

---

УДК 004.8

DOI: 10.25686/978-5-8158-2474-4-2025-442-450

## Синергетическая интеграция методов роевого интеллекта в задачах визуализации и аналитики

**С. А. Третьяков, И. С. Стародубцев**

ИММ УрО РАН, Екатеринбург, Россия

**Аннотация.** Адаптивные мультиагентные роевые системы (AMPC) представляют перспективное направление для решения сложных задач визуализации и аналитики в условиях неопределенности и динамически изменяющейся среды. Данная работа представляет обзор современных подходов к построению AMPC, основанных на синергетической интеграции четырех ключевых технологий: мультиагентного обучения с подкреплением (MARL), графовых нейронных сетей (GNN), методов вычислительного интеллекта (CI) и трансферного обучения (TL). Особое внимание удалено применению синергетического подхода к задачам роевой разведки местности и одновременной локализации и картографирования (SLAM), где визуализация процесса обучения и аналитика траекторий агентов играют критическую роль в оценке эффективности системы. Проведен сравнительный анализ классических и современных методов роевого интеллекта, выявлены ограничения существующих подходов и показаны преимущества их интеграции. Результаты анализа демонстрируют, что синергетический подход MARL+GNN+CI+TL обеспечивает повышенную масштабируемость, устойчивость к отказам и адаптивность по сравнению с классическими методами, открывая новые возможности для интерактивной визуализации и аналитики роевых систем.

**Ключевые слова:** адаптивные роевые системы, мультиагентное обучение с подкреплением, графовые нейронные сети, вычислительный интеллект, трансферное обучение, научная визуализация, SLAM

### Введение

Задачи координации множества автономных агентов в условиях ограниченной информации и децентрализованного управления возникают в широком спектре прикладных областей, включая робототехнику, системы наблюдения и мониторинга окружающей среды. Особый интерес представляют системы, способные адаптироваться к изменениям в структуре задач и условиях функционирования без перепроектирования базовых алгоритмов управления.

Традиционные подходы к построению мультиагентных систем основаны на заранее определенных правилах взаимодействия (алгоритм Boids [1], потенциальные поля [2]) или иерархических схемах координации [3]. Данные методы обеспечивают предсказуемое поведение в стационарных условиях, однако демонстрируют ограниченную способность к адаптации при изменении параметров среды или отказах отдельных агентов. Альтернативным направлением является применение методов машинного обучения, в частности, мультиагентного обучения с подкреплением (Multi-Agent Reinforcement Learning, MARL), позволяющего агентам вырабатывать стратегии поведения на основе взаимодействия со средой [4].

Одной из ключевых проблем MARL является экспоненциальный рост пространства состояний при увеличении числа агентов, что затрудняет масштабирование алгоритмов на большие роевые системы. Частичным решением данной проблемы служит использование графовых нейронных сетей (Graph Neural Networks, GNN) для структурированного представления взаимодействий между агентами [5, 6]. GNN позволяют учитывать локальную топологию связей и обрабатывать системы переменного размера, сохраняя при этом возможность обобщения на новые конфигурации.

Дополнительные возможности для повышения эффективности обучения предоставляют методы вычислительного интеллекта (Computational Intelligence, CI), включающие эволюционные алгоритмы и роевую оптимизацию. Данные подходы эффективны в задачах глобального поиска и настройки гиперпараметров нейросетевых архитектур [7–9]. Трансферное обучение (Transfer Learning, TL) позволяет переносить выученные стратегии между родственными задачами, значительно сокращая время адаптации к новым условиям [10, 11].

Особую значимость методы координации мультиагентных систем приобретают в контексте задач одновременной локализации и картографирования (Simultaneous Localization and Mapping, SLAM), где

группа агентов должна совместно исследовать неизвестную территорию, строя карту местности и определяя собственные координаты. Роевой подход к решению SLAM обеспечивает повышенную отказоустойчивость и возможность распараллеливания вычислений по сравнению с централизованными методами [12].

