УДК 004.056, 303.7.032.4

МЕТОДОЛОГИЯ АНАЛИЗА УГРОЗ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ С ИСПОЛЬЗОВАНИЕМ ЦИФРОВЫХ ДВОЙНИКОВ

И. В. КОТЕНКО[1)]***, И. Б. САЕНКО***[1)]***, Е. С. МИТЯКОВ***[2)] ***В. П. КОЧИН***[3)]

[1)]*Санкт-Петербургский федеральный исследовательский центр РАН,*
14-я линия Васильевского острова, 39, 199178, г. Санкт-Петербург, Россия
[2)]*МИРЭА – Российский технологический университет, пр. Вернадского, 78, 119454, г. Москва, Россия*
[3)]*Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь*

Аннотация. Приводится методология анализа угроз информационной безопасности киберфизических систем на основе цифровых двойников. Предлагаемый подход предусматривает формализацию системы и пространства угроз через многосрезовую структуру, включающую технический, процессный, функциональный, организационный и отраслевой срезы. Далее осуществляется динамическое моделирование угроз в безопасной виртуальной среде цифрового двойника, что позволяет воспроизводить сценарии атак и получать синтетические данные для обучения алгоритмов обнаружения индикаторов угроз. Для выявления аномалий применяются методы частотного анализа, машинного обучения и кластеризации, обеспечивающие адаптивное и точное обнаружение как известных, так и ранее неизвестных атак. Верификация методологии проводится на примере умной энергосети, где показывается эффективность обучения и тестирования алгоритмов на синтетических данных, отражающих реальные и аварийные режимы. Результаты демонстрируют возможность создания самонастраивающихся систем информационной безопасности с высокой степенью адаптивности и точности обнаружения угроз. Представленная методология обеспечивает итеративную обратную связь между этапами, что повышает качество моделирования и обнаружения угроз.

Ключевые слова: киберфизическая система; цифровой двойник; информационная безопасность; моделирование угроз; обнаружение аномалий; машинное обучение; синтетические данные; адаптивная система; умная энергосеть; анализ угроз.

Благодарность. Исследование выполнено при финансовой поддержке Санкт-Петербургского научного фонда (грант № 23-РБ-01-09).

Образец цитирования:
Котенко ИВ, Саенко ИБ, Митяков ЕС, Кочин ВП. Методология анализа угроз информационной безопасности с использованием цифровых двойников. *Журнал Белорусского государственного университета. Математика. Информатика.* 2025;3:76–91.
EDN: RLHKJY

For citation:
Kotenko IV, Saenko IB, Mityakov ES, Kochyn VP. Methodology for information security threat analysis using digital twins. *Journal of the Belarusian State University. Mathematics and Informatics.* 2025;3:76–91. Russian.
EDN: RLHKJY

Авторы:
Игорь Витальевич Котенко – доктор технических наук, заслуженный деятель науки Российской Федерации, профессор; главный научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского института информатики и автоматизации Российской академии наук.
Игорь Борисович Саенко – доктор технических наук, профессор; главный научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского института информатики и автоматизации Российской академии наук.
Евгений Сергеевич Митяков – доктор экономических наук, академик Российской академии естественных наук и Академии инженерных наук имени А. М. Прохорова, профессор; заведующий кафедрой КБ-9 «Предметно-ориентированные информационные системы» Института кибербезопасности и цифровых технологий.
Виктор Павлович Кочин – кандидат технических наук, доцент; проректор по учебной работе и интернационализации образования.

Authors:
Igor V. Kotenko, doctor of science (engineering), honoured scientist of the Russian Federation, full professor; chief researcher at the laboratory of computer security problems, Saint Petersburg Institute for Informatics and Automation, Russian Academy of Sciences.
ivkote@comsec.spb.ru
Igor B. Saenko, doctor of science (engineering), full professor; chief researcher at the laboratory of computer security problems, Saint Petersburg Institute for Informatics and Automation, Russian Academy of Sciences.
ibsaen@comsec.spb.ru
Evgenii S. Mityakov, doctor of science (economics), academician of the Russian Academy of Natural Sciences and the A. M. Prokhorov Academy of Engineering Sciences, full professor; head of the department of KB-9 «Domain-oriented information systems», Institute of Cybersecurity and Digital Technologies.
iyao@mail.ru
Victar P. Kochyn, PhD (engineering), docent; vice-rector for academic affairs and internationalisation of education.
kochyn@bsu.by

