

# Концепция многоагентной нейроподобной когнитивной системы управления мобильным робототехническим комплексом

О. А. Тельминов

АО «Научно-исследовательский институт молекулярной электроники»

otelminov@niime.ru

**Аннотация.** Рассматриваются подходы к разработке системы управления мобильным робототехническим комплексом, в которой сам комплекс и его окружение представляется в виде моделирующих их взаимодействующих агентов. Главный агент выполняет функцию управления комплексом через его цифрового двойника. Особенностью агентов является их уровни реализации в разрезе «скорость принятия решения – размер нейросети»: стратегический, тактический и рефлексный. Низший уровень – нейроморфный, – реализуется на цифро-аналоговом вычислительном ресурсе для реализации спайковых нейронных сетей с диспетчеризацией запуска агентов. Обсуждаются вопросы тестирования и валидации предлагаемой системы управления.

**Ключевые слова:** программные агенты; нейроморфный процессор; когнитивная система управления

## I. ОБЩИЕ ПОЛОЖЕНИЯ

Традиционно реализация управления мобильными робототехническими комплексами (РТК) выполняется на основе подхода, при котором на нижнем уровне разрабатываются, реализуются и настраиваются алгоритмы управления базовыми элементами с последующим усложнением на более высоких структурных уровнях. Современные биологически вдохновленные алгоритмы основаны на моделях, стремящихся воспроизвести требуемое поведение объекта взаимосогласованными действиями элементами всех структурных уровней объекта. При этом такие элементы могут использовать апробированные базовые алгоритмы, их модификации, а также и вновь созданные и адаптированные для выполнения задания алгоритмы.

Целью настоящей работы является нахождение путей реализации требуемого поведения мобильного РТК с применением имеющихся и приобретением новых необходимых навыков. Такая цель может быть выполнена при решении задач, направленных на выявление подходов для: (1) обеспечения самостоятельной постановки задания мобильным РТК самому себе с учетом текущей и прогнозной обстановки; (2) взаимосогласованного по всем уровням структурной иерархии мобильного РТК применения алгоритмов, включая их создание *in situ* при необходимости; (3) повышения прозрачности и объяснимости функционирования мобильного РТК. Совокупность выявленных подходов будет являться концепцией системы управления мобильным РТК.

Решение первой задачи, связанной с постановкой цели системой управления, возможна, по крайней мере, при наличии необходимых компетенций и владения

достаточным объемом информации. Более строгое определение дано в [1]: способность к постановке цели называется целеполаганием, которое основано на потребностях, ценностях, знаниях, установках. Субъект, располагающий такими способностями, способен самостоятельно ставить цели. Если к ним добавляются функции взаимодействия, общение и рассуждение, то у субъекта возникает свойство когнитивности. Система управления, обладающая такими функциями, называется *когнитивной*. После постановки цели следует целереализация, осуществляемая с помощью инструментария и технологии его использования.

Вторая задача подразумевает создание или модификацию алгоритмов в зависимости от требуемой поведенческой функции. С учетом необходимости динамической генерации алгоритмов, а также реализации функции искусственного рассуждения, в качестве инструментария целесообразно использовать нейронные сети.

В основе построения нейросети лежит модель искусственного нейрона: исторически первой была предложена формальная модель, позже – спайковая модель. Особенностью формальной модели нейрона является непрерывное представление сигнала с помощью амплитуды, взвешивание входных сигналов с учетом нейросинаптических весов, последующее суммирование и пропускание через нелинейную функцию. Модель нейрона, основанная на современном представлении биологов о работе естественных нейронов, работает с импульсным (спайковым) кодированием сигналов. Каждый пришедший в нейрон спайк повышает его потенциал, который со временем релаксирует. В случае превышения порогового значения своего потенциала нейрон генерирует спайк на своем выходе. С учетом модели нейрона разрабатывается математический аппарат для обучения нейросетей. Если для формальных нейросетей такой аппарат весьма развит, то для спайковых нейросетей он находится в стадии активного развития: в ряде случаев используют частные подходы, более универсальными являются алгоритмы переноса обученной формальной нейросети на спайковую нейросеть [2].