Анализ и верификация поведения мультиагентных роевых систем требует развитых средств визуализации и аналитики. Визуальное представление траекторий агентов, динамики обучения и топологии взаимодействий позволяет исследователям выявлять закономерности группового поведения и оценивать эффективность различных алгоритмов координации [13–15].

Цель данной работы состоит в систематическом анализе современных подходов к построению адаптивных мультиагентных роевых систем с акцентом на синергетическое использование методов MARL, GNN, CI и TL. Рассматриваются возможности применения интегративного подхода к задачам роевого SLAM, а также вопросы визуализации и аналитики поведения роевых систем.

### **Методы построения адаптивных роевых систем**

Современные подходы к построению роевых систем можно разделить на несколько основных категорий: классические методы, основанные на предопределенных правилах взаимодействия, и адаптивные подходы, использующие машинное обучение и методы искусственного интеллекта. Рассмотрим каждую из этих категорий более подробно.

#### *Классические подходы*

Основу традиционных методов координации роевых систем составляют алгоритмы, основанные на локальных правилах взаимодействия между агентами. Алгоритм Boids [1] формализует групповое поведение через три базовых правила: разделение (избегание столкновений с соседями), выравнивание (согласование направления движения) и сплочение (стремление к центру локальной группы). Данный подход обеспечивает эмерджентное формирование стайного поведения без централизованного управления.

Методы потенциальных полей [2] представляют среду в виде скалярного поля, где каждая точка характеризуется потенциальной энергией. Агенты движутся в направлении уменьшения потенциала, что позволяет реализовать притяжение к целям и отталкивание от препятствий. Консенсусные алгоритмы [16] обеспечивают достижение согласованного состояния группы агентов через итеративный обмен информацией с соседями.

Указанные подходы характеризуются вычислительной эффективностью и предсказуемостью поведения в стационарных условиях. Однако их применимость ограничена при необходимости адаптации к изменяющимся условиям среды или выполнения сложных многоэтапных задач, требующих долгосрочного планирования. Для преодоления этих ограничений в последние годы активно развиваются методы машинного обучения для роевых систем.

#### *Мультиагентное обучение с подкреплением*

Мультиагентное обучение с подкреплением представляет естественное расширение классического обучения с подкреплением на случай множественных взаимодействующих агентов. В отличие от одноагентных формулировок, MARL оперирует в нестационарной среде, где действия каждого агента влияют на функции вознаграждения остальных участников системы.

Основные подходы к MARL можно классифицировать на value-based и policygradient методы. К первой категории относятся алгоритмы VDN (Value Decomposition Networks) [17], QMIX [18] и QTRAN [19, 20], которые аппроксимируют совместную функцию ценности через композицию индивидуальных Q-функций агентов. Данные методы обеспечивают монотонность между локальными и глобальными оценками, что критично для корректности децентрализованного исполнения политик.

Policy-gradient подходы, включая MADDPG [18, 21, 22], COMA [18, 23] и MAPPO [18], оптимизируют стратегии агентов напрямую в пространстве параметров политик. MADDPG использует централизованных критиков для каждого агента при сохранении децентрализованных акторов, что позволяет эффективно обучать непрерывные стратегии. COMA вводит контрафактуальные оценки для решения проблемы распределения вознаграждения между агентами.

Парадигма CTDE (Centralized Training with Decentralized Execution) получила широкое распространение в кооперативных задачах. Агенты имеют доступ к расширенной информации о состоянии системы во время обучения, но полагаются исключительно на локальные наблюдения при исполнении, что обеспечивает практическую применимость в условиях ограниченной коммуникации.

Ключевыми ограничениями MARL являются экспоненциальный рост сложности с увеличением числа агентов, нестационарность среды обучения и сложность обеспечения стабильной сходимости. Эти факторы затрудняют применение чистых MARL-подходов к крупномасштабным роевым системам. Частичным решением указанных проблем является использование графовых представлений для структурирования взаимодействий между агентами.