METHODOLOGY FOR INFORMATION SECURITY THREAT ANALYSIS USING DIGITAL TWINS

***I. V. KOTENKO*[a], *I. B. SAENKO*[a], *E. S. MITYAKOV*[b], *V. P. KOCHYN*[c]**

[a]*Saint Petersburg Federal Research Center of the Russian Academy of Sciences, 39, 14th Line V. O., Saint Petersburg 199178, Russia*
[b]*MIREA – Russian Technological University, 78 Vernadskogo Avenue, Moscow 119454, Russia*
[c]*Belarusian State University, 4 Niezaliezhnasci Avenue, Minsk 220030, Belarus*

Corresponding author: V. P. Kochyn (kochyn@bsu.by)

Abstract. This paper presents a methodology for analysing information security threats in cyber-physical systems based on digital twins. The proposed approach involves formalising the system and threat space through a multi-layered structure, including technical, process, functional, organisational and sectoral layers. Next, dynamic threat modelling is conducted in a secure virtual environment of the digital twin, enabling the reproduction of attack scenarios and generation of synthetic data to train threat indicator detection algorithms. To identify anomalies, frequency analysis, machine learning and clustering methods are applied, ensuring adaptive and accurate detection of both known and previously unknown attacks. The methodology is verified using a smart grid example, demonstrating the effectiveness of training and testing algorithms on synthetic data that reflect normal and emergency operating modes. The results show the potential for creating self-adjusting information security systems with a high degree of adaptability and threat detection accuracy. The presented methodology provides iterative feedback between stages, enhancing the quality of threat modelling and detection.

Keywords: cyber-physical system; digital twin; information security; threat modelling; anomaly detection; machine learning; synthetic data; adaptive system; smart grid; threat analysis.

Acknowledgements. The study was carried out with the financial support of the Saint Petersburg Science Foundation (grant No. 23-RB-01-09).

Введение

Современные информационные системы (ИС) представляют собой высокосложные распределенные киберфизические комплексы со множеством взаимосвязанных компонентов, функционирующих в условиях высокой неопределенности. Эта неопределенность обусловлена не только возрастанием внутренней архитектурной и поведенческой сложности систем, но и постоянной эволюцией угроз информационной безопасности (ИБ), включая появление ранее не наблюдаемых (*zero-day*) и трудноидентифицируемых атак. В таких условиях особую значимость приобретает задача своевременного выявления индикаторов угроз (косвенных признаков наступления нежелательных событий) до их реализации в виде полномасштабных инцидентов. Эффективное решение данной задачи при ограниченном объеме достоверных эмпирических данных требует построения воспроизводимых моделей поведения защищаемых систем и потенциальных сценариев деструктивного воздействия. Одним из наиболее перспективных инструментов в этом контексте выступает технология цифровых двойников (ЦД) – цифровых репрезентаций объектов и процессов, позволяющих моделировать как нормальное функционирование системы, так и ее реакцию на внешние воздействия, включая реализацию сценариев атак.

Технология ЦД в настоящее время формализована в ряде отечественных и международных стандартов, таких как ГОСТ Р 57700.37-2021, ISO/IEC 30173:2023 и ISO/IEC 20924:2024. Однако следует отметить, что в указанных нормативных документах объектом цифрового моделирования преимущественно выступают физические изделия или технические устройства, в то время как вопросы применения ЦД для решения задач ИБ, в особенности динамического моделирования угроз и генерации данных для построения систем обнаружения аномалий, остаются практически неисследованными.

Кроме того, в современной научной и прикладной деятельности отсутствует единая методология, обеспечивающая логически непрерывный переход от формализованного описания архитектуры и поведения защищаемой системы через моделирование сценариев атак к обучению и верификации средств обнаружения индикаторов нарушений. Данный разрыв между моделированием угроз и последующим построением механизмов их выявления в реальных системах существенно ограничивает воспроизводимость, обоснованность и прикладную значимость разрабатываемых решений в области обеспечения ИБ.

