

УДК 621.316.926:621.311.1:004

EDN ZSNMHG

**МУЛЬТИАГЕНТНОЕ УПРАВЛЕНИЕ
РЕЖИМОМ РАБОТЫ ЭНЕРГОРАЙОНА
С РАСПРЕДЕЛЕННЫМИ ЭНЕРГОРЕСУРСАМИ
НА БАЗЕ АКТЕР-КРИТИЧЕСКИХ НЕЙРОСЕТЕЙ**

Н.В. Томин

ORCID: 0000-0001-5898-9649 e-mail: tomin@isem.irk.ru
Институт систем энергетики им. Л.А. Мелентьева СО РАН
Иркутск, Россия

Д.Н. Сидоров

ORCID: 0000-0002-3131-1325 e-mail: sidorov@isem.irk.ru
Институт систем энергетики им. Л.А. Мелентьева СО РАН
Иркутск, Россия

Анализируется проблема обеспечения устойчивости напряжения в энерго-районах с распределенными ВИЭ, функционирующих в условиях неопределенности нагрузок и возможной автономии. Для решения задачи вторичного регулирования напряжения предлагается децентрализованный адаптивный алгоритм на базе мультиагентного обучения с подкреплением. В его основе лежит актер-критиче-ская архитектура, реализующая схему «централизованное обучение – децентрализованное исполнение», где каждый агент, ассоциированный с системообразующим инверто-ром, независимо оптимизирует свою стратегию управления. Ключевой особен-ностью подхода является адаптивная настройка параметров управления по стагизму, при которой уставка напряжения инвертора динамически корректируется по отклю-нению реактивной мощности. Для координации агентов в процессе обучения ис-пользуется обмен информацией о состоянии соседних узлов, а глобальная функция вознаграждения агентов, включающая пространственное дисконтирование, направ-лена на минимизацию отклонений напряжения и перегрузок в сети. Управляющим воздействием для агента выступает уставка напряжения, выбираемая из дискретного набора значений. Проведенные экспериментальные расчеты показали, что предло-женный подход обеспечивает быструю стабилизацию напряжения в утяжеленных и послеаварийных режимах, снижение загрузки линий на 30 % и существенное умень-шение потерь мощности относительно режима без управления. Такой результат де-монстрирует высокую эффективность системы в поддержании стабильного режима как при нормальной, так и при автономной или островной работе энергорайона с РЭР.

Ключевые слова: актер-критическая нейросеть; вторичное управление; воз-обновляемые источники энергии; мультиагентная система; обучение с подкрепле-нием; распределенные энергетические ресурсы; энергорайоны.

Для цитирования: Томин Н.В., Сидоров Д.Н. Мультиагентное управление режимом работы энергорайона с распределенными энергоресурсами на базе актер-критических нейросетей // Интеллектуальная Электротехника. 2026. № 1. С. 117-144. EDN ZSNMHG

MULTI-AGENT CONTROL OF POWER DISTRICT WITH DISTRIBUTED ENERGY RESOURCES BASED ON ACTOR-CRITIC NEURAL NETWORKS

N.V. Tomin

ORCID: 0000-0001-5898-9649 e-mail: tomin@isem.irk.ru
Melentiev Energy Systems Institute of SB RAS
Irkutsk, Russia

D.N. Sidorov

ORCID: 0000-0002-3131-1325 e-mail: sidorov@isem.irk.ru
Melentiev Energy Systems Institute of SB RAS
Irkutsk, Russia

Abstract. This paper addresses the problem of voltage stability in power districts with distributed RESs operating under load uncertainty and possible autonomy. To solve the secondary voltage control problem, a decentralized adaptive algorithm based on multi-agent reinforcement learning (MARL) is proposed. The algorithm is built upon an actor-critic architecture implementing the “centralized training – decentralized execution” scheme, where each agent, associated with a grid-forming inverter, independently optimizes its control strategy. A key feature of the approach is the adaptive tuning of droop control parameters, in which the inverter voltage setpoint is dynamically adjusted based on reactive power deviation. Agent coordination during training is achieved through exchanging state information with neighboring nodes, while a global reward function with spatial discounting is designed to minimize voltage deviations and network overloading. The control action for an agent is the voltage setpoint selected from a discrete set of values. Simulation studies demonstrated that the proposed approach ensures rapid voltage stabilization in stressed and post-fault modes, reduces line loading by 30%, and significantly decreases power losses compared to the uncontrolled mode. These results highlight the high effectiveness of the system in maintaining stable operation under normal conditions as well as during autonomous or islanded operation of a DER-based power district.

Keywords: actor-critical neural network; secondary control; renewable energy sources; multi-agent system; reinforcement learning; distributed energy resources; power districts.

For citation: N.V. Tomin and D.N. Sidorov, “Multi-agent control of power district with distributed energy resources based on actor-critic neural networks”, *Smart Electrical Engineering*, no. 1, pp. 117-144, 2026. EDN ZSNMHG

I. Введение

В современных электроэнергетических системах все большую роль начинают играть энергорайоны, оснащенные распределенными энергетическими ресурсами (РЭР) – локальными источниками генерации и средствами управления, интегрированными в распределительные сети. К РЭР относятся как объекты возобновляемой энергетики (солнечные, ветровые и малые гидроэлектростанции), так и установки малой синхронной генерации, мини-ТЭЦ, дизельные электростанции, газопоршневые и газотурбинные агрегаты, а также системы накопления энергии и управляемые нагрузки [1]. Энергорайон с РЭР представляет собой территориально ограниченный сегмент сети, способный не только потреблять, но и вырабатывать электроэнергию, а при необходимости – функционировать автономно. Мировая практика последних лет демонстрирует устойчивый рост внедрения таких систем, что связано с задачами повышения надежности энергоснабжения, снижения технологических потерь и расширения возможностей маневрирования генерацией. Для России актуальность развития энергорайонов с РЭР особенно велика: масштабная территория с множеством удаленных и труднодоступных регионов, высокая доля изношенных сетей, а также необходимость обеспечения малых населенных пунктов и промышленных объектов стабильным энергоснабжением требуют новых интеллектуальных подходов к управлению режимами работы распределительных сетей [2]. Такие энергорайоны способны повысить устойчивость энергосистемы как в нормальных, так и в аварийных и островных режимах, обеспечивая гибкость и адаптивность локального энергоснабжения.

Общая интеграция РЭР в энергорайоны среднего и низкого напряжения связана со следующими техническими проблемами: перенапряжение в узлах, где «выдается» значительная активная мощность в сеть, перегрузка распределительных трансформаторов и линий, другие проблемы, связанные с напряжением (несбалансированность, фликер и т.п.); некорректная работа систем релейной защиты из-за двунаправленных потоков мощностью. Кроме того, автоматическое управление режимами энергорайонов с РЭР связано с существованием трех режимов: 1) параллельная работа с энергосистемой с выдачей или без выдачи мощности в сеть; 2) изолированный (автономный) режим с обеспечением электроснабжения потребителей; 3) островной режим, когда объект РГ в нормальном режиме работает параллельно с энергосистемой, но в случае возникновения высоких рисков или аварии в распределительных сетях может быть выделен на электроснабжение энергорайона от объектов РГ [2]. При отключении от основной сети в энергорайоне существует два уровня управления: первичное и вторичное [3]. Первичное относится к управлению самого нижнего уровня в РГ для поддержа-

ния заданного уровня напряжения, тогда как вторичное касается уже кооперативной координации локальных опорных напряжений каждого РГ для достижения целей регулирования режима всей сети [4]. Существующие методы вторичного управления можно разделить на два основных класса: централизованные и распределенные. Хотя централизованные методы хорошо себя зарекомендовали для регулирования частоты в больших энергосистемах, в отношении энергорайонов с РЭР, особенно изолированных от централизованной ЭЭС, они могут быть непрактичными, так как связаны с рядом проблем: потери в пропускной способности в линиях связи, проблемы единой точки отказа, а также «проклятия размерности» [5]. Поэтому более эффективными для вторичного управления напряжением и частотой энергорайонами с РЭР представляются распределенные и децентрализованные подходы, построенные на мультиагентных принципах.

В мультиагентном способе решения об участии в осуществлении вторичного регулирования, коррекции мощности для восстановления резервов вторичного регулирования и разрешения на изменение мощности принимаются локальным устройством управления (контроллером) [6]. Например, в [7] рассматриваются подобные решения, адаптированные для энергорайонов с РГ (минигриды или локальные интеллектуальные энергосистемы, ЛИЭС) на основе многоагентных систем, когда задача вторичного регулирования напряжения и частоты реализуется через координацию работы локальных регуляторов напряжения, установленных на источниках малой синхронной генерации или узлах сети, посредством обмена информацией между агентами. При этом классические локальные устройства управления могут здесь иметь ограничения, особенно в тех случаях, когда объекты РГ в энергорайонах относятся в основном к возобновляемым источникам энергии (ВИЭ). В РФ это особенно актуально при развитии изолированных энергорайонов, т.н. активных гибридных энергокомплексов (АГЭК) в зонах децентрализованного энергоснабжения [8]. Для них современные интеллектуальные инверторы способны обеспечивать поддержку активной и реактивной мощности фидера, работая в нескольких режимах управления [9], поэтому потенциально могут быть эффективны для устранения проблем с качеством напряжения, возникающих из-за изменчивости РЭР. В [10] предложены мультиагентные алгоритмы управления инвертором в режиме виртуальной синхронной машины, которые для изолированных энергорайонов позволяют успешно регулировать частоту и обеспечивать устойчивость, достигая баланса активной мощности при пиковых нагрузках в сетях с дефицитом мощности, плавно производить переключение с изолированного режима работы на параллельный с объединенной энергосистемой. В отличие от традиционных многоагентных систем управления, где агенты действуют по заранее заданным правилам и алгоритмам, в данной работе предлагается

новая архитектура мультиагентного управления на основе методов обучения с подкреплением (*Multi-Agent Reinforcement Learning, MARL*) [11]. Такой подход обеспечивает не только координацию между индивидуальными устройствами, но и способность адаптироваться к меняющимся условиям работы сети за счет формирования оптимальных стратегий управления в ходе процесса обучения. Принципиальным преимуществом *MARL* является возможность учитывать стохастическую изменчивость генерации, колебания нагрузки и неполноту информации о состоянии сети, сохраняя высокую устойчивость управления даже в послеаварийных и островных режимах.