#### *Графовые нейронные сети*

Графовые нейронные сети предоставляют эффективный механизм для обработки структурированных данных, где объекты связаны отношениями произвольной топологии. В контексте мультиагентных систем GNN естественным образом моделируют сетевую структуру взаимодействий между агентами.

Базовая архитектура GNN основана на итеративном процессе передачи сообщений (message passing), где каждый узел агрегирует информацию от своих соседей и обновляет собственное представление. Классические варианты включают Graph Convolutional Networks (GCN) [5, 24], GraphSAGE [6] и Graph Attention Networks (GAT) [6, 25], различающиеся способами агрегации и взвешивания соседних узлов.

Применение GNN в мультиагентных системах позволяет решить несколько фундаментальных проблем. Во-первых, GNN инвариантны к перестановкам узлов, что обеспечивает корректную обработку систем с изменяющимся составом агентов. Во-вторых, параметры сети разделяются между всеми узлами, что позволяет масштабировать алгоритмы на произвольное число агентов без переобучения.

Интеграция GNN с MARL реализуется на различных уровнях архитектуры. GNN могут использоваться для построения эмбеддингов состояний, обработки коммуникационных сообщений или формирования централизованных критиков с учетом графовой структуры системы [5, 6, 26, 27]. Такая интеграция особенно эффективна в задачах с явно выраженной пространственной или логической топологией взаимодействий. Дополнительные возможности для повышения эффективности обучения предоставляют методы, не требующие градиентной информации.

#### *Методы вычислительного интеллекта*

Вычислительный интеллект включает семейство алгоритмов оптимизации, не требующих градиентной информации и эффективных в задачах с мультимодальными функциями цели. Основные направления включают эволюционные алгоритмы, роевую оптимизацию и нейроэволюцию.

Эволюционные стратегии (Evolution Strategies, ES) [28, 29] и генетические алгоритмы (Genetic Algorithms, GA) [30] используют принципы естественного отбора для поиска оптимальных решений в пространстве параметров. Алгоритм роя частиц (Particle Swarm Optimization, PSO) [31, 32] моделирует социальное поведение стай для глобальной оптимизации.

В контексте мультиагентных систем CI-методы применяются для настройки гиперпараметров нейросетевых архитектур, оптимизации топологии коммуникационных графов и эволюции стратегий поведения агентов. Нейроэволюция [28] позволяет одновременно оптимизировать архитектуру и веса нейронных сетей, что особенно полезно в задачах с неявно определенной структурой вознаграждения.

Гибридные подходы, сочетающие CI с градиентными методами, демонстрируют перспективные результаты. Например, эволюционные алгоритмы могут использоваться для глобального поиска архитектуры сети с последующей градиентной донастройкой весов [33]. Однако высокая вычислительная стоимость обучения ограничивает применимость CI-методов, что делает актуальными подходы к переносу знаний между задачами.

#### *Трансферное обучение*

Трансферное обучение направлено на повторное использование знаний, полученных при решении одной задачи, для ускорения обучения в родственных задачах. В мультиагентных системах TL особенно актуально ввиду высокой вычислительной стоимости обучения с нуля.

Основные подходы к TL в MARL включают дистилляцию политик [21], где знания экспертных агентов передаются новым участникам системы, и метаобучение [34, 35], позволяющее быстро адаптироваться к новым задачам на основе ограниченного опыта.

Федеративное обучение (Federated Learning) [36] обеспечивает распределенное обучение без централизованного обмена данными, что критично для систем с ограничениями на приватность. Sim2Real-подходы [37] решают проблему переноса политик из симуляции в реальные условия функционирования.

Ключевым вызовом TL является предотвращение негативного переноса, когда знания из исходной задачи ухудшают производительность в целевой области. Для этого применяются методы регуляризации, оценки релевантности задач и архитектурного разделения специфичных и общих компонентов знаний.