Настоящее исследование направлено на преодоление обозначенного методологического дефицита. Его целью является разработка и экспериментальная верификация методологии анализа угроз ИБ на

основе ЦД, обеспечивающей замкнутый контур перехода от формализации структуры и поведения защищаемой системы и пространства угроз к синтезу ЦД, воспроизведению сценариев атак, генерации синтетических данных и обучению алгоритмов обнаружения аномалий.

Ключевая гипотеза исследования заключается в следующем: алгоритмы детектирования аномалий, обученные исключительно на синтетических данных, полученных в результате моделирования типовых сценариев атак в ЦД, способны эффективно выявлять ранее неизвестные угрозы за счет выделения устойчивых поведенческих паттернов, характерных для соответствующего класса воздействий. Для проверки этой гипотезы реализована экспериментальная установка на основе модели умной энергосети (*smart grid*), подверженной различным типам атак. Экспериментальные результаты демонстрируют применимость и эффективность предложенного подхода, подтверждая его научную новизну и практическую значимость.

Состояние исследований

В настоящем разделе дан аналитический обзор современных научных исследований, посвященных применению ЦД в контексте обеспечения ИБ киберфизических систем (КФС). Основное внимание уделено работам, касающимся моделирования угроз, генерации синтетических данных и обнаружения аномалий.

ЦД представляют собой виртуальные аналоги физических систем, применяемые для мониторинга, прогнозирования и оптимизации процессов в самых разных отраслях – от интеллектуального производства и КФС до строительной сферы и аэрокосмической промышленности [1; 2]. Вместе с тем расширение использования ЦД в распределенных информационно-технических и производственных системах выявляет значительные вызовы в области ИБ [3; 4]. Одним из центральных рисков становится высокая степень интеграции ЦД с информационными и операционными системами, что открывает новые векторы атак, такие как атаки типа «человек посередине», отказ в обслуживании, компрометация данных и несанкционированный доступ, комплексные распределенные атаки [5; 6]. В литературе выделены следующие ключевые направления анализа и защиты ЦД [4–7]: анализ угроз в распределенных промышленных ЦД, разработка защищенных платформ и протоколов, а также превентивное моделирование поведения злоумышленников (*human digital twins*, HDT).

По мнению исследователей [8], для эффективного обеспечения безопасности ЦД должны обладать аналитической предсказуемостью, интегрироваться с физическим объектом и обеспечивать динамическую синхронизацию. Ряд авторов [9] отмечают переход ЦД от оптимизационных функций к роли инструмента проактивной ИБ, что связано с возможностями безопасного моделирования атак, тестирования защиты, прогнозирования последствий инцидентов и автоматизации обнаружения аномалий, визуализации информации [10].

В литературе выделяются два основных режима функционирования ЦД в процессе моделирования кибератак [11]: режим репликации, при котором модель синхронизируется с физической системой в реальном времени, и режим изолированного моделирования, позволяющий безопасно выполнять сценарии атак. Количественная оценка рисков и моделирование влияния атак, основанные на динамических байесовских сетях и марковских процессах, подробно рассмотрены в ряде работ (см., например, [12]). Параллельно развиваются игровые среды для обучения ИБ-специалистов реализации сценариев атак и защиты [11].

Отдельное место в исследованиях занимают задачи обнаружения аномалий с использованием ЦД: сравнение прогнозируемого и фактического поведения системы позволяет выявлять вторжения и сбои [13]. HDT-технологии расширяют аналитику, добавляя в моделирование поведенческие паттерны человека [14].

Тем не менее использование синтетических данных, полученных из ЦД, для обучения алгоритмов обнаружения аномалий сопряжено с рядом проблем. Во-первых, ограниченная обобщающая способность упрощенных моделей ЦД снижает реализм выборок [1]. Во-вторых, доменный разрыв между синтетическими и реальными данными, включая распределительные смещения, шум и неопределенность, существенно снижает переносимость моделей [15]. Кроме того, отмечаются структурные и нормативные уязвимости ЦД, отсутствие единых архитектур и дополнительные векторы атак, компрометирующие надежность синтетических данных [16; 17]. В статье [18] предложено структурированное моделирование угроз на основе графовых и таксономических моделей, интегрирующих кибернетические и физические аспекты атак. В настоящей работе представлено развитие данного подхода путем его адаптации к многосрезовому моделированию в ЦД, что позволяет охватить не только технические, но и организационные, процессные и отраслевые аспекты системы.

