

Е.Ю. ШМАЛЬКО, А.И. ДИВЕЕВ, К.Л. ЯМШАНОВ, И.А. ГРОМОВ
**СИМВОЛЬНОЕ МАШИННОЕ ОБУЧЕНИЕ
КВАЗИОПТИМАЛЬНОГО УПРАВЛЕНИЯ КОЛЛАБОРАТИВНОЙ
МНОГОАГЕНТНОЙ РОБОТОТЕХНИЧЕСКОЙ СИСТЕМОЙ**

Шмалько Е.Ю., Дивеев А.И., Ямшанов К.Л., Громов И.А. Символьное машинное обучение квазиоптимального управления коллаборативной многоагентной робототехнической системой.

Аннотация. В статье представлены новые результаты применения методов символьного машинного обучения для автоматизации разработки системы управления применительно к коллаборативной многоагентной системе. Предложена оригинальная математическая постановка задачи оптимального управления коллаборативной многоагентной системой, где коллаборативность подразумевает совместную деятельность роботов и людей в едином сложном фазовом пространстве для достижения цели управления с оптимальным значением заданного критерия качества. В представленной формулировке предлагается рассматривать роль человека в виде фазовых ограничений с неопределенностью и дополнительного вектора управления, с помощью которого уменьшается влияние этих фазовых ограничений в терминах минимизации функционала. Решением задачи является многомерная функция управления, которая переводит технический объект, группу мобильных роботов, в целевое состояние с учетом возможных неопределенностей в состоянии объекта, при этом минимизируя заданный критерий качества процесса. Задача состоит в автоматическом нахождении как оптимальной структуры, так и параметров функции управления. Такая общая постановка задачи управления не может быть решена напрямую существующими аналитическими методами из-за сложности и требует большого ручного труда по выбору соответствующих регуляторов и оптимальной настройке их параметров. В статье предлагается подход, основанный на символьном машинном обучении с помощью символьной регрессии и эволюционной оптимизации для решения такой обобщенно поставленной задачи. Эффективность подхода продемонстрирована численными экспериментами. Приведен пример задачи оптимального управления для трех мобильных роботов с двумя управляемыми фазовыми ограничениями. Прикладная задача рассматривается в общей математической постановке, что позволяет использовать универсальные методы символьного машинного обучения для автоматического поиска решения.

Ключевые слова: машинное обучение, оптимальное управление, фазовые ограничения, коллаборативная многоагентная робототехническая система, символьная регрессия.

1. Введение. В соответствии с современными мировыми тенденциями появляется много новых технологий, которые облегчают работу людей. В этой связи коллаборативное взаимодействие с роботами стало широко исследуемой темой [1–3], объединяющей опыт людей с рабочей нагрузкой роботов, повышающей эффективность и производительность в различных сферах народного хозяйства. Несмотря на относительную молодость данного технологического направления, коллаборативная робототехника прошла путь от нишевой концепции [4]

до одного из самых динамично растущих сегментов автоматизации, о чем свидетельствует официальная статистика Международной федерации робототехники (IRF, International Robotic Federation) [5]. Главная ценность коллаборативной робототехники в гибридности ее концепции, где сильные стороны человека (гибкость, креативность, принятие решений) сочетаются с преимуществами робота (точность, выносливость, повторяемость).

В более ранних работах по коллаборативной робототехнике [6, 7] действия человека рассматривались как действия супервайзера, управляющего роботами. В этом контексте человек не выполняет необходимую задачу совместно с роботом, а только управляет им. Такие системы не являются полностью коллаборативными.

Коллаборативные робототехнические системы подразумевают совместное решение целевой задачи человеком и роботизированной системой в едином рабочем пространстве [8–10]. Первоначально такое взаимодействие было введено в рамках промышленной коллаборативной ячейки [11, 12], где совместная работа выполняется человеком-оператором и роботом. И первостепенной задачей при разработке таких коллаборативных промышленных робототехнических ячеек было обеспечение безопасного взаимодействия человека и робота, для чего необходимо обеспечить роботу достаточно полное восприятие пространства, в котором находится коллаборативный робот. Для этого коллаборативные ячейки оснащаются большим набором датчиков разной модальности, охватывающих весь объем работ. Большой объем научных исследований по коллаборативной робототехнике [12–14] посвящен именно вопросу разработки сенсорно-информационной среды, достаточной для обеспечения безопасного взаимодействия человека и робота, поскольку это направление имеет первостепенное значение. На сегодняшний день эта задача решена на хорошем уровне, и поэтому все большее распространение в производственном секторе получают коллаборативные роботы, сокращенно называемые «коботами» [11], для повышения эффективности и производительности труда.

Сегодня в соответствии с общемировыми тенденциями в робототехнике актуальность приобретают новые задачи, такие как:

1. Переход на новый уровень коллаборативных робототехнических систем, а именно коллаборативных многоагентных робототехнических систем (КолМАРС) [9, 15, 16]. Несколько агентов способны работать над достижением общей цели, выходящей за рамки возможностей каждого из агентов в отдельности.

2. Интеграция искусственного интеллекта и машинного обучения трансформирует коллаборативных роботов из запрограммированных машин в адаптивных партнеров, способных к обучению, прогнозированию намерений и адаптивному планированию траекторий [17, 18].

3. Конвергенция автономных мобильных роботов и коллаборативных манипуляторов создает новый класс – мобильные коллаборативные системы [19], ставя перед разработчиками новые сложные задачи по автономной навигации в сложных средах, динамическому перераспределению задач, координации в гетерогенных группах [19, 20].

Представленные выше тенденции указывают на актуальность настоящей работы, которая посвящена вопросу исследования коллаборативных мультиагентных робототехнических систем с использованием современных методов машинного обучения. Создание нового класса КолМАРС требует решения ряда новых задач, связанных с групповым управлением и взаимодействием человека с группой роботов, а также группы людей и группы роботов, включая не только задачи технического уровня, но и задачи планирования верхнего уровня, связанные с рациональным разделением функций человека и роботов в соответствии с требуемым уровнем совместной работы [21], включением человека с его психоэмоциональными и физическими характеристиками [22] в качестве равноценного агента КолМАРС и т.д.

В настоящей работе исследуется вопрос формализации задач управления КолМАРС и способов решения поставленных задач. Рассматривается технический уровень разработки системы управления коллаборативной многоагентной робототехнической системой.

Сегодня передовые средства навигации и компьютерного зрения позволяют построить сенсорно-информационную систему так, чтобы мобильные роботы в рамках КолМАРС могли выполнять задачу, органично взаимодействуя с людьми. С точки зрения низкоуровневого управления основные подходы направлены на плавное выполнение целевых задач с учетом динамической среды физического взаимодействия человека и робота. Основная цель управления – каким-то образом отслеживать траектории движения человека и избегать столкновений с ним. В зависимости от типа коллаборативных роботов в литературе уже предложены различные стратегии управления, такие как управление механическим импедансом [23, 24], которая позволяет роботу взаимодействовать с окружающей средой, учитывая ее жесткость и сопротивление; стратегия управления на основе скорости, которая минимизирует расстояние между роботами и людьми в зависимости от скорости робота в рабочем пространстве [25],

управление с прогнозированием по модели и метод потенциальных полей для мобильной робототехники [26, 27], вероятностный подход [28], основанный на прогнозировании областей безопасности, занимаемых человеком и др.

Однако, как и при развитии любой технологии, при достижении хорошего уровня инженерной реализации встают вопросы, связанные с повышением эффективности, ускорением процесса разработки, универсализацией подходов и т.д.

С точки зрения задач управления встает вопрос о получении такого управления для коллаборативных роботов, которое было бы оптимальным по критерию качества. И в этой связи возникает необходимость математически сформулировать задачу управления коллаборативными роботами как задачу оптимального управления и формализовать эффективность применяемых законов управления на основе критерия качества, определяемого путем минимизации интегрального функционала. Формализация задачи не только позволяет ввести строгие критерии оценки эффективности работы системы управления, но также открывает широкие возможности применения современных методов искусственного интеллекта для автоматизированной разработки систем управления, предоставляя разработчику возможность перевести процесс синтеза закона управления на вычислительную машину.

В данной статье предлагается постановка основной задачи управления многоагентной коллаборативной системой как задачи оптимального управления. Обсуждается трудоемкость решения такой задачи строгими аналитическими подходами. Далее в работе предлагается и демонстрируется на примере новая интеллектуальная технология синтезированного оптимального управления на основе символического машинного обучения. Представлено математическое описание синтезированного оптимального управления, а также метод символической регрессии для реализации этапа синтеза системы стабилизации и алгоритм эволюционной оптимизации для поиска оптимального расположения управляющих точек равновесия. Предложенный подход в автономном режиме на основе представленной постановки задачи формирует нелинейный закон управления с точки зрения минимизации заданного критерия качества для обеспечения квазиоптимального движения в пространстве со сложными фазовыми ограничениями.