#### *Синергетический подход MARL+GNN+CI+TL*

Анализ отдельных методов построения роевых систем показывает, что каждый из подходов обладает специфическими преимуществами и ограничениями. В данной работе предлагается синергетический подход, основанный на интеграции методов мультиагентного обучения с подкреплением, графовых нейронных сетей, вычислительного интеллекта и трансферного обучения. Ключевая идея заключается в том, что недостатки отдельных технологий могут быть компенсированы за счет их комплементарного взаимодействия, что позволяет получить систему с качественно новыми свойствами.

#### **Архитектура интегрированной системы**

Предлагаемая архитектура основана на модульном принципе, где каждый компонент выполняет специализированную функцию в общей системе координации агентов. В основу архитектурного решения положена парадигма CTDE (Centralized Training with Decentralized Execution), обеспечивающая эффективное обучение при сохранении автономности агентов на этапе исполнения.

MARL обеспечивает основной механизм выработки стратегий поведения через взаимодействие со средой. GNN отвечает за структурированное представление взаимодействий между агентами и эффективную коммуникацию в рамках локальных топологий. CI-модули осуществляют глобальную оптимизацию гиперпараметров системы и архитектурных решений. TL-компоненты обеспечивают быструю адаптацию к новым задачам и условиям функционирования.

Центральным элементом архитектуры является GNN-слой, который формирует графовые эмбеддинги на основе локальных наблюдений агентов и структуры их взаимодействий. Эти эмбеддинги служат входными данными для MARL-политик, что позволяет агентам принимать решения с учетом состояния ближайшего окружения. Параллельно CI-модули оптимизируют параметры GNN-архитектуры (глубина сети, размерность эмбеддингов, механизмы внимания) и MARL-алгоритмов (размер буферов опыта, коэффициенты обучения).

Общая схема взаимодействия компонентов представлена на рисунке 1.



Рисунок 1. Архитектура синергетического подхода MARL+GNN+CI+TL

Особенностью предлагаемой архитектуры является гибкий механизм коммуникации при нарушении прямой связи с централизованным критиком. В случае фрагментации сети агенты используют многошаговую маршрутизацию сообщений через GNN-граф, где промежуточные узлы

выполняют роль ретрансляторов градиентной информации. Это обеспечивает поддержание координации даже при значительных нарушениях топологии сети.

TL-компоненты интегрированы на нескольких уровнях системы. Дистилляция политик позволяет инициализировать новых агентов на основе знаний экспертов.

Метаобучение обеспечивает быструю адаптацию к изменениям в составе роя или параметрах среды. Федеративные схемы позволяют агентам обмениваться обобщенными знаниями без раскрытия локальных данных.

### Компенсация ограничений методов

Синергетическое взаимодействие компонентов системы позволяет преодолеть ключевые ограничения отдельных подходов. Проблема масштабируемости MARL решается через применение GNN, которые обеспечивают линейный рост вычислительной сложности с увеличением числа агентов благодаря локальности обработки сообщений. Нестационарность среды обучения компенсируется CI-модулями, которые адаптивно подстраивают параметры системы при обнаружении изменений в динамике среды.

Высокая вычислительная стоимость CI-методов снижается за счет применения TL для инициализации популяций эволюционных алгоритмов на основе предыдущего опыта. Это значительно сокращает число поколений, необходимых для достижения приемлемого качества решений. Проблема негативного переноса в TL минимизируется через использование GNN-эмбеддингов, которые предоставляют структурированные представления для оценки релевантности между задачами.

Фиксированная структура классических MARL-архитектур преодолевается через нейроэволюцию топологий GNN, что позволяет системе адаптироваться к изменениям в составе роя или типе решаемых задач. CI-модули обеспечивают поиск оптимальных коммуникационных протоколов и схем агрегации информации в GNN-слоях.

