решаемой роботом.

Для задачи фуражировки условия командной работы следующие: 1) для изменения свойств целевых  $Y_A(t)$  объекта  $A$ , робот  $B$  должен изменить свои свойства  $Y_B(t)$  так, чтобы взаимодействие было возможно; 2) робот  $B$  может изменить свойства объекта  $A$  до целевых  $Y^*(t)$  при условии достаточности ресурсов. В задаче фуражировки эти условия означают, что робот  $B$  должен подойти к целевому объекту  $A$  и только после этого изменить его параметры самостоятельно или в команде.

Рассмотрим поведение робота в зависимости от различного сочетания неопределенностей:

- Неопределенностей нет. Известны текущие параметры целевого объекта и его целевое состояние.
- Неопределенность в параметрах целевого объекта.
- Неопределенность в параметрах целевого объекта и в параметрах цели.

Рассмотрим предложенные принципы и правила поведения в условиях неопределенности на простых примерах.

## 5 Пример и эксперименты

Для исследования формирования и функционирования стихийных команд роботов рассмотрим задачу фуражировки стихийными командами роботов. Вектор состояния среды функционирования  $Y(t)=(Y_1(t), \dots, Y_n(t))$  – это вектор параметров робота или объекта –  $Y_i = (x_i, y_i, c_i, w_i)$ , где  $x_i, y_i$  – координаты на плоскости,  $c_i$  – цвет робота и  $w_i$  – его вес. Робот характеризуется четверкой: цель  $G_i$ , ресурсы  $R_i$ , возможность самостоятельного достижения цели  $\mu(Y(n), G_i)$  и полезностью достижения цели  $O(R_i)$ . Цель формулируется как вектор:  $G_i=(G_{i1}, \dots, G_{in})$ , где  $G_{ij}$ – вектор целевых параметров  $j$ -го целевого объекта, которые должен достичь  $i$ -й робот. Ресурсы  $R_i$  связаны с весом  $w_i$  робота. Возможность самостоятельного достижения цели  $\mu(Y(n), G_i)$  представляется как функция от параметров самого робота или других роботов или объектов. При рассмотрении принципов организации стихийных команд возможность достижения цели была представлена как эвристическая функция расстояния до целевого объекта, т.е.  $\rho(g, Y_i(t))$ . Аналогично, как эвристика была определена функция полезность достижения цели  $O(R_i)$ .

Далее будем считать, что роботы могут менять только координаты целевых объектов. Изменения веса  $w_i$  и цвета  $c_i$  недоступны. Условия

фуражировки стихийной командой роботов -  $\bigoplus_{R_i \in K} w_i > w_j$ , т.е. вес целевого объекта меньше агрегированного веса стихийной команды. Рассмотрим возможные варианты поведения роботов, при разных неопределенностях в параметрах целевого объекта и параметрах цели.

- Параметры целевого объекта  $(x_i, y_i, c_i, w_i)$  и цели  $(x_i^*, y_i^*, c_i^*, w_i^*)$  полностью определены. В этом случае может возникнуть стихийная команда роботов, работающая по любому из рассмотренных принципов и переместить целевой объект в заданную точку.
- Есть неопределенность целевого объекта в параметрах целевого объекта  $(x_i, ?, c_i, w_i)$ , а цель определена точно -  $(x_i^*, y_i^*, c_i^*, w_i^*)$ . В этом случае любая неопределенность в параметрах целевого объекта приводит к поиску роботами объекта для любых значений неопределенного параметра. В этом случае роботы, имеющие общие цели могут найти разные целевые объекты, для фуражировки которых у них может оказаться недостаточно ресурсов. В этом случае стихийная команда может не образоваться.

- Целевой объект определен точно  $(x_i, y_i, c_i, w_i)$ , но есть неопределенность в параметрах цели  $(x_i^*, ?, c_i^*, w_i^*)$ . В этом случае, неопределенность в цели позволяет роботам изменять значение неопределенного параметра до любых значений. Стихийная команда может образоваться.

В задаче фуражировки полезность каждого робота для командной работы определяется его близостью к целевому объекту и весом. Информация о параметрах других роботов недоступна. В работе [Кулинич, 2012] рассмотрена модель функционирования стихийных команд роботов, играющих в виртуальный футбол, в которой привлекательность робота для совместной работы оценивалась на основе анализа состояний всех роботов с помощью линейной свертки.

## Заключение

В работе рассмотрены принципы формирования и функционирования стихийных команд роботов, основанные на анализе каждым роботом состояния среды функционирования и не предполагающие обмен роботами информацией о своих целях и ресурсах. Рассмотрены варианты поведения роботов в условиях неопределенности в параметрах целевого объекта и цели. Проводятся анализ возможности образования стихийных команд роботов в условиях неопределенности.

## Список литературы

- [Reynolds, 1987] Reynolds, C.W. Flocks, Herds, and Schools: A Distributed Behavioural Model // Computer Graphics. – 1987. – Vol. 21, No. 4, – PP. 25–34.
- [Бурдун и др., 2010] Бурдун И.Е., Бубин А.Р. Метод самоорганизации стайного поведения малых мобильных роботов гражданского и специального назначения для арктических приложений. Сб. трудов Всероссийской научно-технической конференции «Научное и техническое освоение шельфа Северного Ледовитого океана». СибГУТИ. Новосибирск. 2010. С. 141-149.
- [Карпов и др., 2019] Карпов В.Э., Карпова И.П., Кулинич А.А. Социальные сообщества роботов. М.: URSS, ООО «ЛЕНАНД», 2019. – 352 с.
- [Кулинич, 2016] Кулинич А.А. Модели стайного поведения роботов / Всероссийский научно-практический семинар «Беспилотные транспортные средства с элементами искусственного интеллекта» (БТС-ИИ-2016) , 22-23 сентября 2016, Казань-Иннополис, с.60-69.
- [Кулинич, 2012] Кулинич А.А. Модель поддержки принятия решений для образования коалиций в условиях неопределенности // Искусственный интеллект и принятие решений. № 2, 2012, стр. 95-106.