Целью настоящей статьи является разработка и исследование мультиагентного алгоритма вторичного регулирования напряжения для энергорайонов с РЭР, основанного на актера-критической архитектуре в схеме «централизованное обучение – децентрализованное исполнение». В предложенной системе каждый агент – интеллектуальный контроллер, ассоциированный с системообразующим инвертором – корректирует уставку напряжения на основе локальных измерений и обмена данными с ближайшими узлами, что позволяет эффективно поддерживать профиль напряжения, снижать перегрузки оборудования и потери мощности как в нормальных, так и в автономных или островных режимах работы энергорайона.

II. Принципы вторичного управления режимом работы энергорайонов на базе системообразующих инверторов

Сегодня инвертор, управляемый напряжением (*Voltage-Controlled Voltage Source Inverter*), представляет собой тип системообразующих инверторов (*grid-formed*) и широко используется в энергорайонах с РЭР для обеспечения быстрой поддержки напряжения/частоты [12]. Благодаря функции управления напряжением, он может поддерживать стабильность сети при колебаниях нагрузки, что особенно важно в микросетях, гибридных энергорайонах и изолированных объектах, где надежность энергоснабжения критична. Типичная схема таких инверторов может быть основана на принципе децентрализованной архитектуры управления: для каждого РГ используется вторичный контроллер для координации с соседними РГ и динамической генерации опорного напряжения. Полученное опорное напряжение затем используется первичным контроллером нижнего уровня в качестве опорного для отслеживания. Общая цель вторичного управления в этом случае заключается в поддержании напряжения и частоты всех РГ на заданном уровне, несмотря на возмущения и отказы в электрической сети и отслеживание несовершенств первичного управления.

В рамках задачи регулирования напряжения (рис. 1) первичный контроллер каждого РГ i , $i = 1, \dots, N$ фиксирует опорные напряжения, V_{ni} , от вторичного и регулирует выходное напряжение V_{oi} до желаемой уставки. Обычно это достигается с помощью методов управления по статизму (*droop*

control) [12, 13], но без связи между РГ. Однако одним из перспективных решений может выступать вторичное регулирование напряжения, реализованное посредством согласованной координации доступных контроллеров на основе QU -управления по статизму [14]. Хотя существуют подходы к вторичному управлению на основе моделей [12, 15], их эффективность, как правило, неудовлетворительна из-за упрощений системы, введенных для устранения нелинейности и неопределенных помех.

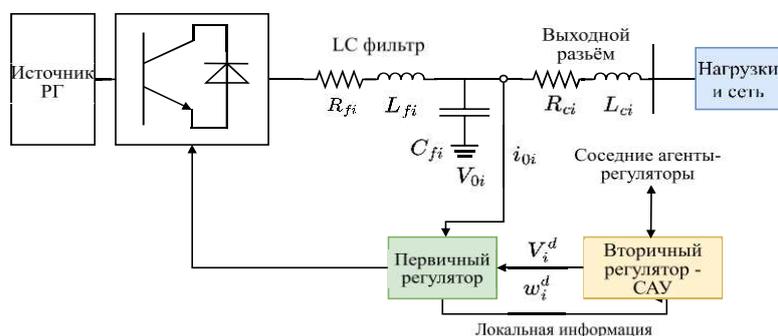


Рис. 1. Блок-схема системы распределенного управления с первичным и вторичным контурами регулирования

Fig. 1. Block diagram of a distributed control system with primary and secondary control loops

Основная мотивация применения управления по статизму на уровне энергорайонов заключается в возможности достаточно легко реализовать децентрализованное пропорциональное управление, и, следовательно, предоставить простую и модульную в реализации схему управления, подобную «plug-and-play», не требующую централизованно-скоординированного управления сетью. В крупномасштабных магистральных сетях управление по статизму обычно применяется только для получения желаемого распределения активной мощности, в то время как амплитуда напряжения на шине генератора регулируется до номинального заданного значения напряжения с помощью АРВ. Однако, в отличие от высоковольтных систем, в компактных энергорайонах линии электропередачи обычно относительно коротки. Поэтому здесь управление по статизму, как правило, применяется для регулирования напряжения с целью достижения желаемого распределения реактивной мощности.

Обоснование использования подобных систем автоматического управления (САУ) по статизму для локальных энергорайонов и микросетей заключается в следующем. Для малых угловых отклонений δ_{ik} следует, что $\sin \delta_{ik} \approx \delta_{ik}$, а $\cos \delta_{ik} \approx 1$. Следовательно, для преимущественно индуктивных

сетей, где $G_{ik} \approx 0$, очевидно, что на реактивную мощность больше всего влияют изменения напряжения. Амплитуды напряжения V_i инверторов меняются в зависимости от отклонений (относительно желаемого значения) реактивной мощности соответственно.

$$u_i^V = V_i^d - k_{Q_i} (Q_i^m - Q_i^d), \quad (1)$$

где u_i^V – управляющий сигнал для амплитуды напряжения; V_i, V_i^d – желаемая (номинальная) амплитуда напряжения; k_{Q_i} – коэффициент усиления по напряжению; Q_i^m – измеренная реактивная мощность; Q_i^d – желаемые уставки реактивной мощности; Q_i – переток реактивной мощности i -го узла.

Для преимущественно индуктивных сетей и для малых угловых отклонений переток реактивной мощности i -го узла Q_i :

$$\begin{aligned} Q_i(\delta_1, \dots, \delta_n, V_1, \dots, V_n) &= G_{ii} V_i^2 + \sum_{k \in N_i} |Y_{ik}| V_i V_k \sin(\delta_{ik} + \varphi_{ik}); \\ Q_i(V_1, \dots, V_n) &= |B_{ii}| V_i^2 + \sum_{k \in N_i} |B_{ik}| V_i V_k \end{aligned} \quad (2)$$

где Q_i – переток реактивной мощности i -го узла; G_{ii} – элемент матрицы проводимости, соответствующий i -му узлу; Y_{ik} – элемент матрицы адмитанса, соответствующий связи между i -м k -м узлами; δ_{ik} – угол между напряжениями i -го и k -го узлов; φ_{ik} – угол между элементом Y_{ik} и δ_{ik} ; B_{ik} – элемент матрицы проводимости, соответствующий i -му узлу; B_{ii} – элемент матрицы проводимости, соответствующий связи между i -м и k -м узлами. Функция Q_i уменьшается от зависимости от углов δ_{ik} и амплитуд напряжения V_i и V_k до зависимости только от амплитуд напряжения V_i и V_k .

Очевидно, что в этом случае реактивной мощностью Q_i можно управлять, управляя амплитудами напряжения V_i и $V_k, \in N_i$. С учетом такого подхода может быть реализовано вторичное управление, целью которого, как известно, является уменьшение отклонения частоты и напряжения электрической сети от их номинальных значений, вызванные первичным управлением. В данной статье предложена реализация распределенного вторичного управления, использующего распределенные протоколы управления, реализованные на отдельных РЭР, которые могут «общаться» друг с другом через распределенную информационно-коммуникационную систему и обмениваться своей локальной информацией с соседними РЭР. Подобное распределенное управление выбирается ω_{ni} и V_{ni} таким образом, чтобы рабочая частота и величина напряжения на выходе каждого РЭР синхронизировались с эталонной частотой и напряжением, ω_{ref} и v_{ref} , т.е.:

$$\begin{cases} \lim_{t \rightarrow \infty} (\omega_i - \omega_{ref}) = 0, & \forall i \in N \\ \lim_{t \rightarrow \infty} (v_{\sigma, mag, i} - v_{ref}) = 0 \end{cases} \quad (3)$$

Вторичное управление гарантирует распределение активной и реактивной мощности РЭР на основе коэффициентов статизма [16, 17] следующим образом:

$$\begin{aligned} m_{P_i} P_i &= m_{P_j} P_j; \\ n_{Q_i} Q_i &= n_{Q_j} Q_j, \end{aligned} \quad (4)$$

где $P_{\max, i}$, $Q_{\max, i}$ и $P_{\max, j}$, $Q_{\max, j}$ – номинальные активная и реактивная мощности i -го и j -го РЭР соответственно.