2. Постановка задачи оптимального управления коллаборативной многоагентной робототехнической системой.

В общем случае, коллаборативная система включает в себя несколько агентов, таких как роботы и люди. Постановки задач управления для

коллаборативных многоагентных систем отличаются от задач с одним роботом, поскольку они учитывают динамику группы роботов.

Математическая модель многоагентной робототехнической системы описывается системой обыкновенных дифференциальных уравнений:

$$\dot{\mathbf{x}}^j = \mathbf{f}^j(\mathbf{x}^j, \mathbf{u}^j), \quad (1)$$

где \mathbf{x}^j является вектором пространства состояний для робота j , $\mathbf{x}^j \in \mathbb{R}^{n_j}$, \mathbf{u}^j - вектор управления для робота j , $\mathbf{u}^j \in U_j \subseteq \mathbb{R}^{m_j}$, U_j - компакт, $m_j \leq n_j$, $j = 1, \dots, N$, N - количество роботов в многоагентной робототехнической системе. Предполагается, что функции правых частей могут и не иметь строгого аналитического описания, а описываться, например, нейронной сетью или некоторой сложной формулой, полученной символьной регрессией.

В задаче оптимального управления для каждого робота заданы начальные состояния:

$$\mathbf{x}^j(0) = \mathbf{x}^{0,j} \in \mathbb{R}^{n_j}, \quad j = 1, \dots, N. \quad (2)$$

Цель управления определена через попадание в заданные целевые терминальные состояния:

$$\mathbf{x}^j(t_{f,j}) = \mathbf{x}^{f,j} \in \mathbb{R}^{n_j}, \quad j = 1, \dots, N, \quad (3)$$

где $t_{f,j}$ - время достижения роботом j терминальных условий,

$$t_{f,j} = \begin{cases} t, & \text{если } t < t^+, \text{ и } \|\mathbf{x}^{f,j} - \mathbf{x}_j(t)\| \leq \varepsilon \\ t^+ & \text{иначе} \end{cases}, \quad (4)$$

t^+ и ε - заданные положительные числа.

Отличительной особенностью постановки задачи оптимального управления коллаборативной системой является необходимость учета влияния человека на систему в общей постановке. Объектом управления (1) в коллаборативной системе являются роботы, как технические объекты. Предположим, что фактор присутствия человека в системе учитывается как фазовые ограничения с элементом влияния на них для обеспечения движений без коллизий и выполнения поставленных целей.

С учетом сделанного предположения, фазовые ограничения могут быть определены для всех роботов следующим образом:

$$\varphi_i(\mathbf{x}^j, \mathbf{v}) \leq 0, \quad i = 1, \dots, S, \quad j = 1, \dots, N, \quad (5)$$

где \mathbf{v} - второй вектор управления, реализуемый человеком путем влияния на фазовые ограничения, $\mathbf{v} \in V \subseteq \mathbb{R}^r$.

В (5) предполагаем, что рассматриваемая многоагентная система имеет N роботов и одного человека, поэтому фазовые ограничения (5) включают только один вектор управления человеком. В дальнейшем это предположение может быть обобщено по мере необходимости.

Кроме рассмотренных фазовых ограничений, для многоагентной робототехнической системы характерно наличие динамических фазовых ограничений, которые возникают ввиду наличия в фазовом пространстве нескольких подвижных объектов в группе:

$$\chi(\mathbf{x}^j, \mathbf{x}^k) \leq 0, \quad j = 1, \dots, N - 1, \quad k = j + 1, \dots, N. \quad (6)$$

Динамические фазовые ограничения (6) описывают условия избегания столкновений между роботами в группе.

Критерий качества для коллаборативной многоагентной робототехнической системы в общем случае можно представить обобщенным интегральным критерием, включая все векторы пространства состояний и все векторы управления:

$$\begin{aligned} J(\mathbf{u}^1, \dots, \mathbf{u}^N, \mathbf{v}) &= \\ &= \int_0^{t_f} f_0(\mathbf{x}^1, \dots, \mathbf{x}^N, \mathbf{u}^1, \dots, \mathbf{u}^N, \mathbf{v}) dt \rightarrow \min_{\mathbf{u}^1 \in U_1, \dots, \mathbf{u}^N \in U_N, \mathbf{v} \in V}, \end{aligned} \quad (7)$$

где $t_f = \max\{t_{f,j} : j = 1, \dots, N\}$.

В задаче оптимального управления необходимо найти вектор управления роботами:

$$\mathbf{u}^j = \mathbf{w}^j(t) \in U_j, \quad j = 1, \dots, N, \quad (8)$$

а также вектор управления, реализуемый человеком, если иных предположений о нем не сделано в постановке задачи:

$$\mathbf{v} = \mathbf{z}(t) \in V, \quad (9)$$

такие что, система дифференциальных уравнений:

$$\dot{\mathbf{x}}^j = \mathbf{f}^j(\mathbf{x}^j, \mathbf{w}^j(t)), \quad j = 1, \dots, N,$$

имела бы частные решения $\mathbf{x}^j(t, \mathbf{x}^{0,j})$, которые бы из заданных начальных состояний (2) достигали терминальных состояний (3):

$$\|\mathbf{x}^j(t_{f,j}, \mathbf{x}^{0,j}) - \mathbf{x}^{f,j}\| \leq \varepsilon,$$

и при этом в любой момент времени все фазовые ограничения выполняются (5), (6):

$$\varphi_i(\mathbf{x}^j(t, \mathbf{x}^{0,j}), \mathbf{z}(t)) \leq 0, \quad i = 1, \dots, S, \quad j = 1, \dots, N,$$

$$\chi(\mathbf{x}^j(t, \mathbf{x}^{0,j}), \mathbf{x}^k(t, \mathbf{x}^{0,k})) \leq 0, \quad j = 1, \dots, N-1, \quad k = j+1, \dots, N,$$

и критерий качества (7) имеет минимальное значение.

3. Машинно синтезированное оптимальное управление.

Задачи оптимального управления с фазовыми ограничениями представляют особую сложность при применении строгих аналитических подходов [29, 30]. Существующие подходы применимы лишь для определённого класса объектов и функционалов, либо требуют введения дополнительных преобразований, как например введение множителей Лагранжа в принципе максимума [31]. Поставленные задачи оптимального управления коллаборативными робототехническими системами обладают повышенной сложностью ввиду наличия как динамических фазовых ограничений, так и элемента неопределенности в фазовых ограничениях, связанных с присутствием человека в области выполнения совместных действий.

В этой связи основной акцент в настоящем научном исследовании сделан на использовании современных методов машинного обучения для численного решения поставленной задачи. Исследования [32–34] дают основание применимости к поставленным задачам подхода на основе принципа синтезированного оптимального управления. Его ключевая идея состоит в том, чтобы не искать функцию управления

напрямую как функцию времени в виде решения исходной задачи оптимального управления, а находить функцию, параметры которой определяются положением устойчивой точки равновесия (аттрактора) для предварительно стабилизированного объекта. Задача сводится к оптимизации этих параметров, периодически пересчитывая их для целенаправленного перемещения аттрактора. Таким образом, управление осуществляется не напрямую, а через оптимальное изменение положения точки равновесия в дискретные моменты времени, что позволяет эффективно автоматизировать поиск управляющих функций методами машинного обучения. Данная параметризация управления открывает возможности для автоматизации процесса поиска управляющих функций с применением методов оптимизации и машинного обучения.

Согласно принципу, численное решение изначально сформулированной задачи оптимального управления реализуется в два этапа:

- сначала решается задача синтеза системы стабилизации в заданную точку пространства состояния;
- затем решается задача параметрической оптимизации, в которой необходимо найти значения параметров, определяющих положение точки устойчивого равновесия на заданных интервалах времени. Путем последовательного переключения между расчетными точками в пространстве состояний объект перемещается из начального состояния в конечное, обеспечивая при этом выполнение всех фазовых ограничений и оптимальность заданного критерия качества.

Такой подход обеспечивает получение реализуемых функций управления, обеспечивая малую чувствительность к внешним возмущениям. Малая чувствительность в случае управления движением объекта за счет последовательной активизации точек устойчивости основана на том, что возможные траектории движения объекта из разных начальных условий сжимаются в окрестности устойчивой точки равновесия, поэтому ошибки по положению объекта уменьшаются, что повышает робастность системы.

В результате такой численный подход позволяет решить задачу без дополнительных аналитических преобразований, функциональных упрощений, линеаризации модели объекта и т.д. Использование машинных методов, хотя и не гарантирует нахождение строгого оптимума, позволяет гарантированно находить хорошие квазиоптимальные решения.

3.1. Символьная регрессия для машинного обучения систем управления. Первый этап синтезированного оптимального управления заключается в нахождении функции стабилизации. Математически это

сводится к нахождению нелинейной многомерной функции управления, аргументом которой служит вектор отклонения от устойчивой точки:

$$\mathbf{u} = \mathbf{h}(\mathbf{x}^* - \mathbf{x}), \quad (10)$$

где \mathbf{x}^* – координаты устойчивой точки в пространстве состояний.