Онтологические и агентно-ориентированные модели, такие как концептуальная структура «Cybonto» (*Cybonto conceptual framework*) [14], дополняют этот подход когнитивным компонентом, позволяя модели-

ровать поведение потенциальных нарушителей. Технические обзоры (см., например, [19]) подчеркивают важность формализации сценариев атак в мультисистемных ЦД-архитектурах и адаптивных системах реагирования, способных обобщать результаты с учетом междисциплинарных связей [18].

Еще один значимый вызов – это обеспечение интероперабельности и стандартизации платформ ЦД, особенно в контексте ИБ. Разнородность форматов, протоколов и семантических моделей мешает унификации взаимодействия. Усложнение возникает из-за потребности в семантическом согласовании, защите и контроле конфиденциальности, а также в управлении жизненным циклом ЦД с учетом отраслевых особенностей [20–22]. Среди стратегий решения обозначенной проблемы можно выделить стандарты ISO/IEC JTC 1/SC 41, ISO 23247, документ NIST IR 8356, а также рекомендации Консорциума цифровых двойников (*Digital Twin Consortium*).

Несмотря на значительный прогресс в указанных направлениях, в научной литературе отсутствует единая методология, обеспечивающая сквозной цикл от формализации системы и пространства угроз до генерации синтетических данных в ЦД, обучения алгоритмов обнаружения аномалий и их верификации. В настоящей работе авторы предлагают такую методологию, она представлена в следующих разделах.

Концептуальная схема методологии

Предлагаемый подход базируется на логически выстроенной последовательности этапов, объединенных одной целью – обеспечить переход от формализации КФС и пространства угроз к выявлению индикаторов реализации этих угроз на основе данных, полученных в ходе безопасного динамического моделирования угроз в ЦД. Рассматриваемая методология носит сквозной и итеративный характер, обеспечивая замкнутый контур анализа, в котором каждая последующая фаза уточняется на основе результатов предыдущей фазы.

В структурном виде методология включает четыре взаимосвязанных этапа.

Этап 1: формализация системы и пространства угроз. Создается формализованная многосрезовая модель ключевых аспектов КФС, отражающая ее архитектуру, функции, процессы и отраслевую специфику. Угрозы соотносятся с компонентами модели с учетом требований к конфиденциальности, целостности и доступности. Полученная модель служит основой для построения ЦД.

Этап 2: моделирование угроз в ЦД. На основе формализованной модели в виртуальной среде реализуются сценарии атак с учетом их динамики и последствий. ЦД синхронизируется с параметрами реальной системы. Результатом этапа является набор синтетических данных, отражающих как нормальное поведение системы, так и ее реакции на реализованные угрозы.

Этап 3: выделение индикаторов угроз. Данные, полученные при моделировании угроз в ЦД, используются для обучения алгоритмов обнаружения индикаторов угроз. Применяются методы частотного анализа и машинного обучения. Извлекаются устойчивые поведенческие паттерны, которые служат основой для построения адаптивных систем ИБ.

Этап 4: верификация подхода. Обученные алгоритмы проходят тестирование на новых сценариях атак и прототипе системы. Оценивается их способность выявлять как известные, так и ранее неизвестные угрозы. Полученные результаты используются для уточнения модели, сценариев и параметров моделирования угроз.

Методология предусматривает обратные связи, обеспечивающие адаптивность и настройку всех компонентов системы. Результаты этапа 3 (выделение индикаторов угроз) и этапа 4 (верификация подхода) используются:

- для пересмотра и уточнения формализованной модели системы и пространства угроз (этап 1);
- актуализации сценариев и параметров моделирования угроз (этап 2);
- совершенствования архитектуры ЦД за счет повышения качества моделирования и генерации данных.

Схематично рассматриваемая методология представлена на рис. 1.

ЦД выступает ядром методологии, обеспечивая интеграцию всех этапов в единый итеративный процесс. Он представляет собой безопасную виртуальную среду для динамического моделирования сценариев атак, генерации синтетических данных и тестирования алгоритмов обнаружения аномалий. Благодаря двунаправленной синхронизации с реальной системой ЦД поддерживает актуальность модели и позволяет адаптировать сценарии атак с учетом результатов анализа.