Вторичное управление энергорайоном, включающего N РЭР, описывается как проблема синхронизации для следующей многокомпонентной системы первого порядка для корректировки первичных входных сигналов управления:

$$\omega_{ni} = V_{ni} = v_i, \quad i = 1, \dots, N; \quad (5)$$

где ω_i^v и v_i^v – распределенные вторичные протоколы управления частотой и напряжением, выбираемые на основе локальной информации каждого РЭР и информации соседних РЭР, которые могут быть записаны как [16]:

$$\begin{aligned} \omega_i &= \omega_i^v - c_\omega \delta_\omega; \\ v_i &= v_i^v - c_v \delta_v, \end{aligned} \quad (6)$$

где c_ω и c_v – коэффициенты управления; δ_ω и δ_v – локальные ошибки отслеживания частоты и напряжения «по соседству», которые могут быть записаны как:

$$\delta_\omega = \sum_{j \in N_i} a_{ij} (\omega_i - \omega_j) + g_i (\omega_i - \omega_{ref}) + \sum_{j \in N_i} a_{ij} (m_{P_i} P_i - m_{P_j} P_j); \quad (7)$$

$$\begin{aligned} \delta_v &= \sum_{j \in N_i} a_{ij} (v_{o, mag, i} - v_{o, mag, j}) + \\ &+ g_i (v_{o, mag, i} - v_{ref}) + \sum_{j \in N_i} a_{ij} (n_{Q_i} Q_i - n_{Q_j} Q_j), \end{aligned} \quad (8)$$

при этом коэффициент закрепления g_i считается ненулевым только для одного РЭР.

III. Мультиагентная система автоматического управления напряжением энергорайона на базе актер-критических нейросетей

Описанный подход к вторичному регулированию на базе системообразующих инверторов и распределенных протоколов управления по статизму обеспечивает согласованное поддержание частоты и напряжения, уменьшение отклонений, а также балансировку активной и реактивной

мощности между РЭР без централизованного диспетчерского вмешательства. Однако эффективность таких систем во многом зависит от заранее заданных параметров и фиксированных коэффициентов управления, что ограничивает их способность адаптироваться к ненормативным возмущениям, меняющейся топологии сети и стохастической генерации распределенных источников. Кроме того, классические протоколы обмена информацией между РЭР предполагают простые правила (условно «жесткую» логику агентов), что не всегда позволяет достигать оптимального распределения ресурсов в сложных и динамичных сценариях. В связи с этим, особый интерес представляет применение интеллектуальных многоагентных систем, реализованных на основе *MARL*, которые могут самостоятельно совершенствовать стратегии управления в ходе эксплуатации, обеспечивая адаптивную координацию инверторов и улучшая устойчивость режима работы энергорайона.

Методы обучения с подкреплением (*Reinforcement Learning, RL*) рассматриваются как важная ветвь машинного обучения для создания автономных самообучающихся агентов, способных взаимодействовать с динамической средой, представленной как марковский процесс принятия решений (МППР) (S, A, T, R) [20]. При этом агент *RL* выбирает действия $a \in A$ в состояниях $s \in S$, и получает вознаграждения $R(s, a, s')$. Цель агента заключается в максимизации суммы дисконтированных вознаграждений в течение времени для поиска оптимальной стратегии управления, π (политики, «*policy*»), что, в общем виде, можно записать в виде выражения (9), согласно которому агент *RL* взаимодействует со средой методом «проб и ошибок».

$$\max_{\pi} E_{\pi} \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t, s_{t+1}) \right], \quad (9)$$

где γ – коэффициент дисконтирования; s_t, a_t – состояние и действие в момент времени t соответственно.

Методы обучения *RL* демонстрирует значительный потенциал для решения этих задач, предлагая адаптивные алгоритмы, способные автоматически оптимизировать управление в условиях стохастичности генерации, изменяющейся топологии и многоагентного взаимодействия РЭР. Зачастую речь идет о т.н. безмодельных алгоритмах *RL*, которые не требуют наличие точной математической модели распределительной сети. Сегодня методология *RL* успешно применяется при управлении нормальными, аварийными и послеаварийными режимами активных распределительных сетей, обеспечивая быстрое регулирование отклонений напряжения и частоты [21], высокие значения индексов надежности (*SAIFI, SAIDI*) при более низких ожи-

даемых затратах [22], учет колебаний напряжения и мощности магистральной сети [20, 23] и т.п.

В последние годы несколько подходов на основе *MARL* были предложены для автономного управления режимом работы локальных энергорайонов и микросетей, работающих как в составе распределительных сетей, так и изолировано [18, 19]. В данной работе был развит безмодельный подход на основе *MARL* для разработки мультиагентной самообучающейся САУ в задаче распределенного вторичного управления по статизму системообразующими инверторами с использованием актер-критических нейросетей [14]. В рамках этого подхода, энергорайон с РГ, подключенными через инверторы, была смоделирована как мультиагентная: $G=(V, E)$, где каждый агент $i \in V$ взаимодействует со своими соседями $N_i: \{j \mid \varepsilon_{ij} \in E\}$. Пусть S и A – глобальные пространства состояний и действий, которые представляют, соответственно, агрегированную информацию о состоянии и координированном управлении для всех агентов-инверторов. Тогда основная динамика электрической сети может быть охарактеризована распределением переходов состояний $P: S \times A \rightarrow [0,1]$.

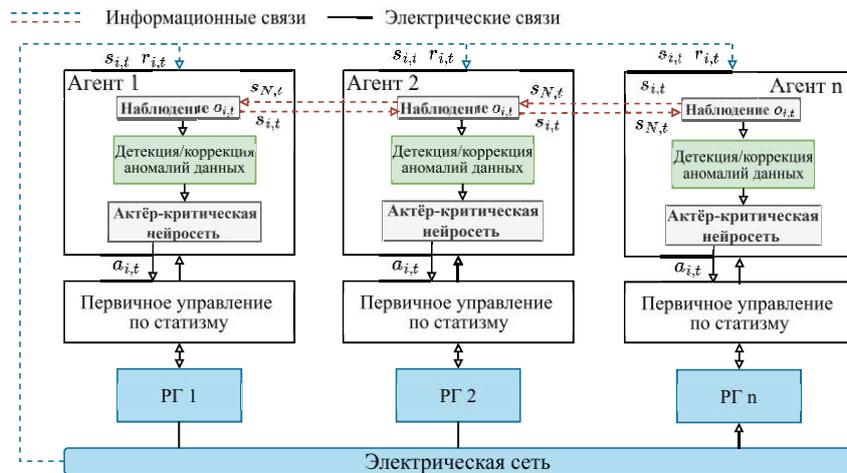


Рис. 2. Общая схема подхода к мультиагентному статическому регулированию напряжения через системообразующие инверторы на базе актер-критического метода *RL*

Fig. 2. General framework of the multi-agent voltage control approach via grid-forming inverters based on the actor-critic reinforcement learning method

В представленной структуре каждый агент-инвертор способен «общаться» только со своими соседями и принимает решение об управлении на основе этих наблюдений (рис. 2). Поскольку каждый агент i наблюдает

только за частью среды (своим состоянием и состояниями своих соседей), это приводит к частичному МППР [24]. В этом случае поставленная задача решается методом *MARL*, где для созданной среды частичного МППР определены следующие ключевые элементы.

1. *Область действий*: управляющее действие для каждого агента-инвертора – уставка вторичного управления напряжением V_n . Были использованы 10 дискретных действий, равномерно распределенных между 1,02 и 1,12 о.е. Общее действие микросети или активной распределительной сети – это совместные действия всех агентов, т.е. $a = v_{n1} \cdot v_{n2} \cdot \dots \cdot v_{nN}$.

2. *Пространство состояний*: состояние каждого агента i выбирается как $s_i = (\delta_i, P_i, Q_i, i_{odi}, i_{oqi}, i_{bdi}, i_{bqi}, v_{bdi}, v_{bqi})$ для характеристики режимов генераторов, подключенных через инверторы, где δ_i – измеренный опорный угол (фаза); P_i, Q_i – активная и реактивная мощности соответственно; $i_{odi}, i_{oqi}, i_{bdi}, i_{bqi}$ – выходные токи d - q генератора i и напрямую подключенные шины, соответственно; а v_{bdi}, v_{bqi} напряжения d - q подключенной шины соответственно.

3. *Пространство наблюдений*: предполагается, что каждый агент может наблюдать только свое локальное состояние, а также сообщения от своих соседей, т.е. $o_{i,t} = S_{i,t} \cup m_{i,t}$, где $m_{i,t}$ – коммуникационное сообщение, полученное от соседних агентов $j \in N$, которое подробнее будет рассмотрено далее;

4. *Вероятности перехода*: вероятность перехода $T(s_0 | s, a)$ характеризует динамику энергорайона с РГ, подключенными через инверторы. В данном случае, для создания платформы моделирования режимов работы распределительной сети был использован безмодельные методы на основе актер-критических сетей, который не требует никаких предварительных знаний о вероятности перехода.

Функция вознаграждения: была использована следующая функцию вознаграждения, чтобы генераторы быстро сходились к эталонным напряжениям (например, 1 отн. ед.):

$$r_{i,t} = \begin{cases} 0,05 - |1 - v_i|, & \text{если } v_i \in [0,8; 0,95] \\ -1|1 - v_i|, & \text{если } v_i \in [0,8; 0,95] \cup [1,05; 1,25] \\ -10, & \text{иначе} \end{cases}, \quad (10)$$

где $r_{i,t}$ – вознаграждение агента i на временном шаге t . Фактически, мы разделяем диапазон напряжений на 3 рабочие зоны: зона нормального режима ($|0,95; 1,05|$ отн.ед.), зона утяжеленного режима $|0,8; 0,95| \cup |1,05; 1,25|$ отн.ед.) и аварийная зона ($|0; 0,8| \cup |1,25; \infty|$ отн.ед.). При сформулированном вознаграждении агенты с «аварийными» напряжениями получают

большой штраф, а агенты с напряжением, близким к 1 отн. ед., получают положительное вознаграждение.

