

УДК 681.51

doi: 10.18698/0536-1044-2024-01-41-56

Методы искусственного интеллекта для задач управления робототехническими и мехатронными системами: обзор

Ю.С. Зайцева

Институт проблем машиноведения РАН; Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»

Methods of artificial intelligence in tasks of managing the robotic and mechatronic systems: review

Y.S. Zaytseva

Institute of Mechanical Engineering Problems of the Russian Academy of Sciences; Saint Petersburg Electrotechnical University "LETI"

Развитие робототехники и мехатроники ставит перед инженерами новые задачи, привлекая для их решения методы искусственного интеллекта. Закономерно, что сложные задачи управления требуют современных решений. Дан обзор последних разработок по управлению мехатронными системами. Показана связь методов классической теории автоматического управления и машинного обучения. Кратко описаны такие известные методы классической теории управления, как оптимизация, адаптация и нечеткая логика, на основе которых построены искусственные нейронные сети и обучение с подкреплением. Рассмотрены последние достижения по применению интеллектуального управления для актуальных задач в различных областях техники. Анализ литературы показал, что будущие исследования направлены на все большую степень автоматизации и автономии объектов управления, а их свойства и характер функционирования должны приблизиться к человеческим чертам интеллекта.

Ключевые слова: алгоритм обучения, нейронные сети, многокритериальная оптимизация, адаптивное управление, интеллектуальное управление

Development of robotics and mechatronics creates new problems for engineers and attracts the artificial intelligence methods to solve them. It is natural that the complex management problems require modern solutions. The paper presents an overview of the latest developments in managing the mechatronic systems. It demonstrates connection between methods of the classical theory of automatic control and the machine learning. The well-known methods of the classical management theory are briefly described. They include optimization, adaptation and fuzzy logic, and form the basis for artificial neural networks and reinforcement learning. The latest achievements in the intelligent management application for solving the current problems in various technology areas are presented. Literature analysis shows that future research is aimed at the increasing degree in automation and autonomy of the control objects, while their functional properties and nature should approach the human intelligence.

Keywords: learning algorithm, neural networks, multicriteria optimization, adaptive management, intelligent management

С точки зрения автоматизации стратегию интеллектуального управления можно трактовать достаточно гибко. Эта концепция отличается от классического управления с обратной связью тем, что не требует математического описания системы и может быть реализована в режиме реального времени. Так, классическое адаптивное управление можно рассматривать как интеллектуальное, если закон управления системой постоянно обновляется. В соответствии с классификацией такие системы относят к пограничным [1].

При анализе литературы, посвященной интеллектуальному управлению мехатронными системами, хорошо прослеживается развитие методов, берущих начало из вычислительной математики, переходящих в область программирования, обработки и анализа больших массивов данных и активно используемых для управления физическими инженерными системами.

В робототехнике и мехатронике применяют различные подходы и направления искусственного интеллекта: машинное обучение, искусственные нейронные сети (ИНС), нечеткую логику, эволюционные вычисления и генетические алгоритмы. В состав машинного обучения входят такие методы, как классическое обучение (с учителем и без учителя), ансамблевые методы, обучение с подкреплением (ОсП), ИНС и глубокое обучение.

Алгоритмы интеллектуального управления используют различные оптимизационные, статистические и вероятностные методы для извлечения полезных закономерностей из проведенных действий и анализа больших наборов данных [2]. Если классическое машинное обучение и ансамбли в основном применяют для обработки и классификации наборов данных, то ИНС и ОсП — для автоматического управления, планирования и логистики [3–5]. Причем наибольшая эффективность ИНС наблюдается при работе со сложными данными и невыделенными признаками.

Цель работы — систематизация методов искусственного интеллекта для решения задач управления, выявление тенденций развития науки в этом направлении и нерешенных вопросов.

Методы интеллектуального управления. Кратко рассмотрим наиболее используемые методы управления, относящиеся к интеллектуальным.

Методы адаптивного управления. Адаптивное управление, как оптимальное, имеет хорошо разработанную теоретическую базу и строгое математическое обоснование [6]. В рамках классической теории автоматического управления этот метод стал отправной точкой к интеллектуальному управлению, так как синтез адаптивных регуляторов обеспечивает заданное качество работы системы при изменении ее параметров и влиянии внешней среды.

Принцип адаптации прошел эволюцию от самонастраивающихся регуляторов до адаптивного обучения [7]. Однако, как показано в работе [8], наличие теории не гарантирует ее успешное применение на практике.

Методы оптимизации. Теория оптимальности легла в основу большинства задач машинного обучения, которые в общем виде можно сформулировать как минимизацию некоторого функционала по параметру [9]. Вид минимизируемой величины зависит от метода машинного обучения. Так, при решении задачи классификации или регрессии минимизируют ошибку предсказания на имеющейся выборке; в теории ОсП находят максимальную выгоду от действий агента. Этой цели можно достичь с помощью какого-либо поискового алгоритма. Таким образом, у математической оптимизации появилось множество видов, методов и приложений.

Искусственные нейронные сети. История развития ИНС переживала взлеты и падения, пока к началу 90-х годов не сформировался их теоретический фундамент. Структура нейрона описывается функцией $R^n \rightarrow R$ [10]. Взвешенная сумма нейронов находится как

$$net = \sum_{i=1}^n w_i x_i = \mathbf{w}^T \mathbf{x},$$

где i — номер нейрона; n — количество нейронов; w_i — вес i -го нейрона; x_i — i -й вход; \mathbf{w}^T — матрица весов; \mathbf{x} — вектор входа.

Далее эта взвешенная сумма проходит через функцию активации, которая преобразует ее в число, подаваемое на выход нейрона.

Задача обучения сводится к определению весов, при которых ошибка обучения стремится к минимуму для всех возможных обучающих выборок. Процесс обучения нейронной сети — это поиск набора весов, преобразующих входной сигнал в нужный выходной. Так как функция ошибки обучения может иметь произвольный вид, обучение ИНС в общем случае описы-

вается как многоэкстремальная невыпуклая задача оптимизации.

Применение ИНС эффективно для управления в нелинейных и зашумленных системах, обеспечивает адаптивность, а после предварительного обучения ИНС может работать в режиме реального времени. Непрерывное развитие структур и свойств ИНС направлено на преодоление имеющихся недостатков. Так, эвристические подходы могут привести к неоднозначным решениям и тупиковым ситуациям, требуют подготовки обучающей выборки.

Недостатки ИНС заключаются в продолжительных временных затратах на выполнение обучения, а однозначно непредсказуемый результат обучения увеличивает риск для управления дорогостоящими объектами. Также реализация ИНС на данный момент возможна в виде сверхбольших интегральных схем.

Методы нечеткой логики. Аппарат нечеткой логики, предложенный Л. Заде [11, 12], базируется на понятии нечеткого множества как объекта с функцией принадлежности элемента ко множеству, принимающей значения в интервале $[0, 1]$. Нечеткие логические рассуждения оказались возможным представить в виде нейронной сети. Для этого роль функции активации нейрона должна выполнять функция принадлежности, связи между нейронами — передача сигнала, а логические t-нормы и k-нормы представляют как специальные виды нейронов.

На сегодняшний день разработано много нейро-нечетких сетей, которые приближенно описываются универсальной формой аппроксиматоров [13]. Их характерной чертой является человекоподобный стиль рассуждений и коннекционистская структура.

Обучение с подкреплением. ОпС содержит понятия агента и среды. Агент в процессе взаимодействия с неопределенной динамической средой изучает так называемую оптимальную стратегию принятия решений, отображающую его состояние после выполнения оптимальных действий [14–16].

В работе [17] описан общий подход к разработке стратегии управления для различных роботов, сред и манипулятивных задач. Среда может быть как физической, так и виртуальной. Важным аспектом во взаимодействии агента и среды является режим реального времени [18]. Агент получает наблюдение реакции среды в режиме реального времени, когда предпринимаются случайные или неоптималь-

ные действия, и оценку этих действий в виде вознаграждения.

В литературе ОсП называют по-разному: приближенным динамическим программированием (approximate dynamic programming) [19], нейродинамическим программированием (neuro-dynamic programming) [20] и адаптивным критиком (adaptive critic) [21].

Вербос использовал широкий термин «приближенное динамическое программирование обучения с подкреплением» [22]. Этот термин в широком смысле включает в себя все доступные с вычислительной точки зрения инструменты для нахождения наиболее точного возможного решения уравнения Беллмана. С точки зрения теории автоматического управления ОсП можно рассматривать как адаптивное оптимальное управление, которое обеспечивает адаптивный регулятор, сходящийся к оптимальному [23].