### Применение к задачам роевого SLAM

Задача одновременной локализации и картографирования для группы агентов представляет комплексную проблему, требующую координации исследовательского поведения, обмена информацией о наблюдениях и совместного построения карты местности. Синергетический подход естественным образом адаптируется к специфике SLAM-задач.

MARL-компонент обучает агентов стратегиям исследования, балансирующим между детальным изучением известных областей и поиском новых территорий. Функция вознаграждения формируется на основе метрик покрытия территории, качества локализации и эффективности коммуникации между агентами. Централизованный критик в фазе обучения позволяет координировать глобальную стратегию исследования, учитывая состояние всего роя и прогресс картографирования.

GNN-слой моделирует пространственные и информационные связи между агентами. Узлы графа соответствуют агентам и ключевым точкам карты, ребра отражают возможности коммуникации и пространственную близость. Механизм передачи сообщений позволяет агентам обмениваться наблюдениями и координировать планы исследования без централизованной координации.

При фрагментации роя на изолированные подгруппы (что характерно для SLAM в сложных средах с препятствиями) каждая группа способна поддерживать локальное согласование стратегий через peer-to-peer-дистилляцию политик. Агенты используют ранее синхронизированные параметры для построения согласованных локальных карт. При восстановлении связи между подгруппами происходят синхронизация и слияние карт на основе общих ориентиров и доверительных метрик качества наблюдений.

CI-модули оптимизируют параметры алгоритмов SLAM (пороги обнаружения петель, стратегии слияния карт) и конфигурацию роя (количество агентов, распределение ролей). Эволюционные алгоритмы позволяют найти компромисс между точностью картографирования и вычислительными ресурсами.

TL-компоненты обеспечивают адаптацию к новым типам местности и условиям наблюдения. Политики, обученные в симуляции, адаптируются к реальным условиям через механизмы доменной адаптации. Метаобучение позволяет быстро настраивать параметры системы при переходе между помещениями и открытыми пространствами.

### Сравнительный анализ подходов

Для систематического сравнения различных подходов к построению роевых систем рассмотрим ключевые характеристики по критериям масштабируемости, устойчивости, скорости обучения и адаптивности (табл. 1).

Таблица 1. Сравнительный анализ подходов к построению роевых систем

Подход	Масштабируемость	Устойчивость	Скорость обучения	Адаптивность
MARL	Средняя: экспоненциальный рост сложности ограничивает большие рои	Низкая: критик снижает неустойчивость, но сбои агентов требуют перезапуска	Низкая: требуется много эпизодов, особенно для off-policy методов	Средняя: обучаются новые стратегии, но без TL требуется долгое дообучение
GNN	Хорошая: инвариантность к числу агентов, теоретически масштабируемые	Высокая при структурной связи: работает при отказе отдельных узлов	Средняя: добавляет коммуникацию, замедляя шаги обучения	Хорошая: обрабатывает изменяющиеся связи и новых агентов
CI	Высокая для параллелизма: хорошо масштабируется на кластерах	Умеренная: стойки к шуму, но при сбоях требуется переэволюция	Низкая: требуют большого количества оценок	Низкая: переобучение при изменениях среды
Transfer Learning	Высокая: переносит опыт, снижая нагрузку при масштабах	Высокая: улучшает устойчивость, агенты быстро переобучаются	Очень высокая: использование предобученных компонентов ускоряет обучение	Очень высокая: перенос политик обеспечивает адаптацию «на лету»
MARL + GNN + CI + TL	Очень высокая: комбинированное решение масштабируется на большие роевые структуры	Очень высокая: устойчивость за счет GNN-коммуникации и CI-модулей	Высокая: TL ускоряет инициализацию, CI оптимизирует архитектуру	Очень высокая: адаптивная коммуникация и совместное обучение политик

Представленное в таблице сравнение демонстрирует преимущества предлагаемого синергетического подхода по всем рассматриваемым критериям. Новым является подход, при котором четыре различные технологии интегрированы в единую архитектуру, где каждый компонент не только выполняет свою функцию, но и активно компенсирует ограничения других методов. Ключевым преимуществом является способность системы адаптироваться к изменениям масштаба, состава роя и условий функционирования без существенной деградации производительности, что принципиально отличает предлагаемое решение от существующих подходов, использующих максимум два-три из рассматриваемых компонентов.