В отличие от традиционных алгоритмов консенсуса, которые обычно страдают от медленной сходимости и неудовлетворительной производительности, предложенная мультиагентная система включает информацию от соседних агентов для повышения эффективности обучения. На каждом шаге времени t агент i обновляет свое скрытое состояние $h_{i,t}$ следующим образом:

$$h_{i,t} = f_i \left(h_{i,t-1}, q_o \left(e_s \left(o_{i,t} \right) \right), q_h \left(h_{N_i,t-1} \right) \right), \quad (11)$$

где $h_{i,t-1}$ – скрытое состояние из предыдущего шага; $o_{i,t}$ – наблюдение агента в момент времени t , т.е. его внутреннее состояние; $h_{N_i,t-1}$ – объединенное скрытое состояние от его соседей; e_s, q_o, q_h – дифференцируемые функции кодирования и извлечения сообщений, где используются однослойные полносвязные слои с 64 нейронами, которые дают возможность каждому агенту адаптироваться к различным объемам и структурам входных данных; f_i – функция кодирования для скрытых состояний и информации о коммуникации, где используется рекуррентная нейросеть с долгой краткосрочной памятью (*LSTM*), которая моделирует временные зависимости, запоминая предыдущие скрытые состояния $h_{i,t-1}$ и взаимодействия, обеспечивая более согласованное поведение.

Чтобы улучшить масштабируемость и устойчивость, наблюдения агентов $o_{i,t}$ соотносятся в соответствии с их физическими единицами. Например, наблюдение $o_{i,t}$ делится на четыре группы: $o_{i,t}^1 \cup o_{i,t}^2 \cup o_{i,t}^3 \cup o_{i,t}^4$ в соответствии с их единицами, а именно: опорный угол, мощность, напряжение и ток. Эти сгруппированные поднаблюдения кодируются отдельно и затем объединяются как:

$$e_s \left(o_{i,t} \right) = \text{cat} \left(e_s^1 \left(o_{i,t}^1 \right), e_s^2 \left(o_{i,t}^2 \right), e_s^3 \left(o_{i,t}^3 \right), e_s^4 \left(o_{i,t}^4 \right) \right), \quad (12)$$

где $e_s^j, j = 1, 2, 3, 4$ – однослойные полносвязные кодирующие слои.

Полученное сообщение $m_{i,t}$ агентом i означает закодированные скрытые состояния его соседей, т.е. $m_{i,t} = h_{N_i,t-1}$, где $h_{N_i,t-1}$ – скрытые состояния соседей агента i в момент времени $t-1$. Поскольку скрытое состояние $h_{i,t-1}$ закодировано нейросетью, эти сообщения гораздо более защищены по сравнению с прямой передачей необработанных состояний. Кодированное наблюдение $e_s(o_{i,t})$ и скрытые состояния соседей $h_{N_i,t-1}$ извлекаются функциями q_o и q_h соответственно. После этого закодированное сообщение объединяется как: $\hat{o}_{i,t} = \text{cat} (q_o(e_s(o_{i,t})), q_h(h_{N_i,t-1}))$. Подобная операция конкатенации

представляется более эффективной по сравнению с операцией суммирования при уменьшении потерь информации.

Полученное скрытое состояние затем используется в актер-критических нейросетях (*Actor-Critic Neural Network*) для генерации случайных действий и предсказания функции ценности, соответственно, т.е. $\pi_{\theta}(h_{i,t})$ и $V_{\omega}(h_{i,t})$. Актер-критическая нейросеть – это архитектура *RL*, в которой обе компоненты, «актер» и «критик», реализованы в виде параметрических нейросетевых моделей. Нейросеть-актер генерирует управляющее действие a_t на основе входных данных наблюдения $\hat{o}_{i,t}$, а нейросеть-критик оценивает ожидаемое качество (ценность) этого действия в заданном состоянии сети. Критик вычисляет градиент ошибки ценности, который используется для обновления весов нейросети-актера, направляя ее обучение к улучшению стратегии. В данной мультиагентной структуре была включена информация о действиях соседей в сеть критика $V_{\omega}(h_{i,t}, a_{Ni,t})$ для улучшения обучения и поток данных между компонентами мультиагентной системы реализован следующим образом (рис. 3).

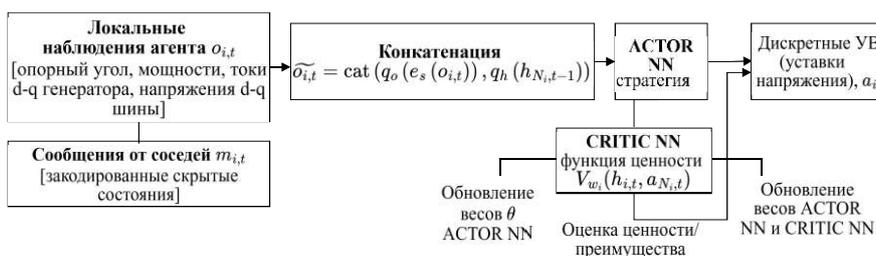


Рис. 3. Общая схема обработки наблюдений и действий в актер-критической нейросети агента *MARL* для вторичного регулирования напряжения

Fig. 3. General scheme of observation and action processing in a *MARL* agent's actor-critic neural network for secondary voltage control

1. Локальные наблюдения агента $o_{i,t}$ (опорный угол, мощности, токи, напряжения) и закодированные сообщения от соседей $m_{i,t}$ объединяются и поступают на вход нейросети-актера (*Actor NN*).

2. Нейросеть-актер выдает дискретное управляющее действие $a_{i,t}$ – выбор новой уставки напряжения инвертора из допустимого набора.

3. Пара (наблюдение, действие), дополненная данными о действиях соседей, передается в нейросеть-критика (*Critic NN*).

4. Критик вычисляет оценку ценности $V_{\omega}(h_{i,t}, a_{Ni,t})$ или функцию преимущества, отражающую ожидаемое влияние данного действия на устойчивость режима и качество напряжения.

5. Полученная оценка используется для корректировки параметров обеих нейросетей: актер обновляет стратегию выбора действий, критик –

модель оценки ценности.

Необходимо отметить, что рассматривалось именно дискретное пространство действий, когда действие выбирается из последнего слоя нейросети как $a_{i,t} \sim \pi_{\theta_i}(h_{i,t})$. Такой подход в итоге позволяет реализовать схему централизованного обучения с децентрализованным исполнением, где у каждого агента есть свои собственные актер-критические нейросети, и их стратегии обновляются независимо, а не согласованно [25], что может лишь ухудшить скорости сходимости. Это также позволяет каждому агенту не только реагировать на локальные изменения режима, но и адаптироваться к поведению соседей, достигая координированного и устойчивого регулирования напряжения в энергорайоне без необходимости иметь точную физическую модель всей сети.

В итоге глобальной целью мультиагентной системы управления максимизация общего глобального вознаграждения $R_{g,t} = \sum_{i \in V} R_{i,t}$, где $R_{i,t} = \sum_{k=0}^{\infty} \gamma^k r_{i,t+k}$ обозначает совокупное вознаграждение для i -агента. Однако подобная математическая постановка связана с типичными проблемами мультиагентного обучения [20]: потери пропускной способности, возможное снижение эффективности обучения, ограничения на количество агентов, медленная сходимость глобального решения. Для решения этих проблем используется фактор пространственного дисконтирования, когда каждый агент i , использует следующее вознаграждение:

$$R_{i,t} = \sum_{k=0}^T \gamma^k \sum_{j \in \mathcal{N}} \alpha(d_{i,j}) r_{i,t+k} \quad (13)$$

где $\alpha(d_{i,j}) \in [0,1]$ – пространственная функция дисконтирования, $d_{i,j}$ – расстояние между агентом i и j . Расстояние может быть евклидовым, характеризующим физическое расстояние между двумя агентами (инверторами) или между двумя вершинами на графе района.

Кроме того, для повышения устойчивости и робастности предложенной схемы мультиагентного статического регулирования к различным сбоям информационно-коммуникационной системы (ИКС) в нее был интегрирована процедура достоверизации входных данных. В предлагаемой процедуре в качестве входных данных x , выступают наблюдения агента $o_{i,t} = S_{i,t} \cup m_{i,t}$, когда, например, потенциальная кибератака на ИКС системы управления может вызывать аномалии как в локальных входных данных агента $S_{i,t}$, так и в сигналах, которые приходят от соседних агентов $m_{i,t}$. В процессе обучения агентов по методу *MARL* параллельно обучается модель изоляционного леса на данных $o_{i,t}$, которые приходят к агентам. В итоге каждому набору $o_{i,t}$ присваивается числовое значение $s(o_{i,t}, m)$, представляющее собой оценку аномалии.

При детекции аномалии в каком-либо наборе $o_{i,t}$ нельзя его просто

исключить, так как это фактически означает исключение контроллера как части мультиагентной САУ из процесса регулирования напряжения. В этом случае, помимо обнаружения аномалий, требуется решение задачи восстановления данных в «поврежденном» наборе $o_{(i,t)}$. Для этого реализуется второй этап процедуры для восстановления поврежденных данных на базе метода k -ближайших соседей. В данном случае понимаются «поврежденные» значения вектора наблюдения агента $o_{i,t} = S_{i,t} \cup m_{i,t}$. Например, после кибератаки могут возникнуть аномалии в каких-либо значениях вектора локального состояния агента $s_i = (\delta_i, P_i, Q_i, i_{odi}, i_{oqi}, i_{bdi}, i_{bqi}, v_{bdi}, v_{bqi})$ (например, в $i_{odi}, i_{oqi}, i_{bdi}, i_{bqi}$) и тогда их восстановление будет зависеть от «нормальных» значениях других параметров вектора, лежащих в окрестности.