Аппаратная реализация нейросети выполняется в специализированном для нее сопроцессоре. Управлением сопроцессором, подготовкой данных, обработкой результатов вычисления и другими функциями занимается центральный процессор, построенный по архитектуре фон Неймана. Итоговый нейропроцессор реализует формальные нейросети, нейроморфный (природоподобный) процессор —

спайковые и другие новые биологически вдохновленные модели нейрона. Классические процессоры и нейропроцессоры изготавливаются по традиционной КМОП-технологии [3], в то время как в нейроморфных процессорах все чаще используется элементная база на новых физических принципах, в том числе позволяющая выполнять вычисления и хранение данных в одном элементе. Нейроморфные процессоры стремятся по критерию производительности к нейропроцессорам и в ближайшие несколько лет имеют шанс превзойти их [4], при этом принципиально и существенно (на 2 порядка и более) выигрывают в части энергопотребления.

Взаимодействие внутри системы управления можно реализовать путем разбиения ее на определенное количество автономных агентов. Каждый агент представляет собой нейросетевую модель определенной физической сущности — мобильного РТК и значимых объектов и субъектов из его окружения. Главным является агент системы управления, выполняющий перечисленные выше когнитивные функции и целереализацию, воздействуя на объект управления через соответствующего агента объекта управления. К остальным агентам относятся объект влияния, на который воздействует объект управления; соратники, имеющие общую с объектом управления цель; соперники, соревнующихся с объектом управления и его соратниками; окружающую среду. Главный агент принимает управляющие решения с учетом взаимодействия с остальными агентами.

Объектом управления является мобильный РТК — специализированная подвижная платформа с вычислителем и устройствами сопряжения, датчиками, исполнительными устройствами и средствами коммуникации. Многоагентная нейроподобная когнитивная система управления взаимодействует с мобильным РТК по двусторонней связи, чем обеспечивается ее своевременное информирование о параметрах и динамике движения объекта управления.

Решению третьей задачи — повышению прозрачности и объяснимости принимаемых системой управления решений, способствует деление такой системы на агентов и анализ информационного обмена между ними «по горизонтали». Следует проработать и структуру деления агентов «по вертикали» с дифференциацией по скорости их реакции на необходимость выполнить определенный вид действия.

Ниже рассматривается детализированное представление такой системы управления, являющееся концепцией ее построения.

## II. ПРЕДЛАГАЕМАЯ АРХИТЕКТУРА СИСТЕМЫ УПРАВЛЕНИЯ

### A. Структура объекта управления — мобильного РТК

Мобильный РТК оснащен исполнительными устройствами, управление которыми в итоге приводят к достижению цели его функционирования (рис. 1). Для принятия решения используется целеполагание и информация с датчиков различных физических величин. Устройства сопряжения позволяют согласовать электрические сигналы по уровню, мощности и способу кодирования как от датчиков к вычислителю, так и от вычислителя к исполнительным устройствам, как правило, с высокой потребляемой мощностью.

Оptionальное взаимодействие с внешним миром осуществляется с помощью устройств связи и терминала оператора со средствами отображения информации и средствами ввода данных.