Результатом применения методологии является комплекс моделей и инструментов: формализованная многосрезовая модель системы и пространства угроз, синтетические данные, обученные алгоритмы обнаружения аномалий и выделенные индикаторы угроз. Методологический каркас объединяет направления моделирования угроз, анализа данных, обеспечивает основу для построения самообучающихся адаптивных систем ИБ.

Рис. 1. Схема методологии
Fig. 1. Methodology scheme

Формализация системы и пространства угроз

Формализация КФС и пространства угроз является ключевым этапом методологии анализа угроз ИБ с использованием ЦД. На этом этапе формируется воспроизводимая модель, объединяющая архитектуру системы, ее поведенческие особенности и потенциальные векторы атак, что закладывает основу для цифрового моделирования и генерации синтетических данных.

Полученная модель описывает защищаемую систему через пять взаимосвязанных срезов:

• технический срез (S_T), включающий аппаратные средства, встроенные устройства, сети передачи данных, программное обеспечение и средства защиты;

• процессный срез (S_P), охватывающий эксплуатационные процедуры, сценарии функционирования, мониторинг и реагирование на инциденты;

• функциональный срез (S_F), отражающий назначение компонентов, их взаимодействие и участие в реализации функций управления и безопасности;

• организационный срез (S_O), включающий роли, ответственность, внутренние регламенты и механизмы управленческого контроля;

• отраслевой срез (S_I), учитывающий особенности применения системы в конкретной предметной области, включая нормативные требования и характерные угрозы.

Каждому компоненту системы $c_i \in S_k$, $k \in \{T, P, F, O, I\}$, сопоставляется множество угроз $U(c_i) \subseteq U$, где U – общее множество рассматриваемых угроз, сформированное на основе авторитетных классификаций (например, Банка данных угроз безопасности информации Федеральной службы по техническому и экспортному контролю (далее – БДУ ФСТЭК России)). Для удобства и формального анализа вводится бинарная матрица соответствия $M \in \{0, 1\}^{n \times m}$, где $n = \sum_k |S_k|$ – общее количество компонентов во всех срезах, а m – количество учитываемых угроз. Значение $M_{ij} = 1$ указывает на наличие связи между компонентом c_i и угрозой u_j, а значение $M_{ij} = 0$ – на ее отсутствие.

В описанной модели каждый срез S_k представляет собой формализованное представление защищаемой КФС в пределах одного из структурных срезов – технического, процессного, функционального, организационного или отраслевого. Однако реальные угрозы ИБ часто затрагивают несколько срезов одновременно. Например, реализация уязвимости в программном обеспечении может не только нарушить техническую целостность системы, но и повлиять на процессы эксплуатации и реагирования.

Для формализации взаимосвязей между срезами вводится матрица влияний $V \in [0, 1]^{5\times 5}$, где каждый элемент V_{ab} отражает долю угроз, которые одновременно затрагивают срезы S_a и S_b. Она вычисляется следующим образом:

$$V_{ab} = \frac{\left|\left\{u \in U \mid \exists c_a \in S_a,\ c_b \in S_b : M_{c_a,u} = 1 \wedge M_{c_b,u} = 1\right\}\right|}{\left|\left\{u \in U \mid \exists c_a \in S_a : M_{c_a,u} = 1\right\}\right|}.$$

В числителе находится количество угроз u, которые одновременно воздействуют хотя бы на один компонент из среза S_a и хотя бы на один компонент из среза S_b, а в знаменателе – количество угроз, воздействующих хотя бы на один компонент из среза S_a.

Матрица влияний $V = [V_{ab}]$ позволяет оценить, насколько реализация угрозы в одном аспекте системы (например, техническом) может повлиять на другой аспект системы (например, процессный). Практическое значение показателя V_{ab} заключается в выявлении каскадных путей распространения угроз, что критически важно для построения многоуровневой защиты. Например, если значение V_{TP} (технический срез $\rightarrow$ процессный срез) равно 0,7, это означает, что 70 % угроз, связанных с техническими компонентами, также воздействуют на эксплуатационные процессы. Данная информация позволяет выявить потенциальные каскадные пути распространения атак и принять упреждающие меры на других уровнях системы. Таким образом, значение V_{ab} отражает условную вероятность того, что угроза, влияющая на срез S_a, затрагивает и срез S_b. При обнаружении аномалии, например, в технической подсистеме это позволяет автоматически активировать мониторинг процессных и организационных процедур, предотвращая каскадный эффект.