Общая стохастическая система дискретного времени в задаче классического ОсП выражается в виде одношаговой марковской модели, в которой полностью наблюдаемая система меняет свое состояние $s_{t+1} = F(s_t, a_t, w_t)$ на основе действия a_t согласно стратегии π_θ (w_t — случайная величина). Цель обучения состоит в том, чтобы найти стратегию действий для достижения наилучшего результата, выражаемого через функцию награды.

Сформулированная задача марковского процесса принятия решений решается с помощью временной рекурсивной формы — оптимального уравнения Беллмана. Однако его точное решение нельзя получить для большинства практических задач из-за вычислительной сложности, связанной с огромными размерами пространства состояний и действий.

Поэтому для обучения инженерных систем без описания их модели нашли применение алгоритмы на основе оценочных функций, оценки градиента функции вознаграждения по параметрам стратегии и метаэвристики. Также иногда применяют алгоритмы, использующие модель среды, для предварительного обучения.

Применение интеллектуальных методов для решения различных задач управления. Задачи стабилизации и программного управления. Эти задачи подразумевают наличие обратной связи в контуре. К стандартным задачам мехатроники и робототехники относятся стабилизация скорости и отслеживание траектории. Скорость и

координаты достаточно легко поддаются изменению на выходе системы.

В работе [24] рассмотрено управление позицией квадрокоптера методом машинного обучения, в котором стратегия действий дифференцируема по параметрам θ и находится градиентным методом. Оценка градиента записывается в следующей форме:

$$\hat{g} = \hat{E}_t[\nabla_{\theta} \log_{\pi_{\theta}}(a_t | s_t) \hat{A}_t], \quad (1)$$

где \hat{E}_t — эмпирическое среднее по конечному набору выборок, полученное путем чередования процессов генерации выборок и улучшения стратегии действий; \hat{A}_t — оценка функции вознаграждения.

Для снижения дисперсии оценок градиента авторы разработали более надежный метод оптимизации путем монотонного повышения производительности агента. С этой целью предложена целевая функция, дифференцирование которой приводит к выражению (1):

$$J(\theta) = \hat{E}_t[\min(r(\theta)\hat{A}_t, \varepsilon)],$$

где $r(\theta)$ — функция оценки обновления стратегии; ε — гиперпараметр.

В работе [25] рассмотрено энтропийно-адаптивное безмодельное ОсП ходьбе мобильного робота-треноги, оснащенного вибрационным приводом с мягкой мембраной. Регулятор определяется как политика нормального распределения Гаусса и задается как ИНС.

Функция награды предложена как сумма некоторой заданной константы и среднеквадратической ошибки по положению и углу робота. Оптимальная стратегия действий для максимизации энтропии определяется как

$$\pi_{\alpha} = \arg \max_{\pi} E_{\tau \sim P, \pi} \left\{ \sum_{t=0}^{\infty} \gamma^t [\hat{r}(s_t, a_t) + \alpha H(\pi(\cdot | s_t))] \right\},$$

где $\hat{r}(s_t, a_t)$ — функция вознаграждения; α — температура энтропии;

$$H(\pi(\cdot | s_t)) = E_{\alpha \sim \pi} [-\log(\pi(a_t, s_t))].$$

Согласно адаптивному мягкому алгоритму актер-критик (actor-critic), необходимо минимизировать функционалы для каждой функции ценности методом стохастического градиентного спуска. Если цель управления меняется во времени, то описанный подход ОсП не применим.

Решение этой проблемы изложено в работе [26], где предложено управлять набором состо-

яние — действие — вознаграждение для обучения имитации постоянной цели в каждом наборе. Такой подход проверен на системе управления скоростью движения автомобиля с различными заданными значениями на модели и реальном транспортном средстве.

Интеллектуальное производство электроэнергии стремительно развивается, привлекая внимание исследователей. Классические методы становятся недостаточными в этой области из-за растущих масштабов систем и различных неопределенностей. Авторы статьи [27] видят альтернативу этим методам в управлении с использованием собранных данных.

Задачи управления, требующие учета стохастических внешних воздействий, тяжело решить классическими методами; они имеют сложные структуры, не всегда удовлетворяющие качеству работы [28, 29]. К таким устройствам относятся ветряные генераторы, преобразующие энергию ветра в механическую энергию.

В трудах [30, 31] исследована проблема управления шагом ветряной турбины, зависящим от нелинейной динамики ветроприемника, связи переменных и неопределенностей внешней среды. Для ее решения предложено улучшенное по скорости гибридное обучение, состоящее из пропорционально-интегрально-дифференциального (ПИД) регулятора и обучающегося наблюдателя. Такая структура управления оказалась эффективнее, чем нечеткий регулятор и ИНС, и позволила снизить погрешность выходной мощности на 41 %.

Состояние системы оценивают по ошибке мощности, скорости ветра и их производных. Функция вознаграждения определена в виде сигнум-функции, знак которой зависит от знака производной от ошибки мощности. Обучающийся наблюдатель представляет собой модуль стратегии управления, который получает на вход ошибку выходной мощности и политику действий. Он генерирует наилучшую политику и настраивает входные параметры. Модуль выбора действий выбирает случайное действие с вероятностью ε и действие с вероятностью $1-\varepsilon$ для исключения локальных минимумов.

ИНС используют для решения задач, в которых может присутствовать полная неопределенность как динамики агента, так и среды. Пример управления такой сложной нелинейной системой высокого порядка можно найти

в работе [32]. Здесь показано выполнение процесса лазерной сварки при одновременном применении ИНС для предсказания качества сварочного шва по обработанным видеоизображениям и зависящей от нее подаваемой мощности лазера. Актор-критик следует стохастической политике управления, выбранной из распределения Гаусса. За меру вознаграждения выбрана оценка ошибки по глубине шва.

В статье [33] предложена система управления автономным колесным мобильным роботом, состоящая из адаптивного нейросетевого контура управления скоростью движения и самообучающегося контура отслеживания траектории с учетом динамических ограничений. ИНС состоит из трех слоев, в которой каждый скрытый нейрон имеет гиперболическую тангенциальную функцию активации.

Много исследований с применением ИНС посвящено автоматизации пилотируемых и беспилотных авиационных систем за счет реализации адаптивных стратегий управления полетом [34, 35]. Особенно это актуально для автономной работы в сложных атмосферных условиях и системных отказах.

Методы машинного обучения позволяют повысить точность управления и реагировать на неизвестные нелинейные условия полета. Как показано в работе [35], ИНС в таком случае можно использовать для реализации обучаемой системы управления полетом. Большие группы нейронных элементов управления обучаются требуемому поведению на основе измеренных данных. Качество обучения можно оценить статистическими методами.

Благодаря способностям к обучению ИНС хорошо подходят для решения сложных задач нелинейного управления и отказоустойчивости.

В статье [36] предложен подход нелинейного адаптивного управления полетом летательного аппарата с линеаризацией обратной связи и использованием онлайн-обучения в скользящем режиме. Алгоритм обучения в скользящем режиме основан на теории управления переменной структурой и рассматривает ИНС и ее обучение как процесс управления. Такой подход позволяет преодолеть летательному аппарату негативное влияние нелинейностей в динамике и внешних возмущений.

В работе [37] рассмотрена программируемая логическая интегральная схема в скользящем режиме на основе рекуррентной вейвлет-нейронной сети для управления линейным уль-

тразвуковым двигателем. Для оценки нелинейных функций в модели системы, параметров системы и внешних возмущений использованы ИНС. Алгоритм адаптивного обучения ИНС в режиме реального времени получен с помощью теоремы устойчивости Ляпунова. Высокая эффективность алгоритма проверена для задач позиционирования и отслеживания.

Адаптивный скользящий режим на основе метода опорных векторов с нечеткими правилами для повышения быстродействия и точности управления двухзвенным роботом-манипулятором описан в статье [38]. Возможности этого метода позволяют хорошо обрабатывать нелинейные сигналы и имеют способность к обобщению, применяемую для уменьшения дрейфа при управлении в скользящем режиме.

В работе [39] синтез регулятора для существенно нелинейной системы выполнен с помощью спайковой ИНС, которая использует временные последовательности импульсов. Регуляторы на основе спайковых ИНС способны к онлайн-обучению и самоадаптации при переходе от моделирования к реальной работе. Такие структуры ИНС имеют меньше нейронов и нейронных связей по сравнению с предыдущими поколениями ИНС. Обучение спайковой ИНС в рамках проводилось с использованием эволюционных алгоритмов. Эффективность предложенного подхода проверена на модели движения роботизированной руки.

Настройка регулятора. Этот процесс не вызывает сложностей при условии имеющегося достоверного описания системы и доступности измерений. В некоторых случаях (особенно, когда система неустойчива) возможны только измерения по обратной связи, вследствие чего их настройка становится трудоемким и неэффективным процессом в меняющихся условиях работы системы. Поэтому прибегают к автоматизированным методам, которые дают возможность быстро определять параметры регулятора в соответствии с поставленной задачей без участия человека.