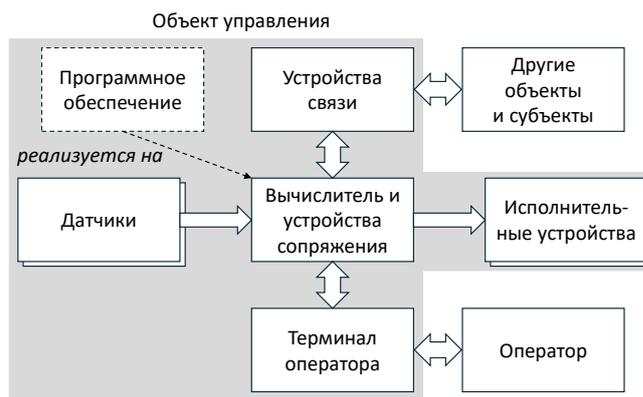


Рис. 1. Структурная схема объекта управления — мобильного РТК

### B. Система управления: цифровой двойник и модели реальных объектов с реализацией в виде агентов

Мобильный РТК, реализуя целеполагание, воздействует на объекты влияния и взаимодействует с ними. В общем случае в этом ему могут помочь соратники и противодействовать соперники. Взаимодействие происходит в окружающей среде, которая также влияет на статическое и динамическое состояние рассматриваемых участников. Система управления мобильным РТК может быть вынесена за пределы объекта управления, быть встроенной в него, либо иметь полную интеграцию с его вычислителем.

Для получения представления об объектах реального мира используются различные модели. В [5] определены понятия компьютерной и математической модели, а также цифровой модели на основе системы таких моделей. На верхнем уровне иерархии располагается цифровой двойник (ЦД) — цифровая модель с двусторонними информационными связями с объектом моделирования. Отсюда следует, что при изменении цифрового двойника изменяется и объект; верно и обратное. Для рассматриваемых объектов только агент объекта управления может выступать в роли ЦД, поскольку с ним возможно установление двусторонних информационных связей. Для объектов влияния, соратников, соперников и окружающей среды разрабатываются компьютерные модели — как основанные на физических законах, так и на данных. Последние реализуются с помощью методов машинного обучения на нейросетях.

Программно-аппаратные агенты представляют собой поведенческие модели всех вышеперечисленные объектов. Главный агент (рис. 2) системы управления принимает конечные решения и воздействует на объект управления. В ходе подготовки таких решений главный агент взаимодействует с другими агентами — моделями объекта управления, объектов влияния, соратников, соперников и окружающей среды. Для последней необходимо реализовать несколько моделей в зависимости от вида среды, уровня детализации и прочих параметров. Агенты, цифровой двойник и модели образуют систему управления.

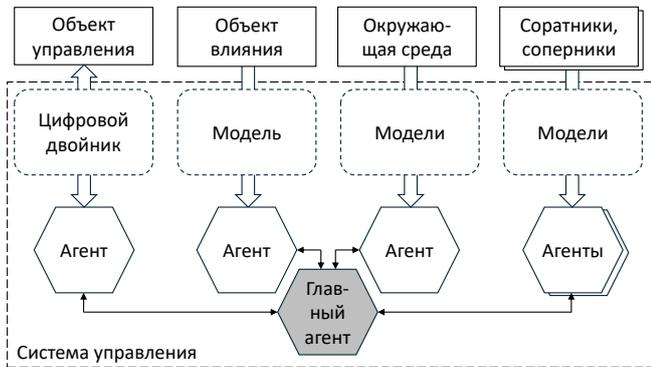


Рис. 2. Система управления во взаимодействии с окружающим миром

К достоинствам многоагентного подхода следует отнести *масштабируемость* и *отказоустойчивость*. Добавление новых объектов в картину мира мобильного РТК приводит к введению соответствующих агентов. Отказ отдельного агента не парализует работу всей системы, как это может возникнуть в целостной монолитной системе. Такой подход позволяет независимо обновлять каждый из агентов по мере необходимости. *Адаптивность* многоагентного решения связана с возможностью перестройки работы отдельных агентов, компенсацией окружения объекта управления количеством соответствующих агентов. Взаимодействие агентов можно *отлаживать в среде виртуального тестирования* [6], позволяя агенту объекта управления выполнять различные сценарии с другими агентами, обучаться и тестировать гипотезы без ущерба для реальных объектов. Единая среда обмена информацией между агентами и принятие решений главным агентом обеспечивает *прозрачность* и *объяснимость* [7] работы системы в целом.