Поскольку искомая функция управления реализуется на борту в виде компьютерной программы, такая цифровая реализация управления позволяет отказаться от традиционных структурных ограничений при определении функции управления и применять методы машинного обучения для их поиска в более широком классе нелинейных функций.

Среди методов машинного обучения наибольшей популярностью пользуются методы обучения с учителем, для успешного применения которых необходим значительный объем обучающих данных, а в качестве критерия оптимизации выступает минимизация ошибки, что реализовано, в частности, в нейросетевых методах [35, 36]. Однако на практике инженер редко располагает готовыми наборами данных для обучения. Теоретически их можно получить, регистрируя действия оператора при ручном управлении с последующей аппроксимацией, однако этот подход имеет ограниченное применение. Во-первых, такое управление обычно не является оптимальным, а во-вторых, требует сбора значительного объема данных, что затруднительно в реальных условиях.

К методам машинного обучения без учителя, которые показали свою эффективность для задач синтеза систем управления, относятся методы символьного машинного обучения, такие как символьная регрессия [37]. Символьная регрессия обладает несколькими принципиальными преимуществами. В отличие от параметрической оптимизации, как в линейной регрессии, символьная регрессия ищет как саму структуру, так и параметры функции, что и определяет её название. Процесс поиска полностью автоматизирован. Важнейшим преимуществом перед нейронными сетями является интерпретируемость получаемых решений – функция может быть представлена в аналитическом виде, что даёт ценное понимание физических механизмов управления.

В контексте дискуссии о слабом и сильном искусственном интеллекте, ключевым критерием является способность системы к обобщению - то есть к действиям в условиях, не представленных в обучающей выборке. Методы символьной регрессии демонстрируют этот признак: они синтезируют управляющие функции, руководствуясь исключительно значением целевого функционала и заданным алфавитом элементарных операций. Это позволяет находить непредсказуемые

человеком аналитические решения, чья работоспособность не привязана к конкретным данным обучения, в отличие от моделей типа нейронных сетей. Данное свойство позволяет рассматривать символьную регрессию как шаг в направлении создания систем сильного искусственного интеллекта.

Методы символьной регрессии осуществляют одновременный оптимизационный поиск структуры математического выражения и его параметров. Структура формируется из элементов заданного алфавита примитивов, который включает элементарные функции с различной аргументностью, аргументы и свободные параметры.

Алфавит элементарных функций можно представить как:

$$\begin{aligned}
 F_0 &= \{f_{0,1}, f_{0,2}, \dots, f_{0,v_0}\}, \\
 F_1 &= \{f_{1,1}(z), f_{1,2}(z), \dots, f_{1,v_1}(z)\}, \\
 F_2 &= \{f_{2,1}(z_1, z_2), f_{2,2}(z_1, z_2), \dots, f_{2,v_2}(z_1, z_2)\}, \\
 F_3 &= \{f_{3,1}(z_1, z_2, z_3), f_{3,2}(z_1, z_2, z_3), \dots, f_{3,v_3}(z_1, z_2, z_3)\}, \\
 &\dots
 \end{aligned} \tag{11}$$

где v_i - количество элементов во множестве F_i , i - количество аргументов.

Для поиска математических выражений функций из примитивов алфавита используют композиции вложенных друг в друга элементарных функций:

$$\begin{aligned}
 f_{a_1, b_1} f_{a_2, b_2} f_{a_3, b_3} &= f_{a_1, b_1} \circ f_{a_2, b_2} \circ f_{a_3, b_3} = \\
 &= f_{a_1, b_1}(f_{a_2, b_2}(f_{a_3, b_3}(\dots))).
 \end{aligned} \tag{12}$$

Искомое математическое выражение в методах символьной регрессии ищется в виде кода. Различные методы символьной регрессии используют различные способы кодирования [37], так, например, генетическое программирование представляет функцию в виде двумерного вектора, а декартово генетическое программирование – в виде целочисленной матрицы.

Приведем пример кодирования математического выражения методом сетевого оператора, который будет использоваться в вычислительном эксперименте данной работы.

Рассмотрим следующее математическое выражение:

$$y = (x_1^2 - x_2^2) \cos(q_1 x_1 + q_2) + x_1 x_2 \exp(-q_3 x_1). \tag{13}$$

Математическое выражение (13) кодируется в виде целочисленной верхнетреугольной матрицы сетевого оператора:

$$\Psi = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (14)$$

Для кодирования данного математического выражения использовался следующий алфавит примитивов:

$$F_0 = \{x_1, x_2, q_1, q_2, q_3\}, \quad (15)$$

$$F_1 = \{f_{1,1}(z) = z, f_{1,2}(z) = -z, f_{1,3}(z) = \cos(z), f_{1,4}(z) = \exp(z), f_{1,5}(z) = z^2\}, \quad (16)$$

$$F_2 = \{f_{2,1}(z_1, z_2) = z_1 + z_2, f_{2,2}(z_1, z_2) = z_1 z_2\}. \quad (17)$$

В матрице (13) над диагональю расположены номера унарных операций (16), на диагонали нули фиксируют элементы множества переменных и параметров (15), а ненулевые диагональные элементы соответствуют номерам бинарных операций (17). Именно в таком закодированном матричном виде (14) функция управления представлена на бортовом вычислителе робота. Более подробно со способами кодирования и вычисления математических выражений по матрице сетевого оператора можно ознакомиться в [37, 38].

Задача синтеза функций путём поиска в пространстве их кодовых представлений представляет собой вычислительно сложную проблему нечисловой оптимизации. Сложность возникает из-за необходимости оптимизационного поиска в пространстве структур при одновременной

оценке их эффективности в метрическом числовом пространстве по значению целевого функционала. Данная проблема является одной из наиболее трудоёмких в современной теории оптимизации.

Среди современных оптимизационных методов генетический алгоритм занимает особое положение благодаря способности работать с нечисловыми пространствами поиска. Ключевое преимущество генетического алгоритма в контексте машинного обучения систем управления заключается в том, что он оперирует не арифметическими вычислениями, а эволюционными операторами скрещивания и мутации, что позволяет эффективно исследовать пространство возможных решений и обеспечивает принципиально иной подход к поиску оптимальных решений.

3.2. Эволюционный поиск квазиоптимального положения управляющих точек равновесия. Второй этап принципа синтезированного оптимального управления направлен на поиск оптимального расположения точек равновесия и решается как задача параметрической оптимизации. Данная задача может быть решена различными методами глобальной оптимизации. Однако учитывая, что функционалы в задачах управления робототехническими системами, как правило, являются многоэкстремальными, особенно при наличии фазовых ограничений, а в контексте задачи многоагентного коллаборативного управления фазовые ограничения имеют дополнительную сложность, в этой связи в работе предлагается применение эволюционных алгоритмов.

4. Вычислительный эксперимент. Рассмотрим пример модельной задачи коллаборативного оптимального управления тремя колесными роботами с дифференциальным приводом. Такая задача совместного выполнения работы может встретиться при уборке территорий, обработке сельскохозяйственных полей, выполнении технологических задач на заводских территориях и т.д.

Положим, что математические модели роботов одинаковы и описываются следующей системой обыкновенных дифференциальных уравнений:

$$\begin{aligned}\dot{x}_1^j &= 0.5(u_1^j + u_2^j) \cos(x_3^j), \\ \dot{x}_2^j &= 0.5(u_1^j + u_2^j) \sin(x_3^j), \\ \dot{x}_3^j &= 0.5(u_1^j - u_2^j),\end{aligned}\tag{18}$$

где x_1^j, x_2^j - линейные координаты на плоскости в дм, x_3^j - угол поворота продольной оси робота в рад, u_1^j, u_2^j - линейные скорости движения колес в дм/с, $j=1,2,3$.

Заданы ограничения на управления:

$$-10 = u^- \leq u_i^j \leq u^+ = 10, \quad i = 1, 2, \quad j = 1, 2, 3, \quad (19)$$

начальные состояния роботов:

$$\mathbf{x}^{0,1} = [0 \ 0 \ 0]^T, \quad \mathbf{x}^{0,2} = [5 \ 0 \ 0]^T, \quad \mathbf{x}^{0,3} = [10 \ 0 \ 0]^T, \quad (20)$$

и терминальные состояния роботов:

$$\mathbf{x}^{f,1} = [0 \ 10 \ 0]^T, \quad \mathbf{x}^{f,2} = [5 \ 10 \ 0]^T, \quad \mathbf{x}^{f,3} = [10 \ 10 \ 0]^T. \quad (21)$$

Заданы управляемые фазовые ограничения:

$$\varphi_k(\mathbf{x}^j, z) = r_k - \sqrt{(x_{k,1} + z - x_1^j)^2 + (x_{k,2} - x_2^j)^2} \leq 0, \quad (22)$$

где $k = 1, 2, j = 1, 2, 3, r_1 = r_2 = 2, x_{1,1} = 2, x_{2,1} = 8, z$ - параметр, управляемый человеком,

$$\dot{z} = az + v_1, \quad (23)$$

$a = 0.1, v_1$ - компонента управления человеком, $-4 \leq v_1 \leq 4$. В качестве (23) было выбрано одно из простейших дифференциальных уравнений, так как целью было просто ввести изменения фазовых ограничений.