Методы адаптивного управления позволяют изменять как структуру, так и параметры регулятора в зависимости от изменения параметров объекта управления или внешних возмущений. Историческая перспектива методов адаптивного управления и обучения приведена в обзорной работе [7]. Несмотря на простоту, ПИД-регулятор по-прежнему используют в современных системах управления, выполняя

настройку его параметров с помощью методов машинного обучения.

Так, для настройки ПИД-регулятора системы управления статического компенсатора реактивной мощности, подключенного к системе преобразования энергии ветра, применен алгоритм безградиентного поиска [40].

В работе [41] использован алгоритм конкуренции колоний для настройки параметров ПИД-регулятора системы управления испарителя, при этом минимизирована абсолютная квадратичная ошибка слежения системы. Показана эффективность алгоритма по сравнению с методом Зиглера — Никольса и генетическим алгоритмом.

В статье [42] рассмотрена автоматизированная настройка пропорционально-интегрального (ПИ) регулятора для системы подачи топлива в автомобиле. Проведено сравнение двух методов настройки: безопасного активного обучения и байесовской оптимизации. Безопасность подразумевает ввод ограничений на такие параметры, как перерегулирование, которое может привести к потере устойчивости системы и аварийному режиму работы. Активное обучение предполагает минимизацию функции потерь, которая зависит только от параметров ПИ-регулятора, ограничивает большую ошибку управления и колебания.

Для обеспечения выборки безопасных для системы параметров авторы предложили комбинированный подход, заключающийся в использовании байесовской оптимизации с дискриминантной моделью. Отличие этого подхода от активного обучения состоит в ограничении нижнего доверительного предела оцениваемым риском. Оба подхода показали хорошие результаты и имеют дальнейшие перспективы улучшения.

Метод безопасной байесовской оптимизации также использован для настройки ПИ-регулятора системы управления квадраторами [43]. Предложен алгоритм вычисления оценки функционала с ограничениями на область безопасных параметров.

В публикации [44] на основе марковского процесса проведена настройка параметров двигателя внутреннего сгорания в режиме реального времени для достижения оптимального режима работы при любом стиле вождения.

Для моделирования и прогнозирования динамических характеристик релейно-реактивной машины предложено использовать адаптивную

нечеткую нейронную систему на основе радиальных базисных функций [45]. В исследуемой системе доступно косвенному измерению только фазовое потягосцепление, по которому рассчитывается фазовый крутящий момент. Обучение машины знанию электромагнитных характеристик проводится с помощью иерархического самоорганизующегося алгоритма.

Направление другой области искусственного интеллекта для синтеза и настройки ПИД-регулятора рассмотрено в трудах [46-49]. Разработана структура планирования траектории с помощью обучения с подкреплением нечеткого регулятора [49]. Модель динамики робота и ПИД-регулятора описана нечеткими правилами. ПИД-регулятор, построенный на основе метода нечеткой логики с Гауссовой функцией принадлежности, показал эффективность при управлении шаговым двигателем [46]. Нечеткая структура регулятора позволяет адаптировать его усиление к изменениям в системе.

Аналогичный подход использован в работе для активного магнитного подшипника, в котором дополнительно проведена многокритериальная оптимизация центров треугольных функций принадлежности и коэффициента усиления с помощью генетического алгоритма [48].

В работе [47] применена нечеткая модель Сугэно нулевого порядка, параметры которой настраиваются в автономном режиме с помощью алгоритма гибридного обучения, представляющего собой комбинацию методов наименьших квадратов и обратного распространения ошибки. Компьютерное моделирование на примере системы стабилизатора напряжения показало, что нечеткая модель хорошо справляется с устойчивостью системы при малых сигналах.

Задачи идентификации. В статье [50] предложен подход к идентификации линейной системы, описанной разностными уравнениями, с использованием алгоритма пчелиного роя. Задача идентификации сформулирована как оптимизация, целью которой является получение модели и набора параметров, для которых ошибка прогнозирования между измеренными выходными данными объекта и модели будет минимальной. Результаты работы алгоритма продемонстрированы на модели двигателя постоянного тока.

Тенденция к развитию интеллектуальных методов управления направлена на комбинацию различных методов. Так, для идентифика-

ции параметров робота использованы фильтр Кальмана, рекуррентная ИНС и метод градиентного спуска [51]. Динамика робота описана уравнениями Эйлера — Лагранжа. Декартовы силы приняты за выходы системы, а положение, скорость и ускорение — за компоненты вектора входного регрессора.

Для задачи обратной динамической идентификации рассмотрены три самообучающиеся ИНС прямого действия с многоуровневой структурой нейронов. Они включали в себя один скрытый слой и достаточное количество нелинейных нейронов и выходной слой с линейной функцией активации.

Задачи оптимизации. Важное развитие теории оптимальности для практических приложений отражено в работах Р.Б. Статникова, где описан метод многокритериальной оптимизации, согласно которому строится множество допустимых решений поставленной задачи [52, 53]. Метод, активно применяемый в автомобилестроении, проиллюстрирован на примере многокритериальной оптимизации кинематических схем автомобиля, его кинематических характеристик механизмов и устройств [53].

Благодаря универсальности алгоритмы безградиентного поиска получили широкое распространение для решения любых задач оптимизации, в том числе для ИНС [11, 54–56]. Их особенность состоит в том, что найденное решение задачи оптимизации является приемлемым, а не оптимальным.

В последнее время стали популярными различные биоинспирированные решения, идеи которых заимствованы у природы [11]. К ним относятся эволюционные алгоритмы, алгоритмы роя и колонии [54, 55], популяционные [56] и др. Известен также алгоритм летучей мыши, относящийся к роевому интеллекту, основанный на эхолокации.

В статье [57] предложен подход оптимального управления на основе алгоритма опорных векторов для минимизации энергопотребления при ходьбе двуногих роботов в условиях неизвестной динамической модели системы и небольших размеров выборки данных. Новый регулятор встроен в оптимальный регулятор и направлен на минимизацию связанной с энергией функции затрат с ограничениями на углы сочленений робота.

В работе [58] проведена многокритериальная оптимизация параметров ПИД-регулятора по улучшенному алгоритму поиска учитель —

ученик, который исключает попадания значений функции в локальный минимум. Для этого фаза ученика не ограничивалась двумя учениками и сочеталась с дополнительной фазой. Также была добавлена фаза заблокированного устройства, где блокируются параметры, близкие к несовместимым целям. Это приводит к тому, что каждая цель не может конфликтовать с другими.

В публикации [59] проведен сравнительный анализ между результатами оптимизации параметров ПИД-регулятора системы управления двигателем постоянного тока, полученными с помощью алгоритма пчелиной колонии, метода роя частиц и учитель — ученик, последний из которых показал лучшие результаты.

В работе [60] приведены результаты большого исследования по самооптимизирующимся измерительным инструментам, предназначенным для разработки и проверки инновационного производственного оборудования и систем управления, основанных на понимании, оценке и контроле производственных характеристик крупных станков с помощью самообучающегося регулятора, способного управлять большим количеством измеренных датчиками данных. Предлагаемая архитектура основана на методе опорных векторов и аппарате нечеткой логики для сравнения собранных данных с фактическим текущим состоянием процесса.

Задача повторяющегося управления. Повторяющееся управление, изначально сформулированное для задачи преодоления повторяющихся возмущений с известным периодом, характерных при вращении двигателей [61], получило дальнейшее развитие для случаев, когда период возмущений неизвестен или меняется во времени.

Задача повторяющегося управления сформулирована следующим образом. Рассматривается асимптотически устойчивая система дискретного времени с периодическим возмущением. Целью закона управления является асимптотическое стремление ошибки управления к нулю. Регулятор можно синтезировать разными методами. Наиболее популярным и изученным среди них является метод на основе принципа внутренней модели, который состоит в построении независимых блоков идентификации параметров гармоник [62].

В работе [63] описаны некоторые фундаментальные проблемы разработки регуляторов на основе повторяющегося управления для обра-

ботки периодических возмущений, возникающих в электромеханических системах. Для преодоления параметрической неопределенности возмущений предложено использовать инструменты адаптивного управления, μ -синтеза, Q-параметризации и антивиндап компенсации (anti-windup).

В процессе выполнения итераций система способна обучаться, внося на каждом шаге коррективы, что позволяет быстрее и точнее достигать желаемых результатов [64-66]. Примерами таких систем являются роботы в промышленных конвейерных системах, наносистемы, химические периодические процессы, испытательные установки надежности и др.