Практическая реализация синергетического подхода требует тщательной балансировки между компонентами системы. Чрезмерная сложность архитектуры может приводить к увеличению времени обучения и снижению интерпретируемости поведения агентов. Необходимо развитие методов автоматической настройки весов различных компонентов в зависимости от специфики решаемых задач.

### Заключение

В данной работе представлен систематический анализ современных подходов к построению адаптивных мультиагентных роевых систем с акцентом на их применение в задачах визуализации и

аналитики. Рассмотрены ключевые ограничения классических методов координации роев и показаны возможности их преодоления за счет интеграции методов искусственного интеллекта.

Предложенный синергетический подход, объединяющий мультиагентное обучение с подкреплением, графовые нейронные сети, методы вычислительного интеллекта и трансферное обучение, демонстрирует качественные преимущества по сравнению с использованием отдельных технологий. Ключевые достоинства интегративного решения включают повышенную масштабируемость за счет локальности GNN-обработки, улучшенную устойчивость к отказам через адаптивную реконфигурацию системы и сокращение времени адаптации к новым задачам благодаря механизмам трансферного обучения.

Особое внимание уделено применению предлагаемого подхода к задачам роевой разведки местности и одновременной локализации и картографирования. Рассмотрена адаптация синергетической архитектуры к специфике SLAM-задач, включая координацию исследовательского поведения агентов и распределенное построение карт местности. Кроме того, предлагаемый подход открывает новые возможности для моделирования и визуализации сложных динамических систем, включая физические процессы, биологические популяции и социальные сети, где поведение множества взаимодействующих элементов определяет макроскопические свойства системы.

Вместе с тем практическая реализация комплексных систем, интегрирующих четыре различные технологии, сопряжена с рядом технических вызовов. К основным ограничениям относятся увеличение архитектурной сложности, необходимость тщательной балансировки параметров различных компонентов и снижение интерпретируемости поведения системы. Данные аспекты требуют дальнейших исследований в направлении разработки методов автоматической настройки и анализа гибридных архитектур.

Перспективными направлениями развития являются исследование применимости синергетического подхода к другим классам задач коллективного интеллекта, разработка специализированных метрик для оценки эффективности интегрированных роевых систем, а также создание инструментов интерактивной визуализации для анализа динамики обучения и поведения больших групп агентов. В контексте научной визуализации особый потенциал представляет использование адаптивных роевых систем для моделирования и визуального анализа многочастичных систем, турбулентных течений, эволюционных процессов и других явлений, характеризующихся сложной пространственно-временной динамикой. Важным направлением является интеграция методов объяснимого искусственного интеллекта для повышения прозрачности принятия решений в мультиагентных системах и разработка визуальных интерфейсов для интерактивного анализа коллективного поведения.