Заданы динамические фазовые ограничения, которые определяют условия столкновения роботов:

$$\chi(\mathbf{x}^j, \mathbf{x}^k) = r_0 - \sqrt{(x_1^j - x_1^k)^2 + (x_2^j - x_2^k)^2} \leq 0, \quad (24)$$

$r_0 = 1.25, j = 1, 2, k = j + 1, \dots, 3$.

Задан критерий качества:

$$\begin{aligned}
 J_4 = t_f + p_1 \sum_{j=1}^3 \sqrt{\sum_{i=1}^3 (x_i^{f,j} - x_i^j)^2} + p_2 \int_0^{t_f} \sum_{j=1}^3 \sum_{i=1}^2 \vartheta(\varphi_i(\mathbf{x}^j, z)) dt + \\
 + p_3 \int_0^{t_f} \sum_{j=1}^2 \sum_{k=j+1}^3 \vartheta(\chi(\mathbf{x}^j, \mathbf{x}^k)) dt \rightarrow \min, \quad (25)
 \end{aligned}$$

где $p_1 = 1$, $p_2 = 3$, $p_3 = 3$, t_f – время достижения терминальных состояний всеми роботами, определяемое из следующих соотношений:

$$t_f = \max\{t_{f,1}, t_{f,2}, t_{f,3}\}, \quad (26)$$

$$t_{f,j} = \begin{cases} t, & \text{если } t < t^+ \text{ и } \|\mathbf{x}^{f,j} - \mathbf{x}^j(t)\| \leq \varepsilon \\ t^+ & \text{иначе} \end{cases}, \quad (27)$$

$$t^+ = 1.4, \varepsilon = 0.025.$$

Первоначально, используя прямой подход к решению задачи оптимального управления в классической постановке, найдем функции управления как функции времени. Методика предполагает дискретизацию времени с постоянным шагом $\Delta t = 0.2$ с. Далее для каждой компоненты управления каждого робота ищем на границах интервалов значения постоянных параметров в ограниченном диапазоне значений. Внутри интервала значения параметров соединяем прямыми отрезками. Если функция управления внутри интервала выходит за границы ограничений на управления (19), то обрезаем значение управления до заданного ограничения. В результате искомая кусочно-линейная функция имеет следующий вид:

$$w_i^j = \begin{cases} u_i^-, & \text{если } \tilde{u}_i^j \leq u_i^- \\ u_i^+, & \text{если } u_i^+ \leq \tilde{u}_i^j \\ \tilde{u}_i^j, & \text{иначе} \end{cases} \quad (28)$$

где,

$$\tilde{u}_i^j = q_{i+3(j-1)+6(k-1)} + \frac{t - (k-1)\Delta t}{\Delta t} (q_{i+3(j-1)+6k} - q_{i+3(j-1)+6(k-1)}), \quad (29)$$

где $\mathbf{q} = [q_1 \dots q_p]^T$ – искомый вектор управления, k – номер интервала, $k\Delta t \leq t \leq (k-1)\Delta t$, $k = 1, \dots, \lfloor t^+/\Delta t \rfloor$.

Так как в задаче $t^+ = 1.4$, $\Delta t = 0.2$, то интервалов 7, границ интервалов 8. Для каждой границы интервала необходимо найти 6 значений управлений, по два управления для каждого робота. Поэтому размерность вектора искомых параметров составляет величину $p = 6 \cdot 8 = 48$. Значения параметров ограничены $-10 = q^+ \leq q_i \leq q^- = -20$.

Для поиска решения используем гибридный эволюционный алгоритм [39]. Полученное решение имеет следующие значения:

$$\mathbf{q} = [12.5231 \quad -16.8999 \quad -18.4437 \quad 9.8319 \quad 18.6029 \quad -16.4049 \\ 11.2881 \quad 3.1847 \quad -8.0429 \quad -1.0858 \quad 15.4253 \quad 6.5292 \quad 17.1205 \\ 19.9991 \quad -10.2189 \quad -13.5536 \quad 11.9678 \quad 17.6286 \quad 7.9180 \quad 13.6077 \\ -11.5984 \quad -13.6049 \quad 18.7199 \quad 19.2897 \quad 17.6816 \quad 11.4286 \quad -16.7217 \\ -12.0699 \quad 17.2628 \quad 1.3383 \quad 14.5977 \quad 16.4364 \quad -9.2731 \quad 15.7315 \\ 9.9467 \quad 18.1820 \quad -4.1368 \quad 10.4489 \quad -3.3799 \quad -7.2057 \quad 3.8314 \\ 19.5919 \quad 1.6495 \quad 12.2760 \quad 12.1812 \\ -11.1262 \quad -18.9317 \quad 15.1346]^T.$$

Движение роботов согласно найденному оптимальному управлению представлено на рисунке 1.

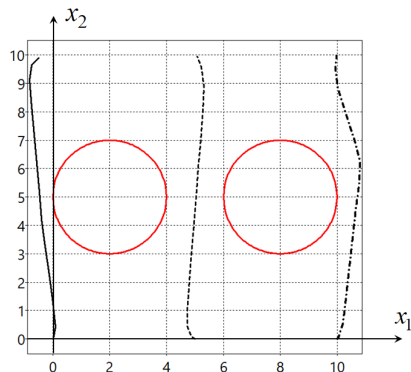


Рис. 1. Решение задачи оптимального управления в классической постановке

Для проверки реализуемости полученного решения подвергнем случайным возмущениям начальные состояния роботов:

$$x_i^{0,j} \leftarrow x_i^{0,j} + B_0(2\xi - 1), \quad (30)$$

где $B_0 = 0.4$, ξ - генератор случайных чисел, который при каждом обращении возвращает случайное равномерно распределенное число в диапазоне от 0 до 1. На рисунке 2 представлены по 4 траектории движения каждого робота для найденного оптимального управления (28), (29) из возмущенных начальных состояний.

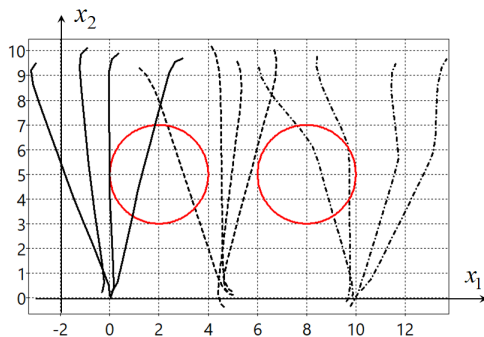


Рис. 2. Траектории движения роботов при оптимальном управлении для возмущенных начальных состояний

Как видно из результатов вычислительных экспериментов, полученное оптимальное управление, как функция времени, не может быть напрямую реализовано на реальном объекте, так как малые возмущения начального состояния приводят к отсутствию достижения объектом терминального состояния и к существенному нарушению фазовых ограничений.

Для решения задачи оптимального коллаборативного управления в классе реализуемых на практике функций управления используем принцип синтезированного оптимального управления. Согласно данному методу первоначально делаем объект управления устойчивым относительно точки в пространстве состояний. Для этой цели необходимо синтезировать систему стабилизации.

Для решения задачи синтеза системы стабилизации и нахождения математического выражения как функции (10) используем численный метод символьной регрессии на основе сетевого оператора [38]. Исследование метода синтезированного оптимального управления [33] показало, что качество синтезированного управления и его чувствительность к внешним возмущениям, а, следовательно, и его реализуемость зависят от качества системы стабилизации. Для обеспечения необходимого качества системы стабилизации, определяемого заданным функционалом, для объекта управления первоначально было задано множество начальных состояний и одно терминальное состояние, которое определяло положение точки устойчивости. Далее для всех начальных состояний была решена задача оптимального управления и были найдены оптимальные траектории движения объекта в точку устойчивости. Метод символьной регрессии использовался для нахождения функции аппроксимации всех оптимальных траекторий.