Итеративное обучение в общем виде можно рассматривать как систему из линейных уравнений состояния, одно из которых описывает динамику системы во времени, второе — динамику итерационного обучения [67, 68]:

$$\begin{aligned} \mathbf{x}(t+1, k) &= \mathbf{A}\mathbf{x}(t, k) + \mathbf{B}\mathbf{u}(t, k); \\ \mathbf{y}(t, k) &= \mathbf{C}\mathbf{x}(t, k), \end{aligned}$$

где \mathbf{x} , \mathbf{u} и \mathbf{y} — вектор состояния, управления и выхода соответственно; t — дискретное время; k — номер шага обучения; \mathbf{A} , \mathbf{B} , \mathbf{C} — матрицы.

Алгоритм итерационного обучения состоит в том, что на каждом шаге измеряемые переменные системы записываются в память программы, затем обрабатываются для достижения заданного качества критерия управления. Цель управления заключается в достижении вектором выходных параметров \mathbf{y} заданного значения с установленной точностью.

Закон итеративного обучающего управления с коррекцией на каждом шаге имеет вид

$$\mathbf{u}(t, k) = \mathbf{u}(t, k-1) + \Delta\mathbf{u}(t, k-1),$$

где $\Delta\mathbf{u}$ — корректирующая добавка к входной переменной системы, выбираемая из условия, что ошибка по выходу на каждом шаге стремится к нулю.

В статье [69] предложен подход к итеративному обучающему управлению нейротеозом, при котором формируется закон изменения интенсивности электростимуляции в каждый момент времени. Угол наклона стопы оценивают в режиме реального времени с помощью гироскопов и акселерометров. В начале каждого следующего шага человека определяют рассогласование угла наклона стопы как поэлементное отклонение между измеренной

угловой траекторией и первыми несколькими отсчетами соответствующей эталонной траектории.

В работе [70] рассмотрен метод отказоустойчивого управления на основе итеративного управления обучением контура управления током для восстановления производительности многофазных приводов с постоянными магнитами в случае обрыва цепи. Преимущество такого метода перед другими заключается в отсутствии необходимости проводить обнаружение и диагностику неисправностей (достаточно выполнить измерения крутящего момента). Таким образом, итеративное управление обучением дает исчерпывающую информацию о системе, обеспечивая устойчивость к неопределенностям модели.

Для движения системы облачных колесных роботов по заданной траектории с одновременным транспортированием груза разработана гибкая схема управления отслеживанием траектории посредством итеративного управления обучением, и проведен анализ устойчивости системы [71].

В работе [72] рассмотрено управление итеративным обучением под руководством человека для задач слежения за траекторией, когда регулятор получает данные от действий человека-оператора.

Выводы

1. Тенденция перехода к высокотехнологичным видам организации производства отражена в научных публикациях по применению интеллектуальных методов в электромеханических системах. Методы искусственного интеллекта позволяют достичь цели и решить задачи управления мехатронными системами, которые раньше вызывали затруднения, при этом повысилась вычислительная эффективность и простота реализации. Сложность задач управления существенно нелинейными, неопределенными или многоагентными системами в условиях воздействия внешней среды требует индивидуального подхода к решению конкретной задачи, для чего на данный момент уже предложено большое количество инструментов.

2. Эффективность этих алгоритмов для обучения сложных систем устанавливается, как правило, только при проведении практического эксперимента. Развитие алгоритмов направле-

но не только на повышение скорости и точности обучения, но и на все большую независимость от жестко заданной человеком стратегии обучения и ее адаптации к различным целям. Разработчики заимствуют идеи из природы, пытаются воспроизвести поведение живых организмов в алгоритмах.

3. Будущие исследования ставят перед собой задачу, которая в международных публикациях получила название «обучение для обучения» [73], когда агенту приходится самому подбирать стратегию обучения и регулировать ее метопараметры.

Литература

- [1] Vepa R. A review of techniques for machine learning of real-time control strategies. *Intell. Syst. Eng.*, 1993, vol. 2, no. 2, pp. 77–90, doi: <https://doi.org/10.1049/ise.1993.0009>
- [2] Mitchell T. *Machine learning*. McGraw-Hill, 1997.
- [3] Mosheiov G. Scheduling problems with a learning effect. *Eur. J. Oper. Res.*, 2001, vol. 132, no. 3, pp. 687–693, doi: [https://doi.org/10.1016/S0377-2217\(00\)00175-2](https://doi.org/10.1016/S0377-2217(00)00175-2)
- [4] Gurel S., Akturk M. Scheduling preventive maintenance on a single CNC machine. *Int. J. Prod. Res.*, 2008, vol. 46, no. 24, pp. 6797–6821, <https://doi.org/10.1080/00207540701487833>
- [5] Dittrich M.-A., Uhlich F., Denkena B. Self-optimizing tool path generation for 5-axis machining processes. *CIRP J. Manuf. Sci. Technol.*, 2019, vol. 24, pp. 49–54, doi: <https://doi.org/10.1016/j.cirpj.2018.11.005>
- [6] Фомин В.Н., Фрадков А.Л., Якубович В.А. *Адаптивное управление динамическими объектами*. Москва, Наука, 1981. 447 с.
- [7] Annaswamy A.M., Fradkov A.L. A historical perspective of adaptive control and learning. *Annu. Rev. Control*, 2021, vol. 52, pp. 18–41, doi: <https://doi.org/10.1016/j.arcontrol.2021.10.014>
- [8] Перельман И.И. Анализ современных методов адаптивного управления с позиций приложения к автоматизации технологических процессов. *Автоматика и телемеханика*, 1991, № 7, с. 3–32.
- [9] Андриевский Б.Р., Фрадков А.Л. Метод скоростного градиента и его приложения. *Автоматика и телемеханика*, 2021, № 9, с. 3–72, doi: <https://doi.org/10.31857/S0005231021090014>
- [10] Поляк Б.Т. *Введение в оптимизацию*. Москва, Наука, 1983. 384 с.
- [11] Карпенко А.П. *Современные алгоритмы поисковой оптимизации*. Москва, Изд-во МГТУ им. Н.Э. Баумана, 2014. 446 с.
- [12] Николенко С., Кадурин А., Архангельская Е. *Глубокое обучение. Погружение в мир нейронных сетей*. Санкт-Петербург, Питер, 2021. 476 с.
- [13] Заде Л. *Понятие лингвистической переменной и его применение к принятию приближенных решений*. Москва, Мир, 1976. 165 с.
- [14] Новак В. *Математические принципы нечеткой логики*. Москва, Физматлит, 2006. 347 с.
- [15] Cybenko G. Approximation by superpositions of a sigmoidal function. *Math. Control Signal. Systems*, 1989, vol. 2, no. 4, pp. 303–314, doi: <https://doi.org/10.1007/BF02551274>
- [16] Lee J., Shin J., Realf M. Machine learning: overview of the recent progresses and implications for the process systems engineering field. *Comput. Chem. Eng.*, 2018, vol. 114, pp. 111–121, doi: <https://doi.org/10.1016/j.compchemeng.2017.10.008>
- [17] Sutton R.S., Barto A.G. *Reinforcement learning*. MIT Press, 2018. 552 p.
- [18] Buoni L., Babuska R., De Schutter B. et al. *Reinforcement learning and dynamic programming using function approximators*. CRC Press, 2010. 280 p.
- [19] Kurrek P., Jocas M., Zoghalmi F. et al. AI motion control — a generic approach to develop control policies for robotic manipulation tasks. *Proc. Int. Conf. Engineering Design*, 2019, vol. 1, no. 1, pp. 3561–3570, doi: <https://doi.org/10.1017/dsi.2019.363>
- [20] Menghal P., Laxmi A. Real time simulation: a novel approach in engineering education. *ICECT*, 2011, vol. 1, pp. 215–219, doi: <https://doi.org/10.1109/ICECTECH.2011.5941592>
- [21] Powell W. *Approximate dynamic programming*. Wiley, 2011. 656 p.