К недостаткам многоагентной системы следует отнести сложность соблюдения *временных ограничений*. С одной стороны, необходимо поддерживать синхронизацию между реальными объектами, моделируемыми с помощью агентов. Своевременное получение необходимых данных о прогнозируемом состоянии объектов реального мира от соответствующих агентов главным агентом влияет на корректность управляющих воздействий на объект управления. Здесь следует учитывать быстродействие реализации агентов, коммуникационные задержки и задержки принятия решения главным агентом. Оценка *согласованности моделей* с функционированием реальных объектов, анализ рисков *безопасности и надежности* получения сведений от датчиков должны учитываться в работе главного агента для корректной выработки управляющего воздействия.

В предлагаемой архитектуре главный агент сотрудничает с агентами, моделирующими поведение соратников для достижения общей цели. Суть коллаборации заключается в совместном построении траектории движения, маневров и избегании конфликтов. Многоагентное обучение с подкреплением приводит к оптимизации получения общего результата путем поддержки действий друг друга и разделения задач.

Конкуренция возникает между главным агентом, агентами-соратниками с одной стороны и агентами-соперниками с другой. Такое взаимодействие приводит к

борьбе за ресурсы, например, за время достижения цели, путем постановки помех или выбора оптимального пути и скорости. Противоборствующие стороны применяют оппонирование, которое сводится к выработке контрмер. В динамической среде такое состязание приводит к развитию стратегии и тактики взаимодействия. В итоге конкуренция приводит к эволюции – адаптивному накоплению опыта и компетенций взаимодействующими сторонами с течением времени и появлением новых форм поведения, не заложенных изначально.

Обучение главного агента нацелено на выбор оптимальных воздействий на объект управления на основе информации его ЦД и других агентов. Стратегия управления вырабатывается с помощью глубокого обучения с подкреплением или предобученной модели (трансферное обучение, *transfer learning*), а также сначала в режиме симуляции, затем – в реальной среде. Развитие главного агента может проводиться и путем популяционного обучения группы его версий с различными стратегиями параллельно с последующим отбором и комбинацией их параметров.

Обучение агентов-соратников строится на кооперативном обучении с подкреплением, в ходе которого вырабатываются взаимная совместимость действий с учетом вклада каждого агента в общий результат. Обучение агентов-соперников, а также агента объекта влияния может быть основано на подражании (*Behavioral Cloning*), при котором нейросеть учится повторять корректные действия из демонстраций с последующим обучением с подкреплением. Совместное обучение агентов-соратников и агентов-соперников целесообразно выполнять с помощью генеративных состязательных нейросетей, в ходе работы которых вырабатываются стратегии противодействия.

Агент окружающей среды обучается на данных о погоде, ветре и других метеоусловиях. В реальной среде модель может обновляться с учетом измеренных данных. Имитация различных препятствий позволяет выполнить предварительное обучение главного агента, агентов-соратников и агентов-соперников.

### С. Цифровой двойник мобильного РТК

Основная задача ЦД мобильного РТК, представленного в виде агента, – выдача информации о состоянии, текущем или прогнозируемом, на основе моделирования кинематики и динамики движения с учетом показаний датчиков. Главный агент системы управления вырабатывает решение и воздействует на такой ЦД, опосредованно изменяя состояние мобильного РТК, что в итоге и обеспечивает двустороннюю информационную связь между ЦД и РТК.

Сначала собирается и обобщается информация о мобильном РТК на основе конструкторских данных, характеристик приводов – двигателей и редукторов (при наличии), движителей (пропеллеров, колес и проч.) и других компонентов. Цифровая копия реального РТК, выполненная с помощью лазерного сканирования или фотограмметрии, дополняет информацию для построения аналитических уравнений движения и имитационных моделей. Анализ данных из протоколов работы (логов) элементов РТК – например, потребляемых двигателями токов, информации с датчиков и по каналу телеметрии, в том числе сопровождаемых синхронными видео- и аудиоданными,

может служить основой для построения нейросетевых моделей, основанных на данных.