Функция системы стабилизации данного типа робота, найденная методом сетевого оператора имеет следующий вид:

$$u_i = h_i(\mathbf{x}^* - \mathbf{x}) = \begin{cases} u^+, & \text{если } u^+ \leq \hat{h}_i(\mathbf{x}^f - \mathbf{x}), \\ u^-, & \text{если } \hat{h}_i(\mathbf{x}^* - \mathbf{x}) \leq u^-, \\ \hat{h}_i(\mathbf{x}^* - \mathbf{x}), & \text{иначе,} \end{cases} \quad (31)$$

где:

$$\hat{h}_1(\mathbf{x}^* - \mathbf{x}) = \mu(H) + G + \mu(F) + \sin(E) + C^3 + \rho_{19}(\sin(r_1(x_1^f - x_1))) + \text{sgn}(x_1^* - x_1), \quad (32)$$

$$\hat{h}_2(\mathbf{x}^* - \mathbf{x}) = \ln(|\hat{h}_1(\mathbf{x}^* - \mathbf{x})|) + H + \text{sgn}(G)\sqrt{G} + \rho_{18}(F) + E^3 + \exp(C) + \tanh(\sin(r_1(x_1^* - x_1))) + \rho_{19}(r_3(x_3^* - x_3)) + (r_1(x_1^* - x_1))^{-1} + \exp((x_1^* - x_1)), \quad (33)$$

$$A = r_3(x_3^* - x_3) + \rho_{19}(r_1(x_1^* - x_1))r_2(x_2^* - x_2),$$

$$B = \arctan(\sin(r_1(x_1^* - x_1))) + \rho_{19}(r_3(x_3^* - x_3)) + \sin(\rho_{19}(r_1(x_1^* - x_1))r_2(x_2^* - x_2)) + \rho_{19}(r_2),$$

$$C = \sin(B) + A + \text{sgn}(\sin(r_1(x_1^* - x_1)))\sqrt{|\sin(r_1(x_1^* - x_1))|},$$

$$D = \sqrt[3]{r_1(x_1^* - x_1)} + \sin(r_1),$$

$$E = \arctan(D) + \sin(r_1(x_1^* - x_1)) + \text{sgn}(r_1(x_1^* - x_1)) + \rho_{19}(r_3),$$

$$F = \text{sgn}(E) + \text{sgn}(D)\sqrt{|D|} + C - C^3 + \sin(\sin(r_1(x_1^* - x_1))) + \tanh(r_3(x_3^* - x_3)) + \arctan(x_1^* - x_1),$$

$$G = \text{sgn}(F)\sqrt{|F|} + \text{sgn}(E) + \rho_{19}(D) + \rho_{18}(C) + \rho_{19}(B) + \rho_{18}(A) +$$

$$\sin(r_1(x_1^* - x_1)) + r_1(x_1^* - x_1) - (r_1(x_1^* - x_1))^3 + (x_3^* - x_3)^2,$$

$$H = F^3 - D + B^3 + \sin(r_1(x_1^* - x_1)) - (\sin(r_1(x_1^* - x_1)))^3 + \rho_{19}(r_1(x_1^* - x_1))r_2(x_2^* - x_2) + \cos(r_1) + \ln(|x_1^* - x_1|),$$

$$\mu(\alpha) = \begin{cases} \alpha, & \text{if } |\alpha| < 1 \\ \text{sgn}(\alpha), & \text{otherwise} \end{cases},$$

$$\rho_{18}(\alpha) = \text{sgn}(\alpha)(\exp(\alpha) - 1), \rho_{19}(\alpha) = \text{sgn}(\alpha) \exp(-|\alpha|),$$

$$r_1 = 0.1646, r_2 = 2.0483, r_3 = 3.6226.$$

Выражения (32), (33) представлены более информативно, чтобы исследователь имел понимание, какая функция описывает систему

управления. Управляющие сигналы для объекта h_i , $i = 1, 2, 3$, в реальном времени формируются бортовым программным модулем [40] на основе расчётов по матрице сетевого оператора.

На втором этапе подхода синтезированного оптимального управления решаем задачу оптимального управления согласно поставленному в исходной задаче критерию. Для этой цели задаем время процесса управления и разбиваем его на равные интервалы. На каждом интервале находим положение точки устойчивого равновесия для каждого объекта управления. В данном случае управлением являются значения постоянных параметров, которые определяют положения устойчивой точки равновесия. Последовательная активация найденных точек устойчивого равновесия заставляет двигаться объект управления к цели с оптимальным значением заданного критерия качества. В данном случае объект управления в процессе движения чаще всего не достигает устойчивой точки равновесия, но при движении к ней обеспечивается компенсация ошибки определения состояния объекта управления, так как при приближении к точке устойчивости все траектории в ее окрестности сближаются. Для рассматриваемой в экспериментальной части задачи было задано время управления $t^+ = 2$ с и длительность интервала времени $\Delta t = 0.5$ с. Таким образом, при решении задачи оптимального управления необходимо было найти по четыре точки устойчивого равновесия, т.е. всего по три координаты для каждой точки равновесия на четырех интервалах времени для трех объектов управления, $3 \times 4 \times 3 = 36$ параметров. Значение параметров определялись в области, заданной ограничениями относительно текущего состояния на начало интервала активации точек устойчивости для каждого объекта управления, $-2 \leq x_1 \leq 2$, $-1 \leq x_2 \leq 5$, $-\pi/6 \leq x_3 \leq \pi/6$.

Для решения задачи был использован также гибридный эволюционный алгоритм, который нашел следующие оптимальные значения параметров:

$$\mathbf{q}^* = [q_1^* \dots q_{36}^*]^T = \begin{bmatrix} 1.4909 & 3.4420 & 0.2946 & -1.8166 & 2.9330 & -0.4168 \\ 1.3954 & 1.4734 & 0.2505 & -1.6667 & 4.2596 & 0.3877 \\ -0.6391 & 4.0990 & 0.2634 & -1.7895 & 4.5150 & -0.2693 \\ 1.6220 & 2.3671 & 0.0598 & -1.7335 & 1.3662 & 0.3996 \\ -0.8450 & 2.3323 & 0.1523 & 3.0837 & -0.6138 & 0.0647 \\ 1.9112 & 2.5906 & 0.5183 & -0.7703 & 2.1947 & -0.2891 \end{bmatrix}^T.$$

На рисунке 3 приведены графики траекторий движения роботов при синтезированном оптимальном управлении и найденные точки устойчивости для каждого робота. На рис. 4 приведены траектории движения роботов при возмущенных начальных состояниях (30). Как видно из сравнения графиков, представленных на рисунках 2 и 4, синтезированное оптимальное управление сохранило формы движения траекторий. Все возмущенные траектории лежат в окрестности оптимальных. Из сказанного следует вывод, что синтезированное оптимальное управление относится к классу непосредственно реализуемых на реальном объекте функций управления.

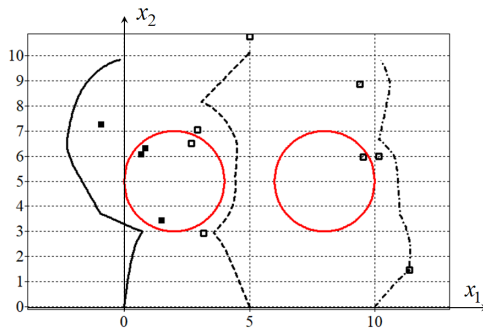


Рис. 3. Траектории движения роботов при синтезированном оптимальном управлении и точки равновесия для каждого робота

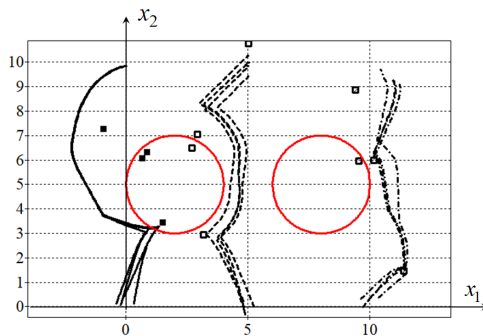


Рис. 4. Траектории движения роботов при синтезированном оптимальном управлении для возмущенных начальных состояний

Для объективной оценки свойства реализуемости был проведен ряд испытаний для функций синтезированного оптимального управления. Результаты испытаний приведены в таблице 1.

Таблица 1. Результаты испытаний синтезированного оптимального управления

| № | Оптимальное | Синтезированное |
|----|-------------|-----------------|
| 1 | 5.8595 | 4.8576 |
| 2 | 7.7290 | 4.3784 |
| 3 | 6.8065 | 4.0595 |
| 4 | 6.3394 | 3.9429 |
| 5 | 8.4089 | 4.7385 |
| 6 | 6.2559 | 5.0455 |
| 7 | 6.4760 | 4.2469 |
| 8 | 8.0961 | 5.7028 |
| 9 | 5.1672 | 3.9700 |
| 10 | 6.7853 | 4.7207 |
| 11 | 6.8748 | 4.9139 |
| 12 | 4.1705 | 4.1690 |
| 13 | 5.1592 | 5.3029 |
| 14 | 6.1010 | 5.6617 |
| 15 | 8.2401 | 3.8764 |
| 16 | 6.6732 | 4.2760 |
| 17 | 4.8250 | 4.7730 |
| 18 | 6.4008 | 3.7630 |
| 19 | 8.5948 | 3.9126 |
| 20 | 3.6610 | 3.8596 |
| 21 | 5.6248 | 4.6421 |
| 22 | 6.0784 | 5.1135 |
| 23 | 5.3511 | 4.2617 |
| 24 | 6.7288 | 4.6106 |

В таблице 1 представлены значения функционала, полученные для оптимального и синтезированного управления со случайными возмущениями начальных состояний роботов. По результатам 24 испытаний среднее значение функционала для оптимального управления составило величину 6.3503, среднеквадратичное отклонение составило величину 1.2546. Для синтезированного управления среднее значение функционала равно 4.5333, среднеквадратичное отклонение 0.5541.