- [22] Bertsekas D., Tsitsiklis J. Neuro-dynamic programming: an overview. *Proc. 34th IEEE Conf. on Decision and Control*, 1995, vol. 1, pp. 560–564, doi: <https://doi.org/10.1109/CDC.1995.478953>
- [23] Prokhorov D., Wunsch D. Adaptive critic designs. *IEEE Trans. Neural Netw.*, 1997, vol. 8, no. 5, pp. 997–1007, doi: <https://doi.org/10.1109/72.623201>
- [24] Werbos P.J. Reinforcement learning and approximate dynamic programming (RLADP) — foundations, common misconceptions, and the challenges ahead. In: *Reinforcement learning and approximate dynamic programming for feedback control*. Wiley, 2012, pp. 1–30.
- [25] Lewis F., Vrable D. Reinforcement learning and adaptive dynamic programming for feedback control. *IEEE Circuits Syst. Mag.*, 2009, vol. 9, no. 3, pp. 32–50, doi: <https://doi.org/10.1109/MCAS.2009.933854>
- [26] Lopes G., Ferreira M., Da Silva Simoes A. et al. Intelligent control of a quadrotor with proximal policy optimization reinforcement learning. *Proc. LARS/SBR/WRE*, 2018, pp. 503–508, doi: <https://doi.org/10.1109/LARS/SBR/WRE.2018.00094>
- [27] Kim J., Hong M., Lee K. et al. Learning to walk a tripod mobile robot using nonlinear soft vibration actuators with entropy adaptive reinforcement learning. *IEEE Robot. Autom. Lett.*, 2020, vol. 5, no. 2, pp. 2317–2324, doi: <https://doi.org/10.1109/LRA.2020.2970945>
- [28] Puccetti L., Kopf F., Rathgeber C. et al. Speed tracking control using online reinforcement learning in a real car. *ICCAR*, 2020, pp. 392–399, doi: <https://doi.org/10.1109/ICCAR49639.2020.9108051>
- [29] Sun L., You F. Machine learning and data-driven techniques for the control of smart power generation systems: an uncertainty handling perspective. *Engineering*, 2021, vol. 7, no. 9, pp. 1239–1247, doi: <https://doi.org/10.1016/j.eng.2021.04.020>
- [30] Tomin N., Kurbatsky V., Guliyev H. Intelligent control of a wind turbine based on reinforcement learning. *ELMA*, 2019, doi: <https://doi.org/10.1109/ELMA.2019.8771645>
- [31] Sierra-Garcia J.E., Santos M., Pandit R. Wind turbine pitch reinforcement learning control improved by PID regulator and learning observer. *Eng. Appl. Artif. Intell.*, 2022, vol. 111, art. 104769, doi: <https://doi.org/10.1016/j.engappai.2022.104769>
- [32] Gunther J., Pilarski P., Helfrich G. et al. Intelligent laser welding through representation, prediction, and control learning: an architecture with deep neural networks and reinforcement learning. *Mechatronics*, 2016, vol. 34, pp. 1–11, doi: <https://doi.org/10.1016/j.mechatronics.2015.09.004>
- [33] Lin W.-S., Yang P.-C. Adaptive critic motion control design of autonomous wheeled mobile robot by dual heuristic programming. *Automatica*, 2008, vol. 44, no. 11, pp. 2716–2723, doi: <https://doi.org/10.1016/j.automatica.2008.03.029>
- [34] Calise A., Rysdyk R. Nonlinear adaptive flight control using neural networks. *IEEE Control Syst. Mag.*, 1998, vol. 18, no. 6, pp. 14–25, doi: <https://doi.org/10.1109/37.736008>
- [35] Kruger T., Schnetter P., Placzek R. et al. Fault-tolerant nonlinear adaptive flight control using sliding mode online learning. *Neural Netw.*, 2012, vol. 32, pp. 267–274, doi: <https://doi.org/10.1016/j.neunet.2012.02.025>
- [36] Lin F.-J., Hung Y.-C., Chen S.-Y. Field-programmable gate array-based intelligent dynamic sliding-mode control using recurrent wavelet neural network for linear ultrasonic motor. *IET Control. Theory Appl.*, 2010, vol. 4, no. 9, pp. 1511–1532, doi: <https://doi.org/10.1049/iet-cta.2009.0066>
- [37] Zhu D., Mei T., Luo M. Adaptive sliding mode control for robots based on fuzzy support vector machines. *ICMA*, 2009, pp. 3469–3474, doi: <https://doi.org/10.1109/ICMA.2009.5246264>
- [38] Perez J., Cabrera J., Castillo J. et al. Bio-inspired spiking neural network for nonlinear systems control. *Neural Netw.*, 2018, vol. 104, pp. 15–25, doi: <https://doi.org/10.1016/j.neunet.2018.04.002>
- [39] Perez-Pena F., Lenero-Bardallo J., Linares-Barranco A. et al. Towards bioinspired close-loop local motor control: a simulated approach supporting neuromorphic implementations. *IEEE ISCAS*, 2017, doi: <https://doi.org/10.1109/ISCAS.2017.8050808>
- [40] Keshta H., Ali A., Saied E. et al. Application of Static Var Compensator (SVC) with PI controller for grid integration of wind farm using harmony search. *Int. J. Emerg. Electr. Power Syst.*, 2016, vol. 17, no. 5, pp. 555–566, doi: <https://doi.org/10.1515/ijeeps-2016-0159>

- [41] Atashpaz Gargari E., Hashemzadeh F., Rajabioun R. et al. Colonial competitive algorithm: a novel approach for PID controller design in MIMO distillation column process. *Int. J. Intell. Comput. Cybern.*, 2008, vol. 1, no. 3, pp. 337–355, doi: <https://doi.org/10.1108/17563780810893446>
- [42] Schillinger M., Hartmann B., Skalecki P. et al. Safe active learning and safe Bayesian optimization for tuning a PI-controller. *IFAC-PapersOnLine*, 2017, vol. 50, no. 1, pp. 5967–5972, doi: <https://doi.org/10.1016/j.ifacol.2017.08.1258>
- [43] Berkenkamp F., Schoellig A., Krause A. Safe controller optimization for quadrotors with Gaussian processes. *IEEE ICRA*, 2016, pp. 491–496, doi: <https://doi.org/10.1109/ICRA.2016.7487170>
- [44] Malikopoulos A., Papalambros P., Assanis D. A learning algorithm for optimal internal combustion engine calibration in real time. *Proc. DETC2007*, 2008, vol. 6-A, pp. 91–100, doi: <https://doi.org/10.1115/DETC2007-34718>
- [45] Liang D., Ding W. Modelling and predicting of a switched reluctance motor drive using radial basis function network-based adaptive fuzzy system. *IET Electr. Power Appl.*, 2009, vol. 3, no. 3, pp. 218–230, doi: <https://doi.org/10.1049/iet-epa.2008.0096>
- [46] Elsodany N., Rezeka S., Maharem N. Adaptive PID control of a stepper motor driving a flexible rotor. *Alex. Eng. J.*, 2011, vol. 50, no. 2, pp. 127–136, doi: <https://doi.org/10.1016/j.aej.2010.08.002>
- [47] Mitra P., Maulik S., Chowdhury S. et al. Anfis based automatic voltage regulator with hybrid learning algorithm. *Proc. Universities Power Engineering Conf.*, 2007, pp. 397–401, doi: <https://doi.org/10.1109/UPEC.2007.4468980>
- [48] Chen H.-C. Optimal fuzzy pid controller design of an active magnetic bearing system based on adaptive genetic algorithms. *Proc. ICMLC*, 2008, vol. 4, pp. 2054–2060, doi: <https://doi.org/10.1109/ICMLC.2008.4620744>
- [49] Wen S., Hu X., Lv X. et al. Q-learning trajectory planning based on Takagi–Sugeno fuzzy parallel distributed compensation structure of humanoid manipulator. *Int. J. Adv. Robot. Syst.*, 2019, vol. 16, no. 1, doi: <https://doi.org/10.1177/1729881419830204>
- [50] Ercin O., Coban R. Identification of linear dynamic systems using the artificial bee colony algorithm. *Turk. J. Elec. Eng. & Comp. Sci.*, 2012, vol. 20, no. 7, pp. 1175–1188, doi: <https://doi.org/10.3906/elk-1012-956>
- [51] Agand P., Shoorehdeli M.A., Khaki-Sedigh A. Adaptive recurrent neural network with Lyapunov stability learning rules for robot dynamic terms identification. *Eng. Appl. Artif. Intell.*, 2017, vol. 65, pp. 1–11, doi: <https://doi.org/10.1016/j.engappai.2017.07.009>
- [52] Соболев И.М., Статников Р.Б. *Выбор оптимальных параметров в задачах со многими критериями*. Москва, Дрофа, 2006. 175 с.
- [53] Черных В.В., Статников Р.Б., Храпов Д.С. и др. Метод исследования пространства параметров в задачах проектирования легковых автомобилей. *Проблемы машиностроения и надежности машин*, 2009, № 4, с. 19–25.
- [54] Balochian S., Baloochian H. Social mimic optimization algorithm and engineering applications. *Expert Syst. Appl.*, 2019, vol. 134, pp. 178–191, doi: <https://doi.org/10.1016/j.eswa.2019.05.035>
- [55] Kumar S., Kumar A., Shankar G. Crow search algorithm based optimal dynamic performance control of SVC assisted SMIB system. *NPSC*, 2018, doi: <https://doi.org/10.1109/NPSC.2018.8771814>
- [56] Abualigah L. Multi-verse optimizer algorithm: a comprehensive survey of its results, variants, and applications. *Neural Comput. & Applic.*, 2020, vol. 32, no. 16, pp. 12381–12401, doi: <https://doi.org/10.1007/s00521-020-04839-1>
- [57] Wang L., Liu Z., Chen C. et al. Support vector machine based optimal control for minimizing energy consumption of biped walking motions. *Int. J. Precis. Eng. Manuf.*, 2012, vol. 13, no. 11, pp. 1975–1981, doi: <https://doi.org/10.1007/s12541-012-0260-7>
- [58] Xiao L., Zhu Q., Li C. et al. Application of modified teaching-learning algorithm in coordination optimization of TCSC and SVC. In: *CCPR*. Springer, 2014, pp. 44–53, doi: https://doi.org/10.1007/978-3-662-45646-0_5
- [59] Shouran M., Habil M. Tuning of PID controller using different optimization algorithms for industrial DC motor. *ICACITE*, 2021, pp. 756–759, doi: <https://doi.org/10.1109/ICACITE51222.2021.9404616>