Предложенная формализация обеспечивает основу для последующего динамического моделирования угроз в ЦД, где учет межсрезовых зависимостей критически важен для построения реалистичных сценариев атак и анализа поведения системы в различных условиях.

Моделирование угроз в ЦД

Следующим этапом методологии является динамическое моделирование угроз ИБ с использованием ЦД исследуемой КФС. На основе модели, сформированной на этапе формализации системы и пространства угроз, создается структура ЦД, включающая компоненты системы, связи между ними и соответствующие классы угроз.

Архитектура ЦД базируется на принципе двунаправленной синхронизации с реальной системой, что обеспечивает передачу актуальных данных о состоянии компонентов системы и получение результатов симуляций в режиме, максимально приближенном к реальному времени. Такая синхронизация гарантирует актуальность и достоверность моделирования, повышая его практическую значимость.

Моделирование реализуется через формирование и проигрывание сценариев атак, соответствующих выявленным угрозам и отраслевой специфике. Сценарии атак строятся на основе следующих факторов:

- таксономий угроз, сформированных в процессе формализации;
- межсрезовых связей угроз, определенных с использованием матрицы влияний V;
- информации о компонентах и процессах КФС.

Ключевым элементом предлагаемого подхода является расширенная динамическая модель угроз, формализуемая функцией

$$T_{\text{extended}} = F\left(T_{\text{base}},\ DT_{\text{impact}}\right),$$

где T_{base} – базовая (статическая) модель угрозы, включающая описание атаки (вектор атаки, цель, эксплуатируемые уязвимости, необходимые условия и технические характеристики); DT_{impact} – динамический компонент, формализующий дополнительную информацию, полученную в результате симуляций и анализа данных в ЦД (динамика развития сценариев атак, реакции системы, поведение защитных механизмов, а также последствия реализации угроз); F – функция интеграции, которая объединяет статическое описание атаки с результатами моделирования угроз в ЦД, формируя расширенную динамическую модель угрозы T_{extended}.

Базовая модель угрозы T_{base} может быть формализована как следующий кортеж:

$$T_{\text{base}} = \langle \text{вектор атаки, цель, уязвимость, условия, технические характеристики} \rangle.$$

Например, для атаки типа «ложные команды управления» в умной энергосети параметры кортежа могут быть следующими:

- вектор атаки – MITM (*man-in-the-middle*);
- цель – контроллер распределенного генератора;
- уязвимость – отсутствие аутентификации команд;
- условия – активность SCADA-сессии;
- технические характеристики – протокол Modbus/TCP.

Динамический компонент DT_{impact} представляет собой временную траекторию развития инцидента, зафиксированную в ЦД, и включает:

- временные метки начала и пика развития атаки;
- реакцию защитных механизмов (например, срабатывание IDS);
- каскадные эффекты (например, отключение смежных узлов);
- изменение ключевых параметров системы (напряжение, частота, задержки).

Функция интеграции F реализуется как расширение статической модели T_{base} с помощью вектора динамических параметров DT_{impact}:

$$T_{\text{extended}} = T_{\text{base}} \cup \left\{ t_{\text{start}}, t_{\text{peak}}, \Delta V(t), ID_{\text{affected}} \right\},$$

где t_{start} – временная метка начала атаки; t_{peak} – временная метка пика развития атаки, указывающая на момент максимального воздействия угрозы на систему; $\Delta V(t)$ – функция изменения напряжения во времени (в контексте умной энергосети), отражающая динамику воздействия угрозы на параметры системы; ID_{affected} – идентификаторы затронутых компонентов системы, позволяющие определить, какие элементы инфраструктуры подверглись воздействию.

Практическая ценность такой модели заключается в генерации обогащенных сценариев атак для обучения систем обнаружения индикаторов угроз, а также в количественной оценке эффективности мер защиты (например, время срабатывания защиты, глубина распространения угрозы).