Результаты испытаний подтверждают ранее установленный вывод о наличии свойства реализуемости синтезированного управления и его

отсутствия у функции оптимального управления, полученной при решении задачи оптимального управления в классической постановке.

5. Заключение. В работе предложено оригинальное формальное описание задач коллаборативного многоагентного управления в терминах оптимального управления, где участие человека определяется через фазовые ограничения. Слабоформализуемое участие человека в коллаборативной системе в предложенном формальном описании выражено через особые фазовые ограничения, имеющие в описании дополнительный свободный вектор управления, реализуемый человеком. Представленный подход, основанный на классической постановке задачи оптимального управления, позволяет искать решения с учетом строгих математических критериев эффективности работы коллаборативной системы. Формальное описание задачи является также основой для применения численных машинных методов ее решения, позволяя автоматизировать процесс синтеза закона управления. Сформулированное описание служит основой для перехода к этапу выбора и разработки алгоритмов решения задачи коллаборативного оптимального управления множеством агентов. В работе представлена возможность применения как прямых численных подходов, так и интеллектуальных методов машинного обучения. Центральным результатом является метод синтеза оптимального управления, позволяющий в автоматическом режиме получать управляющие законы из общей формальной постановки, учитывающей сложность фазового пространства. Как показано в вычислительном примере, машинно синтезированное управление эффективно работает в условиях возможных неопределенностей состояния реального объекта.

Литература

1. Ющенко А.С. Коллаборативная робототехника: состояние и новые задачи // Мехатроника, автоматизация, управление. 2017. Т. 18. №12. С. 812–819. DOI: 10.17587/mau.18.812-819.
2. Simoes A.C., Pinto A., Santos J., Pinheiro S., Romero D. Designing Human-Robot Collaboration (HRC) Workspaces in Industrial Settings: A Systematic Literature Review // Journal of Manufacturing Systems. 2022. vol. 62. pp. 1–14. DOI: 10.1016/j.jmsy.2021.11.007.
3. Hameed A., Ordys A., Mozaryn J., Sibilska-Mroziewicz A. Control System Design and Methods for Collaborative Robots: Review // Applied Sciences. 2023. vol. 13. no. 1. pp. 1–23. DOI: 10.3390/app13010675.
4. Colgate J.E., Wannasupphrasit W., Peshkin M.A. Cobots: Robots for Collaboration with Human Operators // Proceedings of the ASME International Conference on Mechanical Engineering Congress and Exposition. 1996. pp. 433–439. DOI: 10.1115/IMECE1996-0367.
5. Case Studies - Collaborative Robots // IFR. URL: <https://ifr.org/case-studies/case-studies-collaborative-robots> (дата обращения: 08.09.2025).

6. Ajoudani A., Zanchettin A.M., Ivaldi S., Albu-Schaffer A., Kosuge K., Khatib O. Progress and prospects of the human robot collaboration // *Autonomous Robots*. 2018. vol. 42. pp. 957–975.
7. Villani V., Pini F., Leali F., Secchi C. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications // *Mechatronics*. 2018. vol. 55. pp. 1–19.
8. ГОСТ Р 60.1.2.3-2021/ISO/TS 15066:2016. Национальный стандарт Российской Федерации. Роботы и робототехнические устройства. Требования безопасности для роботов, работающих совместно с человеком. URL: <https://docs.cntd.ru/document/1200180499> (дата обращения: 07.07.2025).
9. Серебрянный В.В. Коллаборативные мультиагентные системы – альтернатива полной автоматизации производства // *Мехатроника, автоматизация, управление*. 2020. Т. 21. №7. С. 404–411. DOI: 10.17587/mau.21.404-411.
10. Чикрин Д.Е., Смольникова К.Р. Обзор коллаборативных робототехнических систем и юридико-системные аспекты взаимодействия с ними // *Известия ЮФУ. Технические науки*. 2023. №3. С. 25–35. DOI: 10.18522/2311-3103-2023-3-25-35.
11. Peshkin M.A., Colgate J.E., Wannasuphprasit W., Moore C.A., Gillespie R.B., Akella P. Cobot Architecture // *IEEE Transactions on Robotics and Automation*. 2001. vol. 17. no. 4. pp. 377–390.
12. Rato D., Oliveira M.O., Santos V.M.F., Gomes V., Sappa A.D. A sensor-to-pattern calibration framework for multi-modal industrial collaborative cells // *Journal of Manufacturing Systems*. 2022. vol. 64. pp. 497–507. DOI: 10.48550/arXiv.2210.10365.
13. Marvel J.A. Sensors for Safe, Collaborative Robots in Smart Manufacturing // *IEEE Sensors*. 2017. pp. 1–3. DOI: 10.1109/ICSENS.2017.8234264.
14. Magrini E., Ferraguti F., Ronga A.J., Pini F., Luca A.D., Leali F. Human-Robot Coexistence and Interaction in Open Industrial Cells // *Robotics and Computer-Integrated Manufacturing*. 2020. vol. 61. pp. 1–33. DOI: 10.1016/j.rcim.2019.101846.
15. Gutierrez G.M., Rincon J.A., Julian V. Federated Learning for Collaborative Robotics: A ROS 2-Based Approach // *Electronics*. 2025. vol. 14. no. 7. 1323 p. DOI: 10.3390/electronics14071323.
16. Tamali A., Amardjia N., Tamali M. Distributed and autonomous multi-robot for task allocation and collaboration using a greedy algorithm and robot operating system platform // *IAES International Journal of Robotics and Automation*. 2024. vol. 13. no. 2. pp. 205–219. DOI: 10.11591/ijra.v13i2.pp205-219.
17. Sosa-Ceron A.D., Gonzalez-Hernandez H.G., Reyes-Avendano J.A. Learning from Demonstrations in Human–Robot Collaborative Scenarios: A Survey // *Robotics*. 2022. vol. 11. no. 6. pp. 1–23. DOI: 10.3390/robotics11060126.
18. Semeraro F., Griffiths A., Cangelosi A. Human–robot collaboration and machine learning: A systematic review of recent research // *Robotics and Computer-Integrated Manufacturing* 2023. vol. 79. pp. 1–16.
19. Sandakalum T., Ang M.H.Jr. Motion Planning for Mobile Manipulators – A Systematic Review // *Machines*. 2022. vol. 10. 97 p. DOI: 10.3390/machines10020097.
20. Chen H., Zang X., Zhu Y., Zhao J. Hybrid Sampling-Based Path Planning for Mobile Manipulators Performing Pick and Place Tasks in Narrow Spaces // *Applied Sciences*. 2024. vol. 14. no. 22. 10313 p. DOI: 10.3390/app142210313.
21. Serebrenny V., Shmalko E. Multicriterial analyses of Pareto-efficiency for collaborative multi-agent systems using genetic algorithm with variations // *Procedia Computer Science*. 2021. vol. 186. pp. 793–801. DOI: 10.1016/j.procs.2021.04.201.