III. Экспериментальные исследования

Предложенная модель мультиагентной САУ была использована для децентрализованного управления напряжением на примере локального энергорайона с РЭР, построенного на основе схемы распределительной сети IEEE-34 [14], номинальное напряжение которой составляет 24,9 кВ. Энергорайон характеризуется длинными и малонагруженными двумя рядными регуляторами, линейным трансформатором 20-4 кВ, несимметричной нагрузкой и шунтирующими конденсаторами. Данная схема была разработана для оценки и тестирования алгоритмов решения несбалансированных трехфазных радиальных систем.

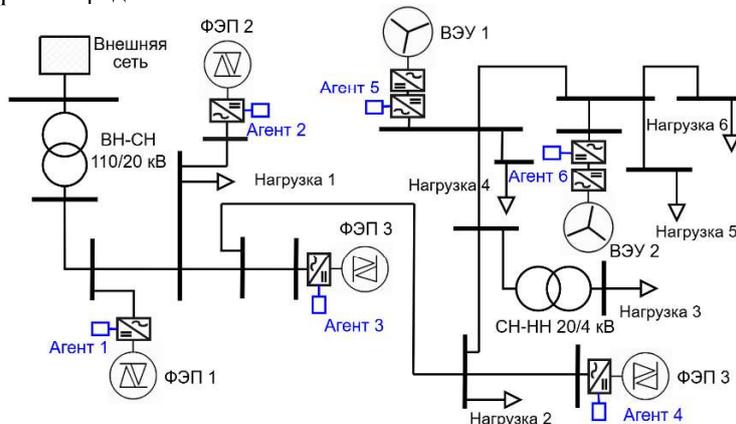


Рис. 4. Энергорайон с РЭР, подключенными через инверторы, на базе модифицированной схемы IEEE-34

Fig. 4. Power district with DERs connected via inverters based on the modified IEEE-34 test feeder

В используемой модификации энергорайон включает в себя следующие РЭР: 4 солнечных ФЭП, подключенных через $DC-AC$ инверторы, 2

ВЭУ, подключенных через *DC-DC*-конвертеры (рис. 4). Основные параметры инверторов, линий и нагрузок сведены в табл. 1.

Таблица 1.
Технические характеристики объектов модифицированной схемы *IEEE-34*

Table 1.
Technical characteristics of modified *IEEE 34* circuit objects

Объект	Исходный параметр		ФЭП 1, ФЭП 2, ФЭП 3, ФЭП 4	ВЭУ 1, ВЭУ 2
Инвертор	Коэффициент статизма по активной мощности, m_p		0,0000565	0,0000750
	Коэффициент статизма по реактивной мощности, n_Q ч		0,052	0,006
	Активное сопротивление фильтра инвертора, R_c (Ом)		0,030	0,030
	Индуктивность фильтра инвертора, L_c (мГн)		0,35	0,35
	Резонансная или частота среза выходного фильтра, ω_c (рад/с)		31,41	31,41
	Пропорциональный коэффициент регулятора, k_p		4	4
	Интегральный коэффициент регулятора, k_i		40	40
	Нагрузка 1	Нагрузка 2	Нагрузка 3	Нагрузка 4
Нагрузки	1,5 Ом	0,5 Ом	1 Ом	0,8 Ом
	1,5 Ом	0,03 Ом	0,05 Ом	0,02 Ом

Подход *MARL* для реализации мультиагентной САУ реализован в среде *Python* с использованием инструментов с открытым исходным кодом для моделирования электрических сетей (*pandapower* и *PowerNet*). Используемая платформа моделирования основана на технических характеристиках линии и нагрузки, описанных в [26]. Для моделирования утяжеленных и аварийных режимов, были добавлены случайные изменения нагрузки по всей сети с отклонениями $\pm 20\%$ от номинальных значений и случайные возмущения в диапазоне $\pm 5\%$ для каждой нагрузки. Все агенты контролировались с дискретностью 0,05 с, и каждый агент мог взаимодействовать со своими соседями через локальные границы связи. Первичное управление нижнего уровня реализовано по аналогии с [27, 28].

Было выполнено сравнение используемого подхода *MARL* с несколькими современными эталонными алгоритмами *MARL*: *IA2L* [29] и *CommNet* [30], чтобы продемонстрировать его эффективность. Каждая модель была обучена более чем на 10 000 эпизодах с $c_\gamma = 0,99$, размер мини-партии $N = 2$, скоростью обучения актера $\eta_\omega = 5 \cdot 10^{-4}$ и скоростью обучения критиков

$\eta_{\omega} = 2,5 \cdot 10^{-4}$. Чтобы обеспечить справедливое сравнение, в каждом эпизоде генерируются разные случайные начальные числа, и в каждом эпизоде одно и то же случайное начальное число используется разными алгоритмами, чтобы гарантировать одинаковую среду обучения/тестирования. Агенты *MARL* контролировались каждые $\Delta T = 0,05$ с (время симуляции) и одна серия длилась $T = 20$ шагов.

На рис. 5а показаны кривые обучения различных алгоритмов *MARL* для рассматриваемой схемы энергорайона, где моделировался мультиагентная САУ. Видно, что используемая мультиагентная система управления на базе предложенного алгоритма *MARL* демонстрирует лучшую сходимость при поиске оптимальной стратегии. После 5000 эпизодов обучения полученная стратегия оценивалась 20 раз при различных нарушениях нагрузки с одним и тем же случайным начальным значением для каждого агента в каждом эпизоде. Результаты этого тестирования представлены на рис. 5б, где показаны профили напряжения для узлов с ФЭП и ВЭУ, подключенными к сети через инверторы, для моделирования одного из утяжеленных режимов схемы (повышение нагрузки на 25 %).

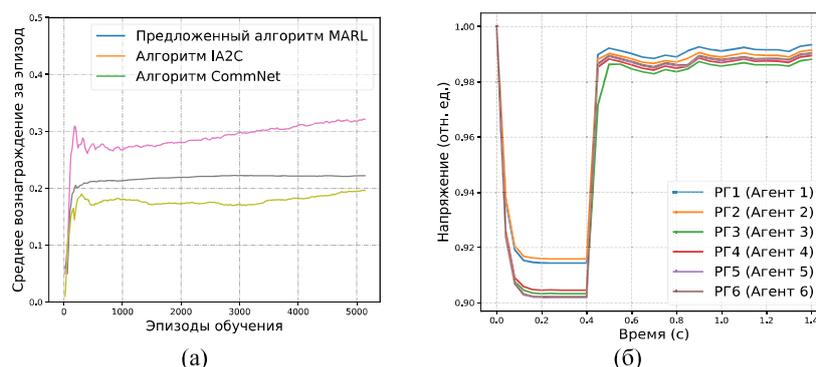


Рис. 5. Результаты тестовых испытаний разработанной мультиагентной системы управления: (а) кривые изменения вознаграждения $R_{g,t}$ при обучении различных алгоритмов *MARL*; (б) изменение напряжения при тестировании мультиагентной САУ на базе предложенного алгоритма *MARL*

Fig. 5. Results of test experiments of the developed multi-agent control system: (a) reward curves $R_{g,t}$ during training of different *MARL* algorithms; (b) voltage variation during testing of the multi-agent automatic control systems based on the proposed *MARL*

Как было отмечено выше, вторичное *QU*-управление направлено на

приведение напряжений всех узлов с подключенными ФЭП и ВЭУ к опорному значению 1 отн.ед. В случае снижения напряжения мультиагентная САУ через 0,4 с после начала возмущения успешно реализует переход из послеаварийного режима в нормальный, когда значения напряжения становятся близки к номинальным (рис. 6). В другом сценарном эксперименте рассмотрены уже случаи возникновения перенапряжений при повышении выработки активной мощности со стороны РЭР. В рамках этого эксперимента на рис. 6 показана возможность предложенного САУ инверторов реализовать функцию автоматического ограничения реактивной мощности в рамках QU -регулирования для предотвращения перенапряжений в сети. Для наглядности ограничения реактивной мощности в зависимости от напряжения в точке подключения ВИЭ была повышена мощность ВЭУ 1 до 16 МВт. В этом случае наблюдается достаточный рост напряжения, чтобы показать, что САУ реагирует на перенапряжения, уменьшая выработку индуктивной реактивной мощности (режим перевозбуждения). Также на левом графике рис. 6 хорошо видно, что реактивная мощность обычно ограничена значением $0,41S_n$, что соответствует коэффициенту мощности $0,925$ в режиме перевозбуждения.

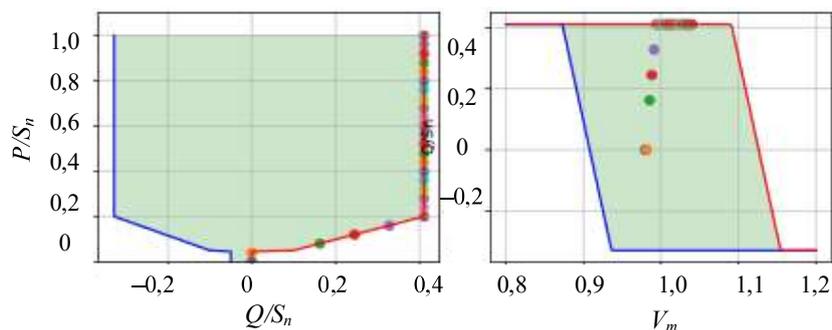


Рис. 6. Применение предложенной мультиагентной САУ инверторов для реализации функции автоматического ограничения реактивной мощности в рамках QU -регулирования для предотвращения перенапряжений в сети

Fig. 6. Application of the proposed multi-agent inverter automatic control systems for automatic reactive power limitation under QU -control to prevent overvoltages in the network

Далее дана большая детализация оценки эффективности предложенных самообучающихся САУ инверторов для автономного QU -регулирования, прежде всего в послеаварийных режимах работы распределительной сети (табл. 2).