Выделение индикаторов угроз

Следующим этапом методологии является разработка системы обнаружения индикаторов угроз – аномалий, свидетельствующих о реализации атак. Основная задача данного этапа состоит в обучении алгоритмов обнаружения характерных признаков угроз на основе синтетических данных, сгенерированных в ЦД.

В отличие от традиционных подходов, базирующихся преимущественно на реальных данных или экспертных оценках, предложенная методика использует синтетические, но достоверные и контролируемые наборы данных, получаемые в безопасной виртуальной среде. Такой подход обеспечивает широкое разнообразие обучающих примеров, включая редкие и ранее неизвестные сценарии атак, что значительно повышает универсальность и адаптивность алгоритмов обнаружения аномалий.

Для выявления аномалий используются взаимодополняющие методы трех типов:

1) анализ спектральных характеристик сигналов, который дает возможность обнаруживать нестандартные изменения в параметрах работы системы;

2) обучение моделей распознавать нормальные паттерны поведения и выявлять отклонения, характерные для атак;

3) кластеризация, позволяющая сгруппировать данные по типичным режимам функционирования и выделить элементы, не попадающие в эти группы, как потенциальные аномалии.

Архитектура системы обнаружения индикаторов угроз построена по принципу многокомпонентного анализа (рис. 2). Входной поток данных, поступающий как из реальной системы, так и из ЦД, проходит этап предобработки, после чего параллельно анализируется несколькими алгоритмами. Полученные результаты подвергаются агрегации и классификации. При классификации аномалий используется информация о видах угроз, сформированная на этапах формализации системы и пространства угроз и моделирования угроз в ЦД, что обеспечивает точное сопоставление выявленных аномалий с конкретными классами атак.

Обучение на данных, сгенерированных в ЦД, существенно расширяет возможности системы по сравнению с традиционными методами обучения, обеспечивая адаптивность к редким и ранее неизвестным видам атак.

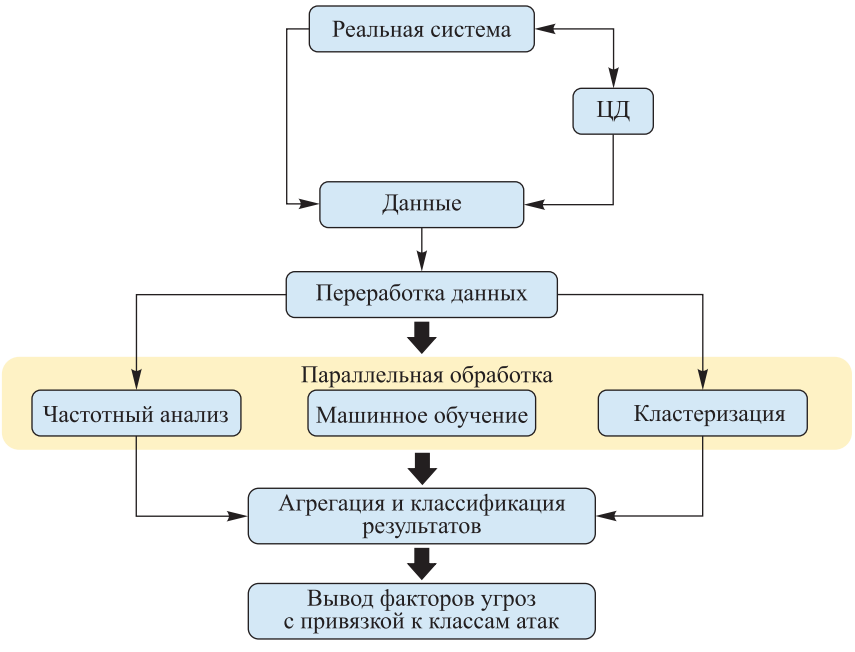

Рис. 2. Архитектура системы обнаружения индикаторов угроз
Fig. 2. Architecture of the threat indicator detection system

Верификация подхода

Экспериментальная апробация предложенной методологии, направленная на ее верификацию, проводилась на примере умной энергосети. Цель этого этапа – подтвердить, что ЦД не только позволяет моделировать сценарии атак, но и генерирует данные, пригодные для обучения алгоритмов обнаружения аномалий, способных эффективно выявлять ранее неизвестные угрозы.