При построении ЦД мобильного РТК возможно использование Robot Operation System (ROS) – программного фреймворка для описания его работы [8]. В его состав входят Узлы, Сообщения, Топики, Сервисы, Действия, Пакеты и Файловая система. Узлы реализуют определенные функции (обработка данных с датчиков, расчет траектории, управление двигателями и др.), общаясь друг с другом через Сообщения – структуры типизированных данных. В свою очередь Сообщения передаются через Топики – каналы передачи между Узлами-издателями и Узлами-подписчиками. При необходимости Узлы запрашивают и получают информацию через Сервисы, реализующие клиент-серверную архитектуру. Сервисы для долгосрочных задач, например, планирования траекторий, представляются в виде Действий: клиент направляет цель, сервер периодически возвращает промежуточные результаты и итоговый результат выполнения. В состав Пакета, представляющего собой папку с файлами, входят Узлы, Сообщения, Конфигурации и другие ресурсы. Файловая система обеспечивает работу всех компонентов, в том числе Узлов и Сервисов, в типовом случае прием информации с датчиков, их обработку и выдачу управляющих воздействий на исполнительные устройства.

Описание физической структуры мобильного РТК выполняется в виде файла со структурой Unified Robot Description Format (URDF) для визуализации, отладки поведения в симуляторе Gazebo. Из модели в ROS можно извлечь геометрические данные, физические параметры, логику управления и алгоритмы, данные симуляции и информацию с датчиков. Полученные сведения следует применить в ЦД на собственном движке как модель, геометрию и логику поведения РТК, а наборы данных – для обучения нейросети. Для завершения построения ЦД необходимо обеспечить обратную информационную связь от реального РТК к ЦД. При этом при выполнении реальных заданий мобильным РТК не следует применять обновление программного обеспечения в реальном времени из соображения мер безопасности: обновленный алгоритм необходимо протестировать и верифицировать после выполнения задания и затем обновить в РТК.

#### D. Иерархия представления агентов

Управление мобильным РТК предлагается выполнять с помощью системы на основе многоуровневой архитектуры из четырех уровней: стратегического, тактического, рефлексного и нейроморфного. Первые три уровня реализуются в программном виде (рис. 3), четвертый – в аппаратном. Каждому уровню характерны специфические функции, в то время как в составе иерархии они способны решать задачи от вычислительно сложной выработки целеполагания до быстрой реакции на внешние воздействия. Такое структурирование каждого агента повышает прозрачность его работы за счет объяснимости принятия решений на каждом уровне.

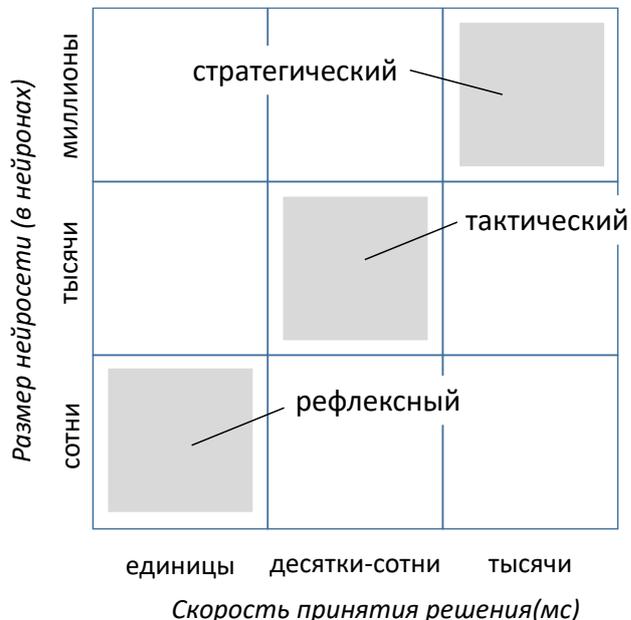


Рис. 3. Вариант реализации уровней иерархии разреза «скорость принятия решения – размер нейросети»