Таблица 2.
Результаты моделирования режима работы энергорайона с РЭР
при разных сценариях управления

Table 2.
Results of modeling the operating condition of a power district with DERs
under different control scenarios

Параметр	Базовое управление	Мультиагентное управление <i>MARL</i>	Разница
Минимальное напряжение, отн. ед.	0,9116	0,9773	+0,0457
Загрузка линий, %	104,27	74,08	-30,19
Потери мощности, МВт	0,3801	0,1013	-0,2788

При этом в отношении ВЭУ выполнен учет локального управления выработкой мощности, подробно рассмотренного выше. Анализ результатов квазидинамического моделирования демонстрирует существенное преимущество САУ инверторов по сравнению с базовым управлением: применение адаптивной системы позволило повысить минимальное напряжение в сети на 0,0457 о.е., снизить максимальную загрузку линий на 30,19 % и уменьшить потери мощности на 0,2788 МВт, что подтверждает эффективность предложенного подхода для стабилизации послеаварийных режимов в энергорайонах с РЭР. Наибольший эффект достигнут в разгрузке перегруженных линий, что особенно важно для предотвращения каскадных аварий и повышения надежности энергорайона. Дополнительно эффекты децентрализованного регулирования напряжения на базе мультиагентной САУ продемонстрированы посредством визуализации расчета режима на граф-схеме энергорайона с РЭР для различных экспериментальных случаев (рис. 7).

Отдельно проведенные испытания мультиагентной системы управления при различных информационных сбоях и кибератаках на ее ИКС показали [31, 32], что система обладает высокой робастностью и критично не снижает эффективность вторичного управления. В частности, при успешно реализованной атаке захвата части контроллеров инверторов распределение вероятностей действий агентов меняется, что указывает влияние на выработку ими управляющих воздействий (например, выбираются уставки с меньшим значением напряжения). В этом случае эффективно работает встроенная процедура достоверизации и восстановления искаженных данных [32]. В табл. 3 показан пример одного из наблюдений $o_{i,t}$ «заражаемого» агента-контроллера № 3 для различных сценариев для определенного момента времени моделирования t .

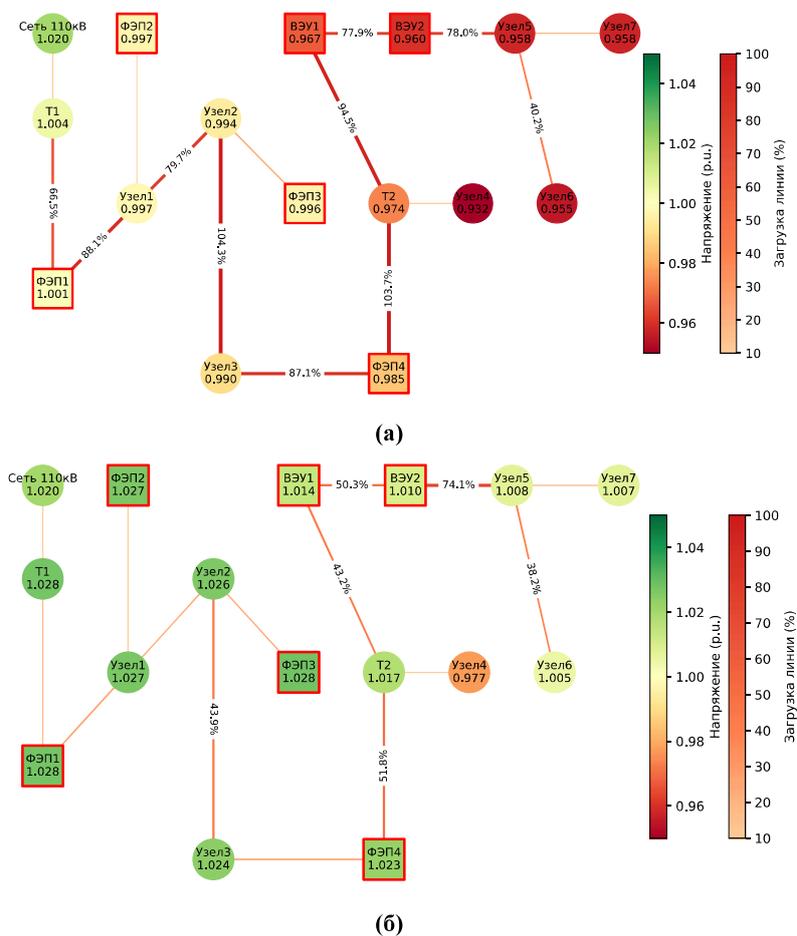


Рис. 7. Визуализация расчета режима на граф энергорайона с РЭР для различных экспериментальных случаев моделирования:
 (а) утяжеленный режим (базовое регулирование);
 (б) утяжеленный режим (с регулирования на базе мультиагентной САУ)

Fig. 7. Visualization of the mode calculation for the power district with DER for various experimental simulation cases:
 (a) heavy loads mode (basic control);
 (b) heavy loads mode (multi-agent automatic control systems)

Таблица 3.

Пример наблюдения $o_{i,t}$ агента-контроллера №3 для различных сценариев

Table 3.

An example of observation $o_{i,t}$ of controller agent №3 for various scenarios

№	Сценарий	Вектор наблюдения агента $o_{i,t}$									
		δ_i	P_i	Q_i	i_{odi}	i_{oqi}	i_{bdi}	i_{bqi}	u_{bdi}	u_{bqi}	
1	Без кибератаки	-2,53	9,18	10,03	4,19	-5,38	-3,42	3,36	-3,45	3,35	
2	Атака «захват контроллера»	-13,6	-7,44	-119	-175	-47,2	-47,1	21,4	-68,1	-75,1	
3	Восстановление данных	-0,17	9,58	7,02	5,02	-3,39	-3,41	3,35	-3,45	3,36	

Хорошо видно, что при атаке «захват контроллера» наблюдение агента по всем параметрам грубо нарушено. Однако применение метода k -ближайших соседей в рамках предложенной процедуры позволяет восстановить «зараженные» данные до вполне близких к потенциально действительным значениям параметров.

Дополнительные испытания также показали [31], что логика управления *MARL* не меняется критически, даже в случаях, когда половина агентов в определенный момент времени «захвачены» злоумышленниками и они начинают пересылать соседним агентам неверную информацию. Это связано с оригинальной процедурой согласованности агентов, описанной выше. В частности, применение нейросетей *LSTM* позволяет агентам запоминать динамику изменений режима и тем самым быть более робастным к временным изменениям в мультиагентной системе, обусловленным кибератаками. В силу того, что каждый агент также имеет свою собственную актер-критическую нейросеть, их стратегии обновляются независимо, а не согласованно, что позволяет быть им относительно автономным и более самостоятельными в случае информационных искажений.

IV. Заключение

Рассмотренный подход к вторичному регулированию напряжения в энергорайонах с РЭР реализует новые принципы сетевого управления, ориентированные на высокочастотные режимы работы современных распределительных сетей. Мультиагентная архитектура на базе актер-критических нейросетей показала способность обеспечивать согласованное поведение инверторов, действующих как системообразующие источники, при наличии неопределенности нагрузок и генерации. В отличие от традиционных многоагентных систем управления, работающих по заранее заданным правилам, предложенный подход на основе методов *MARL* обеспечивает формирование оптимальных стратегий управления непосредственно в процессе обучения. Такая адаптивность позволяет агентам реагировать не

только на текущие возмущения, но и предиктивно учитывать изменения режимов, повышая устойчивость работы электрической сети в непредсказуемых условиях и ненормативных возмущениях.

Проведенные экспериментальные исследования показали, что обмен локальными закодированными состояниями между агентами обеспечивает быстрое восстановление напряжения после аварийных возмущений и уменьшение перегрузок линий. Полученные результаты демонстрируют возможность поддержания уровня напряжения в допустимом диапазоне (0,95-1,05 отн. ед.) и снижения потерь мощности без применения централизованных контроллеров, что подтверждает актуальность децентрализованных решений в условиях роста доли ВИЭ. Результаты тестов на устойчивость к информационным атакам подтвердили, что система сохраняет работоспособность даже при частичном «захвате» агентов, благодаря встроенным процедурам обработки аномалий и независимому обучению актер-критических нейросетей.

Предложенный метод практически применим как для нормальной эксплуатации распределительных сетей, так и для превентивного или аварийного перехода в островной режим. Реализация алгоритма не требует точной математической модели сети, что упрощает его адаптацию к реальным объектам, включая существующие системы управления распределенной генерацией. Результаты работы могут быть использованы при проектировании интеллектуальных систем управления напряжением в распределительных электрических сетях, ориентированных на интеграцию ВИЭ и накопителей, а также для разработки стандартов и регламентов использования децентрализованных средств автоматического регулирования.

Работа выполнена в рамках проектов государственного задания «Централизованно-распределенные многосубъектные киберфизические электроэнергетические системы: Теоретические основы, математические модели и методы формирования, развития, функционирования и управления», (FWEU-2026-0012) и «Фундаментальные математические модели и методы в актуальных энергетических исследованиях с использованием методологии искусственного интеллекта» (FWEU-2026-0015).