22. Jafar F.A., Abdullah N., Muhammad M.N., Zakaria N.A., et al. Investigation of Human Emotional State in Human-Robot Collaboration // *Journal of Computers*. 2014. vol. 9. no. 3. pp. 668–677.
23. De Benedictis C., Franco W., Maffiodo D., Ferraresi C. Control of Force Impulse in Human–Machine Impact // *Mechanisms and Machine Science*. Springer. 2018. vol. 49. pp. 956–964. DOI: 10.1007/978-3-319-61276-8_102.
24. Hogan N. Impedance Control: An Approach to Manipulation: Part I Theory // *Journal of Dynamic Systems, Measurement, and Control*. 1985. 107. pp. 1–7.
25. Parker C.A.C., Croft E.A. Design Personalization of a Cooperative Carrying Robot Controller // *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. 2012. pp. 3916–3921.
26. Woods A.C., La H.M., Ha Q.P. A novel extended potential field controller for use on aerial robots // *Proceedings of the IEEE International Conference on Automation Science and Engineering (CASE)*. 2016. pp. 286–291.
27. Li S., Wang H., Zhang S. Human-Robot Collaborative Manipulation with the Suppression of Human-caused Disturbance // *Journal of Intelligent Robotic Systems*. 2021. vol. 102. 73 p. DOI: 10.1007/s10846-021-01429-8.
28. Haddadin S., Croft E. Physical Human-Robot Interaction // *Springer Handbook of Robotics*. Cham: Springer. 2016. pp. 1835–1874. DOI: 10.1007/978-3-319-32552-1_69.
29. Karamzin D., Pereira F.L. On Higher-Order State Constraints // *SIAM Journal on Control and Optimization*. 2023. vol. 61. no. 4. pp. 1913–1933. DOI: 10.1137/22M1487497.
30. Hartl R.F., Sethi S., Vickson R. A Survey of the Maximum Principles for Optimal Control Problems with State Constraints // *SIAM Review*. 1995. vol. 37. no. 2. pp. 181–218.
31. Карамзин Д.Ю. Некоторые оценки для скачка производной функции-множителя Лагранжа в задачах оптимального управления с фазовыми ограничениями второго порядка // *Известия Иркутского государственного университета. Серия Математика*. 2024. Т. 49. С. 3–15.
32. Шмалько Е.Ю. Машинно синтезированное управление нелинейным динамическим объектом на основе оптимального расположения точек равновесия // *Информатика и автоматизация*. 2023. Т. 22. №1. С. 87–109. DOI: 10.15622/ia.22.1.4.
33. Diveev A., Shmalko E., Serebrenny V., Zentay P. Fundamentals of Synthesized Optimal Control // *Mathematics*. 2020. vol. 9. no. 1. 21 p.
34. Shmalko E. Feasibility of Synthesized Optimal Control Approach on Model of Robotic System with Uncertainties / In: Ronzhin A., Shishlakov V. (eds) // *Electromechanics and Robotics. Smart Innovation, Systems and Technologies*. Springer. 2022. vol. 232. pp. 131–143. DOI: 10.1007/978-981-16-2814-6_12.
35. Sanchez-Sanchez C., Izzo D., Hennes D. Learning the optimal state-feedback using deep networks // *Proceedings of the IEEE Conference on Symposium Series on Computational Intelligence (SSCI)*. 2016. pp. 1–8.
36. Diveev A. Synthesis of Optimal Stabilization System by Supervised Machine Learning of Symbolic Regression // *Lecture Notes in Networks and Systems (LNNS)*. Springer. 2025. vol. 1424. pp. 258–275. DOI: 10.1007/978-3-031-92605-1_17.
37. Diveev A., Shmalko E. Machine Learning Control by Symbolic Regression // Cham: Springer. 2021. 155 p. DOI: 10.1007/978-3-030-83213-1.
38. Дивеев А.И. Численный метод сетевого оператора для синтеза системы управления с неопределенными начальными значениями // *Известия Российской академии наук. Теория и системы управления*. 2012. №2. 63 с.
39. Diveev A. Hybrid Evolutionary Algorithm for Optimal Control Problem // *Lecture Notes in Networks and Systems (LNNS)*. Springer. 2023. vol. 543. pp. 726–738. DOI: 10.1007/978-3-031-16078-3_50.

40. Program module of the network operator // GitHub. URL: https://github.com/KostyaYamshanov/Machine-Learning-Control-kit/tree/main/network_operator.

Шмалько Елизавета Юрьевна — д-р техн. наук, старший научный сотрудник, старший научный сотрудник, Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН). Область научных интересов: численные методы, эволюционные вычисления, оптимальное управление, символьная регрессия. Число научных публикаций — 150. e.shmalko@gmail.com; улица Вавилова, 44, 119333, Москва, Россия; р.т.: +7(964)-636-6669.

Дивеев Асхат Ибрагимович — д-р техн. наук, профессор, главный научный сотрудник, Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН). Область научных интересов: оптимальное управление, символьная регрессия, робототехника. Число научных публикаций — 450. aidiveev@mail.ru; улица Вавилова, 44, 119333, Москва, Россия; р.т.: +7(905)711-4427.

Ямшанов Константин Леонидович — аспирант, Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН). Область научных интересов: робототехника, машинное обучение, управление. Число научных публикаций — 5. yam.kostya@yandex.ru; улица Вавилова, 44, 119333, Москва, Россия; р.т.: +7(923)249-8191.

Громов Иван Анатольевич — аспирант, Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН); инженер, Московский государственный технологический университет «СТАНКИН» (МГТУ «СТАНКИН»). Область научных интересов: робототехника, управление, машинное обучение. Число научных публикаций — 5. 8357743@gmail.com; улица Вавилова, 44, 119333, Москва, Россия; р.т.: +7(905)748-5240.

Поддержка исследований. Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации (проект FSFS-2024-0012).

E.Yu. SHMALKO, A.I. DIVEEV, K.L. YAMSHANOV, I.A. GROMOV
**SYMBOLIC MACHINE LEARNING FOR QUASI-OPTIMAL
CONTROL OF A COLLABORATIVE MULTI-AGENT ROBOTIC
SYSTEM**

Shmalko E.Yu., Diveev A.I., Yamshanov K.L., Gromov I.A. **Symbolic Machine Learning for Quasi-Optimal Control of a Collaborative Multi-Agent Robotic System.**

Abstract. The paper presents new results of applying symbolic machine learning methods to automate the development of a control system as applied to a collaborative multi-agent system. The article proposes an original mathematical formulation of the problem of optimal control of a collaborative multi-agent system, where collaboration implies joint activity of robots and humans in a single complex phase space to achieve the control goal with the optimal value of a given quality criterion. The presented formulation proposes to consider the role of a human in the form of phase constraints with uncertainty and an additional control vector, with the help of which the influence of these phase constraints is reduced in terms of minimizing the functional. The solution to the problem is a multidimensional control function that transfers a technical object, a group of mobile robots, to the target state, taking into account possible uncertainties in the state of the object, while minimizing the specified quality criterion of the process. The task is to automatically find both the optimal structure and the parameters of the control function. Such a general statement of the control problem cannot be solved directly by existing analytical methods due to its complexity and requires a lot of manual work to select the appropriate controllers and optimally adjust their parameters. The paper proposes an approach based on symbolic machine learning using symbolic regression and evolutionary optimization to solve such a generalized problem. The effectiveness of the approach is demonstrated by numerical experiments. An example of an optimal control problem for three mobile robots with two controlled phase constraints is given. The applied problem is considered in a general mathematical statement, which allows using universal methods of symbolic machine learning to automatically find a solution.

Keywords: machine learning, optimal control, phase constraints, collaborative multi-agent robotic system, symbolic regression.

References

1. Yushchenko A.S. [Collaborative robotics: state of art and new challenges]. *Mekhatronika, avtomatizatsiya, upravlenie – Mechatronics, Automation, Control*. 2017. vol. 18. no. 12. pp. 812–819. DOI: 10.17587/mau.18.812-819. (In Russ.).
2. Simoes A.C., Pinto A., Santos J., Pinheiro S., Romero D. Designing Human-Robot Collaboration (HRC) Workspaces in Industrial Settings: A Systematic Literature Review // *Journal of Manufacturing Systems*. 2022. vol. 62. pp. 1–14. DOI: 10.1016/j.jmsy.2021.11.007.
3. Hameed A., Ordys A., Mozaryn J., Sibilska-Mroziewicz A. Control System Design and Methods for Collaborative Robots: Review // *Applied Sciences*. 2023. vol. 13. no. 1. pp. 1–23. DOI: 10.3390/app13010675.
4. Colgate J.E., Wannasuphprasit W., Peshkin M.A. Cobots: Robots for Collaboration with Human Operators // *Proceedings of the ASME International Conference on Mechanical Engineering Congress and Exposition*. 1996. pp. 433–439. DOI: 10.1115/IMECE1996-0367.