### Список литературы

1. C. W. Reynolds, Flocks, herds and schools: A distributed behavioral model, in: Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '87, Association for Computing Machinery, New York, NY, USA, 1987, p. 25–34. URL: <https://doi.org/10.1145/37401.37406>. doi:10.1145/37401.37406.
2. O. Khatib, Real-time obstacle avoidance for manipulators and mobile robots, in: Proceedings. 1985 IEEE international conference on robotics and automation, volume 2, IEEE, 1985, pp. 500–505.
3. I. N. Dubina, N. M. Oskorbin, D. S. Khvalynskiy, Decision-making coordination in hierarchical systems, World of Economics and Management 19 (2019) 5–18. URL: <https://doi.org/10.25205/2542-0429-2019-19-2-5-18>, in Russian.
4. S. V. Albrecht, F. Christianos, L. Schäfer, Multi-Agent Reinforcement Learning: Foundations and Modern Approaches, MIT Press, 2024. URL: <https://www.marl-book.com/>.
5. A. Goeckner, Y. Sui, N. Martinet, X. Li, Q. Zhu, Graph neural network-based multiagent reinforcement learning for resilient distributed coordination of multi-robot systems, in: 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2024, pp. 5732–5739.
6. Z. Liu, J. Zhang, E. Shi, Z. Liu, D. Niyato, B. Ai, X. S. Shen, Graph neural network meets multi-agent reinforcement learning: Fundamentals, applications, and future directions, IEEE Wireless Communications (2024).
7. Z. Huang, S. Batra, T. Chen, R. Krupani, T. Kumar, A. Molchanov, A. Petrenko, J. A. Preiss, Z. Yang, G. S. Sukhatme, Quadswarm: A modular multi-quadrotor simulator for deep reinforcement learning with direct thrust control, 2023. URL: <https://arxiv.org/abs/2306.09537>. arXiv:2306.09537.

8. Охрименко, Методы вычислительной математики в нейронных сетях, in: Международная научно-техническая конференция молодых ученых БГТУ им. ВГ Шухова, посвященная 170-летию со дня рождения ВГ Шухова, 2023, pp. 354–357.
9. Родзин, Вычислительный интеллект: немонотонные логики и графическое представление знаний, Программные продукты и системы (2002) 21–23.
10. M. Zhao, G. Wang, Q. Fu, X. Guo, Y. Chen, T. Li, X. Liu, Mw-maddpg: a metalearning based decision-making method for collaborative uav swarm, *Frontiers in Neurorobotics* 17 (2023) 1243174.
11. A. W. Mamond, M. Kundroo, S.-e. Yoo, S. Kim, T. Kim, Fldqn: Cooperative multi-agent federated reinforcement learning for solving travel time minimization problems in dynamic environments using sumo simulation, *Sensors* 25 (2025) 911.
12. Касаткин, Приходько, Роеевые алгоритмы одновременной локализации и построения карты, in: Прикладная математика: современные проблемы математики, информатики и моделирования, ФГБУ "Российское энергетическое агентство" Минэнерго России Краснодарский ЦНТИ – филиал ФГБУ "РЭА" Минэнерго России, Краснодар, 2022, pp. 124–128. In Russian.
13. A. Dorri, S. S. Kanhere, R. Jurdak, Multi-agent systems: A survey, *Ieee Access* 6 (2018) 28573–28593.
14. Бронников, Адаптивное визуальное управление производственным агентом, Технология приборостроения (2014) 3–6.
15. Фролов, Применение многоагентных технологий при построении интеллектуальных компьютерных обучающих систем, Гаудеамус (2013) 223–225.
16. R. Olfati-Saber, J. A. Fax, R. M. Murray, Consensus and cooperation in networked multi-agent systems, *Proceedings of the IEEE* 95 (2007) 215–233.
17. P. Sunehag, G. Lever, A. Gruslys, W. M. Czarnecki, V. Zambaldi, M. Jaderberg, M. Lanctot, N. Sonnerat, J. Z. Leibo, K. Tuyls, T. Graepel, Value-decomposition networks for cooperative multi-agent learning, 2017. URL: <https://arxiv.org/abs/1706.05296>. arXiv:1706.05296.
18. G. Papoudakis, F. Christianos, L. Schäfer, S. V. Albrecht, Benchmarking multiagent deep reinforcement learning algorithms in cooperative tasks, *arXiv preprint arXiv:2006.07869* (2020).
19. T. Rashid, M. Samvelyan, C. S. De Witt, G. Farquhar, J. Foerster, S. Whiteson, Monotonic value function factorisation for deep multi-agent reinforcement learning, *Journal of Machine Learning Research* 21 (2020) 1–51.
20. K. Son, D. Kim, W. J. Kang, D. E. Hostallero, Y. Yi, Qtran: Learning to factorize with transformation for cooperative multi-agent reinforcement learning, in: International conference on machine learning, PMLR, 2019, pp. 5887–5896.
21. S. Wadhawania, D.-K. Kim, S. Omidshafiei, J. P. How, Policy distillation and value matching in multiagent reinforcement learning, in: 2019 IEEE/RSJ international conference on intelligent robots and systems (IROS), IEEE, 2019, pp. 8193–8200.
22. R. Lowe, Y. I. Wu, A. Tamar, J. Harb, O. Pieter Abbeel, I. Mordatch, Multi-agent actor-critic for mixed cooperative-competitive environments, *Advances in neural information processing systems* 30 (2017).
23. J. Foerster, G. Farquhar, T. Afouras, N. Nardelli, S. Whiteson, Counterfactual multi-agent policy gradients, in: *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.
24. T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, 2017. URL: <https://arxiv.org/abs/1609.02907>. arXiv:1609.02907.
25. P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, Y. Bengio, Graph attention networks, 2018. URL: <https://arxiv.org/abs/1710.10903>. arXiv:1710.10903.
26. L. Ratnabala, A. Fedoseev, R. Peter, D. Tsetserukou, Magnnet: Multi-agent graph neural network-based efficient task allocation for autonomous vehicles with deep reinforcement learning, *arXiv preprint arXiv:2502.02311* (2025).
27. H. Park, B. Seong, S.-K. Ko, Spectra: Scalable multi-agent reinforcement learning with permutation-free networks, *arXiv preprint arXiv:2503.11726* (2025).
28. N. Vaughan, Multi-agent reinforcement learning for swarm retrieval with evolving neural network, in: *Biomimetic and Biohybrid Systems: 7th International Conference, Living Machines 2018, Paris, France, July 17–20, 2018, Proceedings* 7, Springer, 2018, pp. 522–526.
29. T. Salimans, J. Ho, X. Chen, S. Sidor, I. Sutskever, Evolution strategies as a scalable alternative to reinforcement learning, 2017. URL: <https://arxiv.org/abs/1703.03864>. arXiv:1703.03864.
30. Z. Zhu, C. Yu, J. Wang, A hybrid genetic algorithm and proximal policy optimization system for efficient multi-agent task allocation, *Systems* 13 (2025) 453.
31. L. Leuzzi, S. Jones, S. Hauert, D. Bacciu, A. Cossu, Lifelong evolution of swarms, 2025. URL: <https://arxiv.org/abs/2503.17763>. arXiv:2503.17763.