© Томин Н.В., 2026

© Сидоров Д.Н., 2026

Поступила в редакцию 05.12.2025

Принята к публикации 05.12.2025

Received 29.12.2025

Accepted 29.12.2025

Библиографический список

- [1] Илюшин П.В. Перспективные направления развития распределительных сетей при интеграции локальных интеллектуальных энергосистем // Электроэнергия. Передача и распределение. 2021. № 4 (67). С. 70-80.
- [2] Илюшин П.В., Куликов А.Л. Автоматика управления нормальными и аварийными режимами энергорайонов с распределенной генерацией. Н. Новгород: Нижегородский институт управления – филиал РАНХиГС, 2019. – 364 с.
- [3] Гуревич Ю.Е., Илюшин П.В. Особенности расчетов режимов в энергорайонах с распределенной генерацией. Н. Новгород: НИУ РАНХиГС, 2018. – 280 с.
- [4] Bidram A., Davoudi A., Lewis F.L. A multiobjective distributed control framework for islanded AC microgrids // IEEE Transactions on Industrial Informatics. 2014. № 10 (3). P. 1785-1798. DOI: 10.1109/TII.2014.2326917
- [5] Dragicevic T., Wu D., Shafiee Q., Meng L. Distributed and decentralized control architectures for converter-interfaced microgrids // Chinese Journal of Electrical Engineering. 2017. № 3 (2). P. 41-52. DOI: 10.23919/CJEE.2017.8048411
- [6] Фишов А.Г., Осинцев А.А., Какоша Ю.В., Одинабеков М.З. Активные распределительные электрические сети с децентрализованным мультиагентным управлением режимом. Ч. 2 // Электричество. 2022. № 11. С. 29-45. DOI: 10.24160/0013-5380-2022-11-29-45
- [7] Армеев Д.В., Фишов А.Г. Мультиагентное децентрализованное регулирование частоты и управление балансом активной мощности в электрических сетях с распределенной генерацией // Релейная защита и автоматизация. 2024. № 4 (57). С. 24-33.
- [8] Бердников Р.Н., Холкин Д.В., Чаусов И.С. Оптимизация систем энергоснабжения удаленных и изолированных территорий за счет управления энергетической гибкостью // Энергетическая политика. 2023. № 1 (179). С. 94-106. DOI: 10.46920/2409-5516_2023_1179.94
- [9] Sidorov D., Panasetky D., Tomin N., Karamov D., Zhukov A., Muftahov I., Dreglea A., Liu F., Li Y. Toward zero-emission hybrid AC/DC power systems with renewable energy sources and storages: A case study from lake Baikal region // Energies. 2020. № 13 (5). P. 1226. DOI: 10.3390/en13051226
- [10] Астапов В.Ю., Ачитаев А.А., Домышев А.В. Мультиагентное управление виртуальной инерцией ветроустановок. Анализ актуальности и эффективности // Электроэнергия. Передача и распределение. 2024. № 4 (85). С. 22-27.
- [11] Zhang K., Yang Z., Basar T. Multi-agent reinforcement learning: a selective overview of theories and algorithms // Studies in Systems, Decision and Control. Vol. 325. Germany: Springer, 2021. P. 321-384. DOI: 10.1007/978-3-030-60990-0_12
- [12] Bidram A., Davoudi A., Lewis F.L., Qu Z. Secondary control of microgrids based on distributed cooperative control of multi-agent systems // IET Generation, Transmission & Distribution. 2013. № 7 (8). P. 822-831. DOI: 10.1049/iet-gtd.2012.0576
- [13] Guerrero J.M., Vasquez J.C., Matas J., De Vicuna L.G., Castilla M. Hierarchical control of droop-controlled AC and DC microgrids – A general approach toward standardization // IEEE Transactions on Industrial Electronics. 2011. № 58 (1). P. 158-172. DOI: 10.1109/TIE.2010.2066534

- [14] Tomin N., Voropai N., Rehtanz C., Kurbatsky V. Management of voltage flexibility from inverter-based distributed generation using multi-agent reinforcement learning // *Energies*. 2021. № 14 (24). P. 8270. DOI: 10.3390/en14248270
- [15] Laaksonen H., Khajeh H., Parthasarathy C., Shafie-khah M. Towards flexible distribution systems: future adaptive management schemes // *Applied Sciences*. 2021. № 11 (8). P. 3709. DOI: 10.3390/app11083709
- [16] Bidram A., Davoudi A., Lewis F.L. Distributed control systems for small-scale power networks: using multiagent cooperative control theory // *IEEE Control Systems Magazine*. 2014. № 34 (6). P. 56-77. DOI: 10.1109/MCS.2014.2350571
- [17] Duan J., Wang C., Xu H., Liu W., Peng J-C., Jiang H. Distributed control of inverter-interfaced microgrids with bounded transient line currents // *IEEE Transactions on Industrial Informatics*. 2018. № 14 (5). P. 2052-2061. DOI: 10.1109/TII.2018.2791988
- [18] Wang S., Duan J., Shi D., Xu C., Li H., Diao R., Wang Z. A data-driven multi-agent autonomous voltage control framework using deep reinforcement learning // *IEEE Transactions on Power Systems*. 2020. № 35 (6). P. 4644-4654. DOI: 10.1109/TPWRS.2020.2990179
- [19] Liu H., Wu W. Online multi-agent reinforcement learning for decentralized inverter-based volt-var control // *IEEE Transactions on Smart Grid*. 2021. № 12 (4). P. 2980-2990. DOI: 10.1109/TSG.2021.3060027
- [20] Сагтон Р.С., Барто Э.Г. Обучение с подкреплением. М.: ДМК-Пресс, 2020. – 552 с.
- [21] Wu Z., Li Y., Zhang X., Zheng S., Zhao J. Distributed voltage control for multi-feeder distribution networks considering transmission network voltage fluctuation based on robust deep reinforcement learning // *Applied Energy*. 2025. Vol. 379. P. 124984. DOI: 10.1016/j.apenergy.2024.124984
- [22] Shang Y., Wu W., Liao J., Guo J., Su J., Liu W., Huang Y. Stochastic maintenance schedules of active distribution networks based on monte-carlo tree search // *IEEE Transactions on Power Systems*. 2020. № 35 (5). P. 3940-3952. DOI: 10.1109/TPWRS.2020.2973761
- [23] Zhong J., Zhang H., Li Y., Zhang C. Service restoration for distribution system via monte-carlo tree search and learned model // 2022 17th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), June 12-15, 2022, Manchester, United Kingdom: IEEE, 2022. DOI: 10.1109/PMAPS53380.2022.9810609
- [24] Hausknecht M., Stone P. Deep recurrent Q-learning for partially observable MDPs // *arXiv.org*. 2015. [Электронный ресурс]. URL: <https://arxiv.org/abs/1507.06527> (дата обращения: 27.11.2025).
- [25] Zhang K., Yang Z., Liu H., Zhang T., Basar T. Fully decentralized multi-agent reinforcement learning with networked agents // 35th International Conference on Machine Learning (ICML 2018), 2018, Stockholm, Sweden: 2018. P. 5872-5881.
- [26] Thurner L., Scheidler A., Schafer F., Menke J-H., Dollichon J., Meier F., Meinecke S., Braun M. Pandapower – An open-source python tool for convenient modeling, analysis, and optimization of electric power systems // *IEEE Transactions on Power Systems*. 2018. № 33 (6). P. 6510-6521. DOI: 10.1109/TPWRS.2018.2829021

- [27] Bidram A., Davoudi A., Lewis F.L., Qu Z. Secondary control of microgrids based on distributed cooperative control of multi-agent systems // *IET Generation, Transmission and Distribution*. 2013. № 7 (8). P. 822-831. DOI: 10.1049/iet-gtd.2012.0576
- [28] Tomin N.V. Robust reinforcement learning-based multiple inputs and multiple outputs controller for wind turbines // *Mathematics*. 2023. № 11 (14). 3242. DOI: 10.3390/math11143242
- [29] Sukhbaatar S., Szlam A., Fergus R. Learning Multiagent Communication with Backpropagation // *Advances in Neural Information Processing Systems (NIPS 2016)*. 2016. Vol. 29. P. 2244-2252.
- [30] Lowe R., Wu Y., Tamar A., Harb J., Abbeel P., Mordatch I. Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments // *Advances in Neural Information Processing Systems (NIPS 2017)*. 2017. Vol. 30. P. 6379-6390.
- [31] Гурина Л.А., Томин Н.В. Обеспечение информационной безопасности при вторичном регулировании напряжения в мультиагентных системах управления киберфизическими микросетями // *Электричество*. 2024. № 10. С. 34-45. DOI: 10.24160/0013-5380-2024-10-34-45
- [32] Гурина Л.А., Томин Н.В. Интеллектуальные методы обеспечения кибербезопасности мультиагентных систем управления микросетями // *Вопросы кибербезопасности*. 2024. № 6 (64). С. 53-64. DOI: 10.21681/2311-3456-2024-6-53-64

References

- [1] P.V. Ilyushin, "Advanced trends of distribution network development when integrating local intelligent power systems", *Electric Power. Transmission and Distribution*, vol. 4, no. 67, pp. 70-80, 2021.
- [2] P.V. Ilyushin and A.L. Kulikov, *Avtomatika upravleniya normal'nymi i avarijnymi rezhimami ehnergorajonov s raspredelennoj generaciej [Automation of control of normal and emergency modes of power districts with distributed generation]*. N. Novgorod: Nizhegorodskij institut upravleniya – filial RANHiGS, 2019 (in Russian).
- [3] Yu.E. Gurevich and P.V. Ilyushin, *Osobennosti raschetov rezhimov v ehnergorajonah s raspredelennoj generaciej [Features of mode calculations in power districts with distributed generation]*. N. Novgorod: NIU RANHiGS, 2018 (in Russian).
- [4] A. Bidram, A. Davoudi and F.L. Lewis, "A multiobjective distributed control framework for islanded AC microgrids", *IEEE Transactions on Industrial Informatics*, vol. 10, no. 3, pp. 1785-1798, Aug. 2014. DOI: 10.1109/TII.2014.2326917
- [5] T. Dragicevic, D. Wu, Q. Shafiee and L. Meng, "Distributed and decentralized control architectures for converter-interfaced microgrids", *Chinese Journal of Electrical Engineering*, vol. 3, no. 2, pp. 41-52, Sep. 2017. DOI: 10.23919/CJEE.2017.8048411
- [6] A.G. Fishov, A.A. Osintsev, Yu.V. Kakosha and M.Z. Odinabekov, "Active power distribution networks with decentralized multi-agent control mode. Part 2", *Elektrichestvo*, no. 11, pp. 29-45, 2022. DOI: 10.24160/0013-5380-2022-11-29-45
- [7] D.V. Armeev and A.G. Fishov, "Multi-agent frequency control in electric grids with distributed generation", *Relay Protection and Automation*, vol. 4, no. 57, pp. 24-33, 2024.