На *стратегическом* уровне агент выполняет долговременное планирование, например, для главного агента и агентов-соратников к нему относятся изменение целеполагания, коррекция маршрута и места назначения. В случае ограниченного вычислительного ресурса могут быть применены мягкие временные требования к получению результата расчета, что позволяет применить рекуррентные вычисления нейросети. На *тактическом* уровне агент реализуется в реальном времени – например, выполняет управление движением (положением и скоростью), коррекцией траектории. *Рефлексный* уровень подразумевает работу в жестком реальном времени, где верное решение, но полученное с опозданием, является неверным. Здесь отрабатываются задачи с наименьшими задержками, например, стабилизация и ориентация в ответ на воздействие окружающей среды, в частности нейросетевой аналог пропорционально-интегрально-дифференцирующего (ПИД) регулятора.

#### E. Нейроморфный уровень представления агентов

Все рассмотренные выше агенты должны быть реализованы на спайковых нейросетях и выполняться на периферийных (edge) нейроморфных процессорах с возможностью обучения с подкреплением. В них необходимо предусмотреть ключевые модели спайковых нейронов (Ижикевича, Leakage-Integrate-and-Fire и другие), реализацию синаптических связей и правил пластичности для обучения. При существовании и развитии таких процессоров на классической КМОП-базе (TrueNorth [9], Алтай [10]) целесообразным видится выполнение ключевой операции векторно-матричного умножения в аналоговой форме, поскольку оно является вычислительно сложной процедурой. Такие вычисления возможны как с применением транзисторов и конденсаторов как накопителя заряда, так и на мемристорных кроссбарах. Управление умножением выполняется с помощью классического управляющего процессора, построенного на одной из архитектур – на наиболее часто применяемой в таких задачах ARM, редко встречающейся Intel (только в Loihi, Loihi 2) и на

активно набирающую популярность RISC-V. Для мемристормых матриц также необходим блок генерации управляющих последовательностей сигналов начальной формовки (опционально), записи и чтения. Особое внимание следует уделить разработке модуля преобразования цифрового сигнала в аналоговую форму и обратно с точки зрения быстродействия и занимаемой площади на кристалле.

Для масштабирования размерности обрабатываемых нейросетей используется матрица процессоров с межкристалльной маршрутизацией пакетов. С целью сокращения непроизводительных цифро-аналоговых и обратных преобразований следует предусмотреть аналоговые тракты обмена между кристаллами.

Важным является вопрос преобразования цифровых и аналоговых сигналов, в том числе с датчиков, в спайковые сигналы. Хотя наиболее распространенной формой является временное представление, следует оценить варианты частотного и амплитудного кодирования.

Операционная система с диспетчеризацией ресурсов позволит обеспечить релевантный доступ к аппаратным ресурсам рассматриваемых агентов с учетом решаемых ими задач с точки зрения стратегического, тактического и рефлексного уровней. При нехватке ресурсов выполняется приоритизация задач с разделением их по времени. В операционной системе предусмотрена очередь сообщений (чат) для обмена информацией между агентами – как по запросу, так и по наступлению событий. Последующий анализ сообщений позволяет увеличить прозрачность и объяснимость принимаемых решений главным агентом.

Для работы с нейроморфным процессором и матрицей таких процессоров необходимо предусмотреть фреймворк – программу для моделирования, обучения и загрузки нейросети в нейроморфную аппаратуру.

#### *F. Тестирование и валидация*

Для обеспечения корректности работы всей системы необходимо выполнить ее тестирование и валидацию [11], решив следующие задачи. Во-первых, определяется функциональная корректность каждого агента и согласованность совместного поведения агентов. Во-вторых, оценивается адекватность представления реальных объектов с помощью ЦД и семейства моделей. В-третьих, система должна быть устойчивой к сбоям и непредвиденным ситуациям, а объект управления своими действиями не приводит к опасным ситуациям. В-четвертых, оцениваются производительность и задержки при реализации агентов на нейроморфной аппаратуре.