32. S. Lin, J. Wang, B. Huang, X. Kong, H. Yang, Bio particle swarm optimization and reinforcement learning algorithm for path planning of automated guided vehicles in dynamic industrial environments, *Scientific Reports* 15 (2025) 463.
33. E. Real, A. Aggarwal, Y. Huang, Q. V. Le, Regularized evolution for image classifier architecture search, in: *Proceedings of the aaai conference on artificial intelligence*, volume 33, 2019, pp. 4780–4789.
34. A. S. Nipu, S. Liu, A. Harris, Enabling multi-agent transfer reinforcement learning via scenario independent representation, in: *2023 IEEE Conference on Games (CoG)*, 2023, pp. 1–8. doi:10.1109/CoG57401.2023.10333236.
35. C. Finn, P. Abbeel, S. Levine, Model-agnostic meta-learning for fast adaptation of deep networks, in: *International conference on machine learning*, PMLR, 2017, pp. 1126–1135.
36. B. McMahan, E. Moore, D. Ramage, S. Hampson, B. A. y Arcas, Communicationefficient learning of deep networks from decentralized data, in: *Artificial intelligence and statistics*, PMLR, 2017, pp. 1273–1282.
37. J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, P. Abbeel, Domain randomization for transferring deep neural networks from simulation to the real world, in: *2017 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, IEEE, 2017, pp. 23–30.