- [60] Silvestri M., Pedrazzoli P., Boer C. et al. Compensating high precision positioning machine tools by a self learning capable controller. *Proc. EUSPEN*, 2011, vol. 2, pp. 121–124.
- [61] Tomizuka M. Zero phase error tracking algorithm for digital control. *J. Dyn. Sys., Meas., Control.*, 1987, vol. 109, no. 1, pp. 65–68, doi: <https://doi.org/10.1115/1.3143822>
- [62] Tomizuka M. Dealing with periodic disturbances in controls of mechanical systems. *Annu. Rev. Control*, 2008, vol. 32, no. 2, pp. 193–199, doi: <https://doi.org/10.1016/j.arcontrol.2008.07.002>
- [63] Uchiyama M. Formation of high-speed motion pattern of a mechanical arm by trial. *Transactions of the Society of Instrument and Control Engineers*, 1978, vol. 14, no. 6, pp. 706–712, doi: <https://doi.org/10.9746/sicetr1965.14.706>
- [64] Li W., Maier P., Enge H. Self-learning control applied to vibration control of a rotating spindle by piezopusher bearings. *Proc. Inst. Mech. Eng., I: J. Syst. Control Eng.*, 2004, vol. 218, no. 3, pp. 185–196, doi: <https://doi.org/10.1243/095965104773913867>
- [65] Hladowski L., Galkowski K., Rogers E. Further results on dynamic iterative learning control law design using repetitive process stability theory. *10th nDS*, 2017, doi: <https://doi.org/10.1109/NDS.2017.8070621>
- [66] Kurek J., Zaremba M. Iterative learning control synthesis based on 2-D system theory. *IEEE Trans. Automat. Contr.*, 1993, vol. 38, no. 1, pp. 121–125, doi: <https://doi.org/10.1109/9.186321>
- [67] Pakshin P., Emelianova J., Emelianov M. Iterative learning control design for multiagent systems based on 2D models. *Autom. Remote Control*, 2018, vol. 79, no. 6, pp. 1040–1056, <https://doi.org/10.1134/S000511791806005X>
- [68] Seel T., Werner C., Raisch J. et al. Iterative learning control of a drop foot neuro-prosthesis — generating physiological foot motion in paretic gait by automatic feedback control. *Control. Eng. Pract.*, 2016, vol. 48, pp. 87–97, doi: <https://doi.org/10.1016/j.conengprac.2015.11.007>
- [69] Seel T., Werner C., Schauer T. The adaptive drop foot stimulator — multivariable learning control of foot pitch and roll motion in paretic gait. *Med. Eng. Phys.*, 2016, vol. 38, no. 11, pp. 1205–1213, doi: <https://doi.org/10.1016/j.medengphy.2016.06.009>
- [70] Mohammadpour A., Mishra S., Parsa L. Iterative learning control for fault-tolerance in multi-phase permanent-magnet machines. *Proc. American Control Conf.*, 2013, pp. 5929–5934, doi: <https://doi.org/10.1109/acc.2013.6580768>
- [71] Li J., Wang S., Wang J. et al. Iterative learning control for a distributed cloud robot with payload delivery. *Assem. Autom.*, 2021, vol. 41, no. 3, pp. 263–273, doi: <https://doi.org/10.1108/AA-11-2020-0179>
- [72] Warriar R., Devasia S. Iterative learning from novice human demonstrations for output tracking. *Trans. Hum. Mach. Syst.*, 2016, vol. 46, no. 4, pp. 510–521, doi: <https://doi.org/10.1109/THMS.2016.2545243>
- [73] Lansdell B., Kording K. Towards learning-to-learn. *Curr. Opin. Behav. Sci.*, 2019, vol. 29, pp. 45–50, doi: <https://doi.org/10.1016/j.cobeha.2019.04.005>

References

- [1] Vepa R. A review of techniques for machine learning of real-time control strategies. *Intell. Syst. Eng.*, 1993, vol. 2, no. 2, pp. 77–90, doi: <https://doi.org/10.1049/ise.1993.0009>
- [2] Mitchell T. *Machine learning*. McGraw-Hill, 1997.
- [3] Mosheiov G. Scheduling problems with a learning effect. *Eur. J. Oper. Res.*, 2001, vol. 132, no. 3, pp. 687–693, doi: [https://doi.org/10.1016/S0377-2217\(00\)00175-2](https://doi.org/10.1016/S0377-2217(00)00175-2)
- [4] Gurel S., Akturk M. Scheduling preventive maintenance on a single CNC machine. *Int. J. Prod. Res.*, 2008, vol. 46, no. 24, pp. 6797–6821, <https://doi.org/10.1080/00207540701487833>
- [5] Dittrich M.-A., Uhlich F., Denkena B. Self-optimizing tool path generation for 5-axis machining processes. *CIRP J. Manuf. Sci. Technol.*, 2019, vol. 24, pp. 49–54, doi: <https://doi.org/10.1016/j.cirpj.2018.11.005>
- [6] Fomin V.N., Fradkov A.L., Yakubovich V.A. *Adaptivnoe upravlenie dinamicheskimi obektami* [Adaptive control on dynamics objects]. Moscow, Nauka Publ., 1981. 447 p. (In Russ.).