Решение перечисленных задач выполняется в несколько этапов: модульное и интеграционное тестирование агентов в среде симуляции, верификация адекватности цифрового двойника и моделей, тестирование в реальном времени на нейроморфной аппаратуре, тестирование отказоустойчивости и робастности, верификация безопасности и анализ

рисков. Для каждого этапа разрабатываются методы тестирования и критерии прохождения теста.

### III. ВЫВОДЫ И ПЕРСПЕКТИВЫ РАЗВИТИЯ

В работе рассмотрены разноуровневые подходы к разработке системы управления целевым объектом из нескольких агентов, каждый из которых представляется набором поведенческих подзадач в разрезе «скорость принятия решения – размер нейросети» и реализуется на матрице нейроморфных процессоров. Обсуждены составляющие системы управления объектом управления: цифровой двойник мобильного РТК и модели реальных объектов с реализацией в виде агентов; иерархия представления агентов, включая поведенческие уровни – стратегический, тактический, рефлексный и аппаратный – нейроморфный уровень представления агентов с учетом тестирования и верификации.

На основе представленной концепции в дальнейшем планируется проработка ее составляющих: разработка цифрового двойника мобильного РТК и моделей его окружения; многоуровневая нейросетевая реализация обучающихся агентов; нейроморфная реализация агентов – программная и аппаратные части; подходы к тестированию и валидации предлагаемой системы управления с учетом прозрачности и объяснимости принимаемых ей решений.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Locke E.A. Toward a Theory of Task Motivation and Incentives. *Organizational Behavior and Human Performance* 1968, 3, 157–189.
- [2] Тельминов О.А. Особенности архитектур нейро- и нейроморфных процессоров // Математическое моделирование в материаловедении электронных компонентов. МММЭК-2024: Материалы VI международной конференции, Москва: ООО "МАКС Пресс", 2024. С. 168-169.
- [3] Красников Г.Я., Горнев Е.С., Матюшкин И.В. Общая теория технологий и микроэлектроника. М.: ТЕХНОСФЕРА, 2020.
- [4] На пути к реализации высокопроизводительных вычислений в памяти на основе мемристормой электронной компонентной базы / Михайлов А.Н., Грязнов Е.Г., Лукоянов В.И., Коряжкина М.Н., Борданов И.А., Щаников С.А., Тельминов О.А., Иванченко М.В., Казанцев В.Б. // Физмат. 2023. Т. 1, № 1. С. 42-64.
- [5] ГОСТ Р 577000.37-2021. Компьютерные модели и моделирование. Цифровые двойники изделий. Общие положения. М.: Российский институт стандартизации, 2021.
- [6] NVIDIA Isaac Sim // URL: <https://developer.nvidia.com/isaac/sim> (дата обращения: 01.04.2025).
- [7] Hsieh, Weiche, et al. A Comprehensive Guide to Explainable AI: From Classical Models to LLMs. arXiv:2412.00800, arXiv, 8 Dec. 2024.
- [8] Robot Operating System // URL: <https://www.ros.org/> (дата обращения: 01.04.2025).
- [9] Merolla, P.A.; Arthur, J.V.; Alvarez-Icaza, R.; Cassidy, A.S.; Sawada, J.; Akopyan, F.; Jackson, B.L.; Imam, N.; Guo, C.; Nakamura, Y.; et al. A Million Spiking-Neuron Integrated Circuit with a Scalable Communication Network and Interface. *Science* 2014, 345, 668–673.
- [10] Нейроморфный процессор "Алтай" для энергоэффективных вычислений / Н.В. Гришанов, А.В. Зверев, Д.Е. Ипатов и др. // Наноиндустрия. 2020. № S96-2. С. 531-538
- [11] Phillips, Ibukun, and C. Robert Kenley. "Validation Framework of a Digital Twin: A System Identification Approach." *INCOSE International Symposium*, vol. 34, no. 1, July 2024, pp. 249–67.