5. Case Studies - Collaborative Robots // IFR. Available at: <https://ifr.org/case-studies/case-studies-collaborative-robots> (accessed 08.09.2025).
6. Ajoudani A., Zanchettin A.M., Ivaldi S., Albu-Schaffer A., Kosuge K., Khatib O. Progress and prospects of the human robot collaboration // *Autonomous Robots*. 2018. vol. 42. pp. 957–975.
7. Villani V., Pini F., Leali F., Secchi C. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications // *Mechatronics*. 2018. vol. 55. pp. 1–19.
8. GOST R 60.1.2.3-2021/ISO/TS 15066:2016. [Roboty i robototekhnicheskie ustroystva. Trebovaniya bezopasnosti dlya robotov, rabotayushchikh sovместno s chelovekom]. Available at: <https://docs.cntd.ru/document/1200180499> (accessed 07.07.2025).
9. Serebrenny V.V. [Collaborative multi-agent systems as an alternative to full automation of production]. *Mekhatronika, avtomatizatsiya, upravlenie – Mechatronics, Automation, Control*. 2020. vol. 21. no. 7. pp. 404–411. (In Russ.).
10. Chikrin D.E., Smolnikova K.R. [Review of collaborative robotic systems and legal-system aspects of interaction with them]. *Izvestiya YuFU. Tekhnicheskie nauki – Proceedings of Southern Federal University. Engineering Sciences*. 2023. no. 3 (233). pp. 25–35. (In Russ.).
11. Peshkin M.A., Colgate J.E., Wannasuphprasit W., Moore C.A., Gillespie R.B., Akella P. Cobot Architecture // *IEEE Transactions on Robotics and Automation*. 2001. vol. 17. no. 4. pp. 377–390.
12. Rato D., Oliveira M.O., Santos V.M.F., Gomes V., Sappa A.D. A sensor-to-pattern calibration framework for multi-modal industrial collaborative cells // *Journal of Manufacturing Systems*. 2022. vol. 64. pp. 497–507. DOI: 10.48550/arXiv.2210.10365.
13. Marvel J.A. Sensors for Safe, Collaborative Robots in Smart Manufacturing // *IEEE Sensors*. 2017. pp. 1–3. DOI: 10.1109/ICSENS.2017.8234264.
14. Magrini E., Ferraguti F., Ronga A.J., Pini F., Luca A.D., Leali F. Human-Robot Coexistence and Interaction in Open Industrial Cells // *Robotics and Computer-Integrated Manufacturing*. 2020. vol. 61. pp. 1–33. DOI: 10.1016/j.rcim.2019.101846.
15. Gutierrez G.M., Rincon J.A., Julian V. Federated Learning for Collaborative Robotics: A ROS 2-Based Approach // *Electronics*. 2025. vol. 14. no. 7. 1323 p. DOI: 10.3390/electronics14071323.
16. Tamali A., Amardjia N., Tamali M. Distributed and autonomous multi-robot for task allocation and collaboration using a greedy algorithm and robot operating system platform // *IAES International Journal of Robotics and Automation*. 2024. vol. 13. no. 2. pp. 205–219. DOI: 10.11591/ijra.v13i2.pp205-219.
17. Sosa-Ceron A.D., Gonzalez-Hernandez H.G., Reyes-Avendano J.A. Learning from Demonstrations in Human–Robot Collaborative Scenarios: A Survey // *Robotics*. 2022. vol. 11. no. 6. pp. 1–23. DOI: 10.3390/robotics11060126.
18. Semeraro F., Griffiths A., Cangelosi A. Human–robot collaboration and machine learning: A systematic review of recent research // *Robotics and Computer-Integrated Manufacturing*. 2023. vol. 79. pp. 1–16.
19. Sandakalum T., Ang M.H.Jr. Motion Planning for Mobile Manipulators – A Systematic Review // *Machines*. 2022. vol. 10. 97 p. DOI: 10.3390/machines10020097.
20. Chen H., Zang X., Zhu Y., Zhao J. Hybrid Sampling-Based Path Planning for Mobile Manipulators Performing Pick and Place Tasks in Narrow Spaces // *Applied Sciences*. 2024. vol. 14. no. 22. 10313 p. DOI: 10.3390/app142210313.
21. Serebrenny V., Shmalko E. Multicriterial analyses of Pareto-efficiency for collaborative multi-agent systems using genetic algorithm with variations // *Procedia Computer Science*. 2021. vol. 186. pp. 793–801. DOI: 10.1016/j.procs.2021.04.201.

22. Jafar F.A., Abdullah N., Muhammad M.N., Zakaria N.A., et al. Investigation of Human Emotional State in Human-Robot Collaboration // *Journal of Computers*. 2014. vol. 9. no. 3. pp. 668–677.
23. De Benedictis C., Franco W., Maffiodo D., Ferraresi C. Control of Force Impulse in Human–Machine Impact // *Mechanisms and Machine Science*. Springer. 2018. vol. 49. pp. 956–964. DOI: 10.1007/978-3-319-61276-8_102.
24. Hogan N. Impedance Control: An Approach to Manipulation: Part I Theory // *Journal of Dynamic Systems, Measurement, and Control*. 1985. 107. pp. 1–7.
25. Parker C.A.C., Croft E.A. Design Personalization of a Cooperative Carrying Robot Controller // *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. 2012. pp. 3916–3921.
26. Woods A.C., La H.M., Ha Q.P. A novel extended potential field controller for use on aerial robots // *Proceedings of the IEEE International Conference on Automation Science and Engineering (CASE)*. 2016. pp. 286–291.
27. Li S., Wang H., Zhang S. Human-Robot Collaborative Manipulation with the Suppression of Human-caused Disturbance // *Journal of Intelligent Robotic Systems*. 2021. vol. 102. 73 p. DOI: 10.1007/s10846-021-01429-8.
28. Haddadin S., Croft E. Physical Human-Robot Interaction // *Springer Handbook of Robotics*. Cham: Springer. 2016. pp. 1835–1874. DOI: 10.1007/978-3-319-32552-1_69.
29. Karamzin D., Pereira F.L. On Higher-Order State Constraints // *SIAM Journal on Control and Optimization*. 2023. vol. 61. no. 4. pp. 1913–1933. DOI: 10.1137/22M1487497.
30. Hartl R.F., Sethi S., Vickson R. A Survey of the Maximum Principles for Optimal Control Problems with State Constraints // *SIAM Review*. 1995. vol. 37. no. 2. pp. 181–218.
31. Karamzin D.Yu. [Some estimates for the jump of the derivative of the Lagrange multiplier function in optimal control problems with second-order state constraints]. *Izvestiya Irkutskogo gosudarstvennogo universiteta. Seriya Matematika – The Bulletin of Irkutsk State University. Series Mathematics*. 2024. vol. 49. pp. 3–15. (In Russ.).
32. Shmalko E.Yu. [Machine-synthesized control of a nonlinear dynamic object based on optimal location of equilibrium points]. *Informatika i avtomatizatsiya – Informatics and Automation*. 2023. vol. 22. no. 1. pp. 87–109. (In Russ.).
33. Diveev A., Shmalko E., Serebrenny V., Zentay P. Fundamentals of Synthesized Optimal Control // *Mathematics*. 2020. vol. 9. no. 1. 21 p.
34. Shmalko E. Feasibility of Synthesized Optimal Control Approach on Model of Robotic System with Uncertainties. In: Ronzhin A., Shishlakov V. (eds). *Electromechanics and Robotics. Smart Innovation, Systems and Technologies*. Springer. 2022. vol. 232. pp. 131–143. DOI: 10.1007/978-981-16-2814-6_12..
35. Sanchez-Sanchez C., Izzo D., Hennes D. Learning the optimal state-feedback using deep networks // *Proceedings of the IEEE Conference on Symposium Series on Computational Intelligence (SSCI)*. 2016. pp. 1–8.
36. Diveev A. Synthesis of Optimal Stabilization System by Supervised Machine Learning of Symbolic Regression // *Lecture Notes in Networks and Systems (LNNS)*. Springer. 2025. vol. 1424. pp. 258–275. DOI: 10.1007/978-3-031-92605-1_17.
37. Diveev A., Shmalko E. Machine Learning Control by Symbolic Regression // Cham: Springer. 2021. 155 p. DOI: 10.1007/978-3-030-83213-1.
38. Diveev A.I. [Numerical method of network operator for the synthesis of a control system with uncertain initial values]. *Izvestiya Rossiyskoy akademii nauk. Teoriya i sistemy upravleniya – Journal of Computer and Systems Sciences International*. 2012. no. 2. p. 63. (In Russ.).

39. Diveev A. Hybrid Evolutionary Algorithm for Optimal Control Problem // Lecture Notes in Networks and Systems (LNNS). Springer. 2023. vol. 543. pp. 726–738. DOI: 10.1007/978-3-031-16078-3_50.
40. Program module of the network operator // GitHub. Available at: https://github.com/KostyaYamshanov/Machine-Learning-Control-kit/tree/main/network_operator (In Russ.).

Schmalko Elizaveta — Ph.D., Dr. Sci., Associate professor, Senior researcher, Federal Research Center "Computer Science and Management" of the Russian Academy of Sciences (FIC ISU RAS). Research interests: numerical methods, evolutionary computation, optimal control, symbolic regression. The number of publications — 150. e.shmalko@gmail.com; 44, Vavilova St., 119333, Moscow, Russia; office phone: +7(964)-636-6669.

Diveev Askhat — Ph.D., Dr. Sci., Professor, Chief scientific officer, Federal Research Center "Computer Science and Management" of the Russian Academy of Sciences (FIC ISU RAS). Research interests: optimal control, symbolic regression, robotics. The number of publications — 450. aidiveev@mail.ru; 44, Vavilova St., 119333, Moscow, Russia; office phone: +7(905)711-4427.

Yamshanov Konstantin — Graduate student, Federal Research Center "Computer Science and Management" of the Russian Academy of Sciences (FIC ISU RAS). Research interests: robotics, machine learning, management. The number of publications — 5. yam.kostya@yandex.ru; 44, Vavilova St., 119333, Moscow, Russia; office phone: +7(923)249-8191.

Gromov Ivan — Graduate student, Federal Research Center "Computer Science and Management" of the Russian Academy of Sciences (FIC ISU RAS); Engineer, Moscow State Technological University "STANKIN"(MSTU "STANKIN"). Research interests: robotics, control, machine learning. The number of publications — 5. 8357743@gmail.com; 44, Vavilova St., 119333, Moscow, Russia; office phone: +7(905)748-5240.

Acknowledgements. The work was carried out with the financial support of the Ministry of Science and Higher Education of the Russian Federation (project FSFS-2024-0012).