- [7] Annaswamy A.M., Fradkov A.L. A historical perspective of adaptive control and learning. *Annu. Rev. Control*, 2021, vol. 52, pp. 18–41, doi: <https://doi.org/10.1016/j.arcontrol.2021.10.014>
- [8] Perelman I.I. Analysis of current adaptive control methods from the standpoint of the application to the automation of technological processes. *Avtomatika i telemekhanika*, 1991, no. 7, pp. 3–32. (In Russ.). (Eng. version: *Autom. Remote Control*, 1991, vol. 52, no. 7, pp. 891–911).
- [9] Andrievskii B.R., Fradkov A.L. Speed gradient method and its applications. *Avtomatika i telemekhanika*, 2021, no. 9, pp. 3–72, doi: <https://doi.org/10.31857/S0005231021090014> (In Russ.). (Eng. version: *Autom. Remote Control*, 2021, vol. 82, no. 9, pp. 1463–1518, doi: <https://doi.org/10.1134/S0005117921090010>)
- [10] Polyak B.T. *Vvedenie v optimizatsiyu* [Introduction into optimization]. Moscow, Nauka Publ., 1983. 384 p. (In Russ.).
- [11] Karpenko A.P. *Sovremennye algoritmy poiskovoy optimizatsii* [Modern algorithms of search optimization]. Moscow, Bauman MSTU Publ., 2014. 446 p. (In Russ.).
- [12] Nikolenko S., Kadurin A., Arkhangelskaya E. *Glubokoe obuchenie. Pogruzhenie v mir neuronnykh setei* [Deep learning. Diving into the world of neural networks]. Sankt-Petersburg, Piter Publ., 2021. 476 p. (In Russ.).
- [13] Zadeh L. *Ponyatie lingvisticheskoi peremennoi i ego primenenie k prinyatiyu priblizhennykh reshenii* [The concept of a linguistic variable and its application to approximate reasoning]. Moscow, Mir Publ., 1976. 165 p. (In Russ.).
- [14] Novák V., Perfilieva I., Mockor J. *Mathematical principles of fuzzy logic*. Springer, 1999. 320 p. (Russ. ed.: *Matematicheskie printsipy nechetkoi logiki*. Moscow, Fizmatlit Publ., 2006. 347 p.)
- [15] Cybenko G. Approximation by superpositions of a sigmoidal function. *Math. Control Signal. Systems*, 1989, vol. 2, no. 4, pp. 303–314, doi: <https://doi.org/10.1007/BF02551274>
- [16] Lee J., Shin J., Realff M. Machine learning: overview of the recent progresses and implications for the process systems engineering field. *Comput. Chem. Eng.*, 2018, vol. 114, pp. 111–121, doi: <https://doi.org/10.1016/j.compchemeng.2017.10.008>
- [17] Sutton R.S., Barto A.G. *Reinforcement learning*. MIT Press, 2018. 552 p.
- [18] Buoni L., Babuska R., De Schutter B. et al. *Reinforcement learning and dynamic programming using function approximators*. CRC Press, 2010. 280 p.
- [19] Kurrek P., Jocas M., Zoghalmi F. et al. AI motion control — a generic approach to develop control policies for robotic manipulation tasks. *Proc. Int. Conf. Engineering Design*, 2019, vol. 1, no. 1, pp. 3561–3570, doi: <https://doi.org/10.1017/dsi.2019.363>
- [20] Menghal P., Laxmi A. Real time simulation: a novel approach in engineering education. *ICECT*, 2011, vol. 1, pp. 215–219, doi: <https://doi.org/10.1109/ICECTECH.2011.5941592>
- [21] Powell W. *Approximate dynamic programming*. Wiley, 2011. 656 p.
- [22] Bertsekas D., Tsitsiklis J. Neuro-dynamic programming: an overview. *Proc. 34th IEEE Conf. on Decision and Control*, 1995, vol. 1, pp. 560–564, doi: <https://doi.org/10.1109/CDC.1995.478953>
- [23] Prokhorov D., Wunsch D. Adaptive critic designs. *IEEE Trans. Neural Netw.*, 1997, vol. 8, no. 5, pp. 997–1007, doi: <https://doi.org/10.1109/72.623201>
- [24] Werbos P.J. Reinforcement learning and approximate dynamic programming (RLADP)—foundations, common misconceptions, and the challenges ahead. In: *Reinforcement learning and approximate dynamic programming for feedback control*. Wiley, 2012, pp. 1–30.
- [25] Lewis F., Vrable D. Reinforcement learning and adaptive dynamic programming for feedback control. *IEEE Circuits Syst. Mag.*, 2009, vol. 9, no. 3, pp. 32–50, doi: <https://doi.org/10.1109/MCAS.2009.933854>
- [26] Lopes G., Ferreira M., Da Silva Simoes A. et al. Intelligent control of a quadrotor with proximal policy optimization reinforcement learning. *Proc. LARS/SBR/WRE*, 2018, pp. 503–508, doi: <https://doi.org/10.1109/LARS/SBR/WRE.2018.00094>
- [27] Kim J., Hong M., Lee K. et al. Learning to walk a tripod mobile robot using nonlinear soft vibration actuators with entropy adaptive reinforcement learning. *IEEE Robot. Autom. Lett.*, 2020, vol. 5, no. 2, pp. 2317–2324, doi: <https://doi.org/10.1109/LRA.2020.2970945>

- [28] Puccetti L., Kopf F., Rathgeber C. et al. Speed tracking control using online reinforcement learning in a real car. *ICCAR*, 2020, pp. 392–399, doi: <https://doi.org/10.1109/ICCAR49639.2020.9108051>
- [29] Sun L., You F. Machine learning and data-driven techniques for the control of smart power generation systems: an uncertainty handling perspective. *Engineering*, 2021, vol. 7, no. 9, pp. 1239–1247, doi: <https://doi.org/10.1016/j.eng.2021.04.020>
- [30] Tomin N., Kurbatsky V., Guliyev H. Intelligent control of a wind turbine based on reinforcement learning. *ELMA*, 2019, doi: <https://doi.org/10.1109/ELMA.2019.8771645>
- [31] Sierra-Garcia J.E., Santos M., Pandit R. Wind turbine pitch reinforcement learning control improved by PID regulator and learning observer. *Eng. Appl. Artif. Intell.*, 2022, vol. 111, art. 104769, doi: <https://doi.org/10.1016/j.engappai.2022.104769>
- [32] Gunther J., Pilarski P., Helfrich G. et al. Intelligent laser welding through representation, prediction, and control learning: an architecture with deep neural networks and reinforcement learning. *Mechatronics*, 2016, vol. 34, pp. 1–11, doi: <https://doi.org/10.1016/j.mechatronics.2015.09.004>
- [33] Lin W.-S., Yang P.-C. Adaptive critic motion control design of autonomous wheeled mobile robot by dual heuristic programming. *Automatica*, 2008, vol. 44, no. 11, pp. 2716–2723, doi: <https://doi.org/10.1016/j.automatica.2008.03.029>
- [34] Calise A., Rysdyk R. Nonlinear adaptive flight control using neural networks. *IEEE Control Syst. Mag.*, 1998, vol. 18, no. 6, pp. 14–25, doi: <https://doi.org/10.1109/37.736008>
- [35] Kruger T., Schnetter P., Placzek R. et al. Fault-tolerant nonlinear adaptive flight control using sliding mode online learning. *Neural Netw.*, 2012, vol. 32, pp. 267–274, doi: <https://doi.org/10.1016/j.neunet.2012.02.025>
- [36] Lin F.-J., Hung Y.-C., Chen S.-Y. Field-programmable gate array-based intelligent dynamic sliding-mode control using recurrent wavelet neural network for linear ultrasonic motor. *IET Control. Theory Appl.*, 2010, vol. 4, no. 9, pp. 1511–1532, doi: <https://doi.org/10.1049/iet-cta.2009.0066>
- [37] Zhu D., Mei T., Luo M. Adaptive sliding mode control for robots based on fuzzy support vector machines. *ICMA*, 2009, pp. 3469–3474, doi: <https://doi.org/10.1109/ICMA.2009.5246264>
- [38] Perez J., Cabrera J., Castillo J. et al. Bio-inspired spiking neural network for nonlinear systems control. *Neural Netw.*, 2018, vol. 104, pp. 15–25, doi: <https://doi.org/10.1016/j.neunet.2018.04.002>
- [39] Perez-Pena F., Lenero-Bardallo J., Linares-Barranco A. et al. Towards bioinspired close-loop local motor control: a simulated approach supporting neuromorphic implementations. *IEEE ISCAS*, 2017, doi: <https://doi.org/10.1109/ISCAS.2017.8050808>
- [40] Keshta H., Ali A., Saied E. et al. Application of Static Var Compensator (SVC) with PI controller for grid integration of wind farm using harmony search. *Int. J. Emerg. Electr. Power Syst.*, 2016, vol. 17, no. 5, pp. 555–566, doi: <https://doi.org/10.1515/ijeeps-2016-0159>
- [41] Atashpaz Gargari E., Hashemzadeh F., Rajabioun R. et al. Colonial competitive algorithm: a novel approach for PID controller design in MIMO distillation column process. *Int. J. Intell. Comput. Cybern.*, 2008, vol. 1, no. 3, pp. 337–355, doi: <https://doi.org/10.1108/17563780810893446>
- [42] Schillinger M., Hartmann B., Skalecki P. et al. Safe active learning and safe Bayesian optimization for tuning a PI-controller. *IFAC-PapersOnLine*, 2017, vol. 50, no. 1, pp. 5967–5972, doi: <https://doi.org/10.1016/j.ifacol.2017.08.1258>
- [43] Berkenkamp F., Schoellig A., Krause A. Safe controller optimization for quadrotors with Gaussian processes. *IEEE ICRA*, 2016, pp. 491–496, doi: <https://doi.org/10.1109/ICRA.2016.7487170>
- [44] Malikopoulos A., Papalambros P., Assanis D. A learning algorithm for optimal internal combustion engine calibration in real time. *Proc. DETC2007*, 2008, vol. 6-A, pp. 91–100, doi: <https://doi.org/10.1115/DETC2007-34718>
- [45] Liang D., Ding W. Modelling and predicting of a switched reluctance motor drive using radial basis function network-based adaptive fuzzy system. *IET Electr. Power Appl.*, 2009, vol. 3, no. 3, pp. 218–230, doi: <https://doi.org/10.1049/iet-epa.2008.0096>