- [8] R.N. Berdnikov, D.V. Khokin and I.S. Chaurov, "Optimization of energy supply systems for remote and isolated territories through energy flexibility management", *Energy policy*, vol. 1, no. 179, pp. 94-106, 2023. DOI: 10.46920/2409-5516_2023_1179.94
- [9] D. Sidorov, D. Panasetsky, N. Tomin, D. Karamov, A. Zhukov, I. Muftahov, A. Dreglea, F. Liu and Y. Li, "Toward zero-emission hybrid AC/DC power systems with renewable energy sources and storages: A case study from lake Baikal region", *Energies*, vol. 13, no. 5, pp. 1226, 2020. DOI: 10.3390/en13051226
- [10] V. Yu. Astapov, A.A. Achitayev and A.V. Domyshev, "Multi-agent control for virtual inertia modes in a wind power plant-relevance and efficiency analysis", *Electric Power. Transmission and Distribution*, vol. 4, no. 85, pp. 22-27, 2024.
- [11] K. Zhang, Z. Yang and T. Basar, "Multi-agent reinforcement learning: A selective overview of theories and algorithms", in *Studies in Systems, Decision and Control. Vol. 325*, Germany: Springer, 2021, pp. 321-384. DOI: 10.1007/978-3-030-60990-0_12
- [12] A. Bidram, A. Davoudi, F.L. Lewis and Z. Qu, "Secondary control of microgrids based on distributed cooperative control of multi-agent systems", *IET Generation, Transmission & Distribution*, vol. 7, no. 8, pp. 822-831, Aug. 2013. DOI: 10.1049/iet-gtd.2012.0576
- [13] J.M. Guerrero, J.C. Vasquez, J. Matas, L.G. De Vicuna and M. Castilla, "Hierarchical control of droop-controlled AC and DC microgrids – A general approach toward standardization", *IEEE Transactions on Industrial Electronics*, vol. 58, no. 1, pp. 158-172, Jan. 2011. DOI: 10.1109/TIE.2010.2066534
- [14] N. Tomin, N. Voropai, C. Rehtanz and V. Kurbatsky, "Management of voltage flexibility from inverter-based distributed generation using multi-agent reinforcement learning", *Energies*, vol. 14, no. 24, pp. 8270, Dec. 2021. DOI: 10.3390/en14248270
- [15] H. Laaksonen, H. Khajeh, C. Parthasarathy and M. Shafie-khah, "Towards flexible distribution systems: Future adaptive management schemes", *Applied Sciences*, vol. 11, no. 8, pp. 3709, Apr. 2021. DOI: 10.3390/app11083709
- [16] A. Bidram, A. Davoudi and F.L. Lewis, "Distributed control systems for small-scale power networks: using multiagent cooperative control theory", *IEEE Control Systems Magazine*, vol. 34, no. 6, pp. 56-77, Dec. 2014. DOI: 10.1109/MCS.2014.2350571
- [17] J. Duan, C. Wang, H. Xu, W. Liu, J-C. Peng and H. Jiang, "Distributed control of inverter-interfaced microgrids with bounded transient line currents", *IEEE Transactions on Industrial Informatics*, vol. 14, no. 5, pp. 2052-2061, May 2018. DOI: 10.1109/TII.2018.2791988
- [18] S. Wang, J. Duan, D. Shi, C. Xu, H. Li, R. Diao and Z. Wang, "A data-driven multi-agent autonomous voltage control framework using deep reinforcement learning", *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4644-4654, 2020. DOI: 10.1109/TPWRS.2020.2990179
- [19] H. Liu and W. Wu, "Online multi-agent reinforcement learning for decentralized inverter-based volt-var control". *IEEE Transactions on Smart Grid*, vol. 12, no. 4, pp. 2980-2990, 2021. DOI: 10.1109/TSG.2021.3060027
- [20] R.S. Sutton and E.G. Barto, *Reinforcement Learning: An Introduction*. Cambridge; London: A Bradford Book, The MIT Press, 2018.

- [21] Z. Wu, Y. Li, X. Zhang, S. Zheng and J. Zhao, "Distributed voltage control for multi-feeder distribution networks considering transmission network voltage fluctuation based on robust deep reinforcement learning", *Applied Energy*, vol. 379, pp. 124984, 2025. DOI: 10.1016/j.apenergy.2024.124984
- [22] Y. Shang, W. Wu, J. Liao, J. Guo, J. Su, W. Liu and Y. Huang, "Stochastic maintenance schedules of active distribution networks based on Monte-Carlo tree search", *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3940-3952, 2020. DOI: 10.1109/TPWRS.2020.2973761
- [23] J. Zhong, H. Zhang, Y. Li and C. Zhang, "Service restoration for distribution system via Monte-Carlo tree search and learned model", in proc. *2022 17th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, June 12-15, 2022, Manchester, United Kingdom. DOI: 10.1109/PMAPS53380.2022.9810609
- [24] M. Hausknecht and P. Stone, "Deep recurrent Q-learning for partially observable MDPs", *arXiv.org*, pp. 7, 2015. [Abstract]. Available at: <https://arxiv.org/abs/1507.06527>.
- [25] K. Zhang, Z. Yang, H. Liu, T. Zhang and T. Basar, "Fully decentralized multi-agent reinforcement learning with networked agents", in proc. *35th International Conference on Machine Learning (ICML 2018)*, 2018, Stockholm, Sweden, pp. 5872-5881.
- [26] L. Thurner, A. Scheidler, F. Schafer, J-H. Menke, J. Dollichon, F. Meier, S. Meinecke and M. Braun, "Pandapower – An open-source python tool for convenient modeling, analysis, and optimization of electric power systems", *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6510-6521, 2018. DOI: 10.1109/TPWRS.2018.2829021
- [27] A. Bidram, A. Davoudi, F.L. Lewis and Z. Qu, "Secondary control of microgrids based on distributed cooperative control of multi-agent systems", *IET Generation, Transmission and Distribution*, vol. 7, no. 8, pp. 822-831, 2013. DOI: 10.1049/iet-gtd.2012.0576
- [28] N.V. Tomin, "Robust reinforcement learning-based multiple inputs and multiple outputs controller for wind turbines", *Mathematics*, vol. 11, no. 14, p. 3242, 2023. DOI: 10.3390/math11143242
- [29] S. Sukhbaatar, A. Szlam and R. Fergus, "Learning Multiagent Communication with Backpropagation", *Advances in Neural Information Processing Systems (NIPS 2016)*, vol. 29, pp. 2244-2252, 2016.
- [30] R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel and I. Mordatch, "Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments", *Advances in Neural Information Processing Systems (NIPS 2017)*, vol. 30, pp. 6379-6390, 2017.
- [31] L.A. Gurina and N.V. Tomin, "Ensuring cyber security in secondary voltage regulation in multi-agent control systems of cyber-physical microgrids", *Elektrichestvo*, no. 10, pp. 34-45, 2024. DOI: 10.24160/0013-5380-2024-10-34-45
- [32] L.A. Gurina and N.V. Tomin, "Intelligent methods of ensuring cybersecurity multi-agent control systems of microgrid", *Cybersecurity issues*, vol. 6, no. 64, pp. 53-64, 2024. DOI: 10.21681/2311-3456-2024-6-53-64

**ИНФОРМАЦИЯ ОБ АВТОРАХ
INFORMATION ABOUT THE AUTHORS**

Томин Никита Викторович, кандидат технических наук, заведующий лабораторией управления функционированием электроэнергетических систем, старший научный сотрудник Института систем энергетики им Л.А. Мелентьева СО РАН, г. Иркутск, Российская Федерация.

Сидоров Денис Николаевич, доктор физико-математических наук, профессор РАН, главный научный сотрудник Института систем энергетики им Л.А. Мелентьева СО РАН, г. Иркутск, Российская Федерация.

Nikita V. Tomin, Cand. Sci. (Eng.), head of electric power systems operation and control lab, senior research fellow of the Melentiev Energy Systems Institute of SB RAS, Irkutsk, Russian Federation.

Denis N. Sidorov, D. Sci. (Math.), professor of RAS, chief research fellow of the Melentiev Energy Systems Institute of SB RAS, Irkutsk, Russian Federation.