- [46] Elsodany N., Rezeka S., Maharem N. Adaptive PID control of a stepper motor driving a flexible rotor. *Alex. Eng. J.*, 2011, vol. 50, no. 2, pp. 127–136, doi: <https://doi.org/10.1016/j.aej.2010.08.002>
- [47] Mitra P., Maulik S., Chowdhury S. et al. Anfis based automatic voltage regulator with hybrid learning algorithm. *Proc. Universities Power Engineering Conf.*, 2007, pp. 397–401, doi: <https://doi.org/10.1109/UPEC.2007.4468980>
- [48] Chen H.-C. Optimal fuzzy pid controller design of an active magnetic bearing system based on adaptive genetic algorithms. *Proc. ICMLC*, 2008, vol. 4, pp. 2054–2060, doi: <https://doi.org/10.1109/ICMLC.2008.4620744>
- [49] Wen S., Hu X., Lv X. et al. Q-learning trajectory planning based on Takagi–Sugeno fuzzy parallel distributed compensation structure of humanoid manipulator. *Int. J. Adv. Robot. Syst.*, 2019, vol. 16, no. 1, doi: <https://doi.org/10.1177/1729881419830204>
- [50] Ercin O., Coban R. Identification of linear dynamic systems using the artificial bee colony algorithm. *Turk. J. Elec. Eng. & Comp. Sci.*, 2012, vol. 20, no. 7, pp. 1175–1188, doi: <https://doi.org/10.3906/elk-1012-956>
- [51] Agand P., Shoorehdeli M.A., Khaki-Sedigh A. Adaptive recurrent neural network with Lyapunov stability learning rules for robot dynamic terms identification. *Eng. Appl. Artif. Intell.*, 2017, vol. 65, pp. 1–11, doi: <https://doi.org/10.1016/j.engappai.2017.07.009>
- [52] Sobol I.M., Statnikov R.B. *Vybor optimalnykh parametrov v zadachakh so mnogimi kriteriyami* [Choice of optimum parameters in multicriteria problems]. Moscow, Drofa Publ., 2006. 175 p. (In Russ.).
- [53] Chernykh V.V., Statnikov R.B., Khrapov D.S. et al. Method for investigating the parameter space in problems of passenger car design. *Problemy mashinostroeniya i nadezhnosti mashin*, 2009, no. 4, pp. 19–25. (In Russ.). (Eng. version: *J. Mach. Manuf. Reliab.*, 2009, vol. 38, no. 4, pp. 329–334, doi: <https://doi.org/10.3103/S1052618809040049>)
- [54] Balochian S., Baloochian H. Social mimic optimization algorithm and engineering applications. *Expert Syst. Appl.*, 2019, vol. 134, pp. 178–191, doi: <https://doi.org/10.1016/j.eswa.2019.05.035>
- [55] Kumar S., Kumar A., Shankar G. Crow search algorithm based optimal dynamic performance control of SVC assisted SMIB system. *NPSC*, 2018, doi: <https://doi.org/10.1109/NPSC.2018.8771814>
- [56] Abualigah L. Multi-verse optimizer algorithm: a comprehensive survey of its results, variants, and applications. *Neural Comput. & Applic.*, 2020, vol. 32, no. 16, pp. 12381–12401, doi: <https://doi.org/10.1007/s00521-020-04839-1>
- [57] Wang L., Liu Z., Chen C. et al. Support vector machine based optimal control for minimizing energy consumption of biped walking motions. *Int. J. Precis. Eng. Manuf.*, 2012, vol. 13, no. 11, pp. 1975–1981, doi: <https://doi.org/10.1007/s12541-012-0260-7>
- [58] Xiao L., Zhu Q., Li C. et al. Application of modified teaching-learning algorithm in coordination optimization of TCSC and SVC. In: *CCPR*. Springer, 2014, pp. 44–53, doi: https://doi.org/10.1007/978-3-662-45646-0_5
- [59] Shouran M., Habil M. Tuning of PID controller using different optimization algorithms for industrial DC motor. *ICACITE*, 2021, pp. 756–759, doi: <https://doi.org/10.1109/ICACITE51222.2021.9404616>
- [60] Silvestri M., Pedrazzoli P., Boer C. et al. Compensating high precision positioning machine tools by a self learning capable controller. *Proc. EUSPEN*, 2011, vol. 2, pp. 121–124.
- [61] Tomizuka M. Zero phase error tracking algorithm for digital control. *J. Dyn. Sys., Meas., Control.*, 1987, vol. 109, no. 1, pp. 65–68, doi: <https://doi.org/10.1115/1.3143822>
- [62] Tomizuka M. Dealing with periodic disturbances in controls of mechanical systems. *Annu. Rev. Control*, 2008, vol. 32, no. 2, pp. 193–199, doi: <https://doi.org/10.1016/j.arcontrol.2008.07.002>
- [63] Uchiyama M. Formation of high-speed motion pattern of a mechanical arm by trial. *Transactions of the Society of Instrument and Control Engineers*, 1978, vol. 14, no. 6, pp. 706–712, doi: <https://doi.org/10.9746/sicetr1965.14.706>
- [64] Li W., Maier P., Enge H. Self-learning control applied to vibration control of a rotating spindle by piezopusher bearings. *Proc. Inst. Mech. Eng., I: J. Syst. Control Eng.*, 2004, vol. 218, no. 3, pp. 185–196, doi: <https://doi.org/10.1243/095965104773913867>

- [65] Hladowski L., Galkowski K., Rogers E. Further results on dynamic iterative learning control law design using repetitive process stability theory. *10th nDS*, 2017, doi: <https://doi.org/10.1109/NDS.2017.8070621>
- [66] Kurek J., Zaremba M. Iterative learning control synthesis based on 2-D system theory. *IEEE Trans. Automat. Contr.*, 1993, vol. 38, no. 1, pp. 121–125, doi: <https://doi.org/10.1109/9.186321>
- [67] Pakshin P., Emelianova J., Emelianov M. Iterative learning control design for multiagent systems based on 2D models. *Autom. Remote Control*, 2018, vol. 79, no. 6, pp. 1040–1056, <https://doi.org/10.1134/S000511791806005X>
- [68] Seel T., Werner C., Raisch J. et al. Iterative learning control of a drop foot neuro-prosthesis — generating physiological foot motion in paretic gait by automatic feedback control. *Control. Eng. Pract.*, 2016, vol. 48, pp. 87–97, doi: <https://doi.org/10.1016/j.conengprac.2015.11.007>
- [69] Seel T., Werner C., Schauer T. The adaptive drop foot stimulator — multivariable learning control of foot pitch and roll motion in paretic gait. *Med. Eng. Phys.*, 2016, vol. 38, no. 11, pp. 1205–1213, doi: <https://doi.org/10.1016/j.medengphy.2016.06.009>
- [70] Mohammadpour A., Mishra S., Parsa L. Iterative learning control for fault-tolerance in multi-phase permanent-magnet machines. *Proc. American Control Conf.*, 2013, pp. 5929–5934, doi: <https://doi.org/10.1109/acc.2013.6580768>
- [71] Li J., Wang S., Wang J. et al. Iterative learning control for a distributed cloud robot with payload delivery. *Assem. Autom.*, 2021, vol. 41, no. 3, pp. 263–273, doi: <https://doi.org/10.1108/AA-11-2020-0179>
- [72] Warriar R., Devasia S. Iterative learning from novice human demonstrations for output tracking. *Trans. Hum. Mach. Syst.*, 2016, vol. 46, no. 4, pp. 510–521, doi: <https://doi.org/10.1109/THMS.2016.2545243>
- [73] Lansdell B., Kording K. Towards learning-to-learn. *Curr. Opin. Behav. Sci.*, 2019, vol. 29, pp. 45–50, doi: <https://doi.org/10.1016/j.cobeha.2019.04.005>

Статья поступила в редакцию 17.07.2023

Информация об авторе

ЗАЙЦЕВА Юлия Сергеевна — кандидат технических наук, научный сотрудник лаборатории «Управление сложными системами». Институт проблем машиноведения РАН; доцент кафедры «Системы автоматического управления». Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» (197022, Санкт-Петербург, Российская Федерация, ул. Профессора Попова, д. 5, e-mail: juliazaytsev@gmail.com).

Information about the author

ZAYTSEVA Yuliya Sergeevna — Candidate of Science (Eng.), Researcher, Laboratory for Control of Complex Systems. Institute of Mechanical Engineering Problems of the Russian Academy of Sciences; Associate Professor, Department of Automated Control Systems. Saint Petersburg Electrotechnical University “LETI” (197022, St. Petersburg, Russian Federation, Professora Popova St., Bldg. 5, e-mail: juliazaytsev@gmail.com).

Просьба ссылаться на эту статью следующим образом:

Зайцева Ю.С. Методы искусственного интеллекта для задач управления робототехническими и мехатронными системами: обзор. *Известия высших учебных заведений. Машиностроение*, 2024, № 1, с. 41–56, doi: [10.18698/0536-1044-2024-01-41-56](https://doi.org/10.18698/0536-1044-2024-01-41-56)

Please cite this article in English as:

Zaytseva Y.S. Methods of artificial intelligence in tasks of managing the robotic and mechatronic systems: review. *BMSTU Journal of Mechanical Engineering*, 2024, no. 1, pp. 41–56, doi: [10.18698/0536-1044-2024-01-41-56](https://doi.org/10.18698/0536-1044-2024-01-41-56)