Современные угрозы ИБ, включая вредоносное программное обеспечение, фишинг, DDoS-атаки и целенаправленные операции, представляют серьезную опасность для умной энергосети из-за высокой степени цифровизации и сетевой взаимосвязанности компонентов. Ключевая проблема состоит в том, что кибератаки могут маскироваться под естественные искажения, вызванные, например, нелинейными потребителями, погодными условиями или аппаратными сбоями. Это обстоятельство затрудняет их обнаружение и разграничение с неугрожающими отклонениями, создавая риски для устойчивости и безопасности системы [23].

В соответствии с предложенной методологией на первом этапе апробации подхода была построена формализованная многосрезовая модель умной энергосети, включавшая:

- технический срез (аппаратные и программные компоненты – датчики, контроллеры, SCADA, коммуникационная инфраструктура и средства защиты);
- процессный срез (процессы мониторинга, управления нагрузкой, автовосстановления и реагирования на инциденты);
- функциональный срез (функции регулирования напряжения и частоты, балансировки нагрузки и передачи данных);
- организационный срез (роли персонала, регламенты доступа и политики ИБ);
- отраслевой срез (нормативные требования, включая ГОСТ Р ИСО/МЭК 27019-2021, и характерные угрозы для энергетики).

Далее на основе этой модели была построена бинарная матрица соответствия между компонентами системы и угрозами из БДУ ФСТЭК России. Анализ показал, что 78 % угроз, связанных с подменой команд, искажением данных с датчиков и отказами в обслуживании, затрагивают компоненты, участвующие в измерении и регулировании напряжения (в частности, датчики, контроллеры и каналы передачи данных в SCADA-системе). Исходя из этого, можно заключить, что реализация таких угроз с высокой вероятностью приведет к отклонению напряжения от нормы или к его некорректной регистрации. Следовательно, временной ряд напряжения становится чувствительным индикатором кибератак, поскольку в нем отражаются как прямые, так и косвенные (например, искаженные данные, используемые для управления) последствия атак.

На основе этой структуры был построен ЦД, реализованный в виде виртуальной имитационной среды на языке Python с использованием библиотек SimPy, Scapy, Pandas и NumPy. ЦД поддерживает двунаправленную синхронизацию с реальной системой и используется для моделирования сценариев кибератак и аварийных состояний.

Для анализа поведения системы и выявления аномалий был разработан экспериментальный комплекс, интегрированный с ЦД и включающий:

- сбор и предварительную обработку сигналов;
- применение алгоритмов обнаружения аномалий;
- агрегирование результатов и визуализацию;
- автоматическое реагирование на события.

Такая архитектура позволила использовать ЦД в качестве безопасной тестовой платформы, обеспечивающей гибкость и контролируемость эксперимента в условиях, приближенных к реальному времени, но без риска для физической инфраструктуры.

В основу подхода легли три ключевых направления анализа:

- частотный анализ с вейвлет-преобразованием (используется на этапе предобработки для выявления характерных отклонений в сигналах);
- анализ отклонений от штатного поведения (реализуется путем обучения моделей на данных нормального функционирования системы);
- кластеризация режимов работы (заключается в формировании устойчивых кластеров безопасных состояний, где все, что выходит за их пределы, рассматривается как потенциальная аномалия).

Представим поэтапную процедуру апробации предложенной методологии.

Этап 1: генерация в ЦД данных для обучения. На этом этапе эксперимента в ЦД была смоделирована работа энергосистемы в трех режимах:

- нормальном режиме (напряжение сети описывалось синусоидальным сигналом с добавлением случайного шума, что отражало естественные колебания в системе при штатном функционировании);
- режиме реализации угроз (кибератаки) (имитировались резкие скачки напряжения, характерные для кибератак, направленных на дестабилизацию сети; эти сценарии представляли собой сгенерированные паттерны аномалий – индикаторов угроз);
- аварийном режиме (осуществлялось плавное, но устойчивое изменение амплитуды сигнала, что моделировало сбои и технические неисправности, не связанные с вредоносными действиями).

На рис. 3 представлена иллюстрация одного из сгенерированных временных рядов напряжения с маркировкой аномалий, соответствующих различным режимам работы системы.

Каждый режим моделировался многократно (100 раз) с варьированием параметров, что позволило сформировать датасет из нескольких наборов синтетических сигналов.

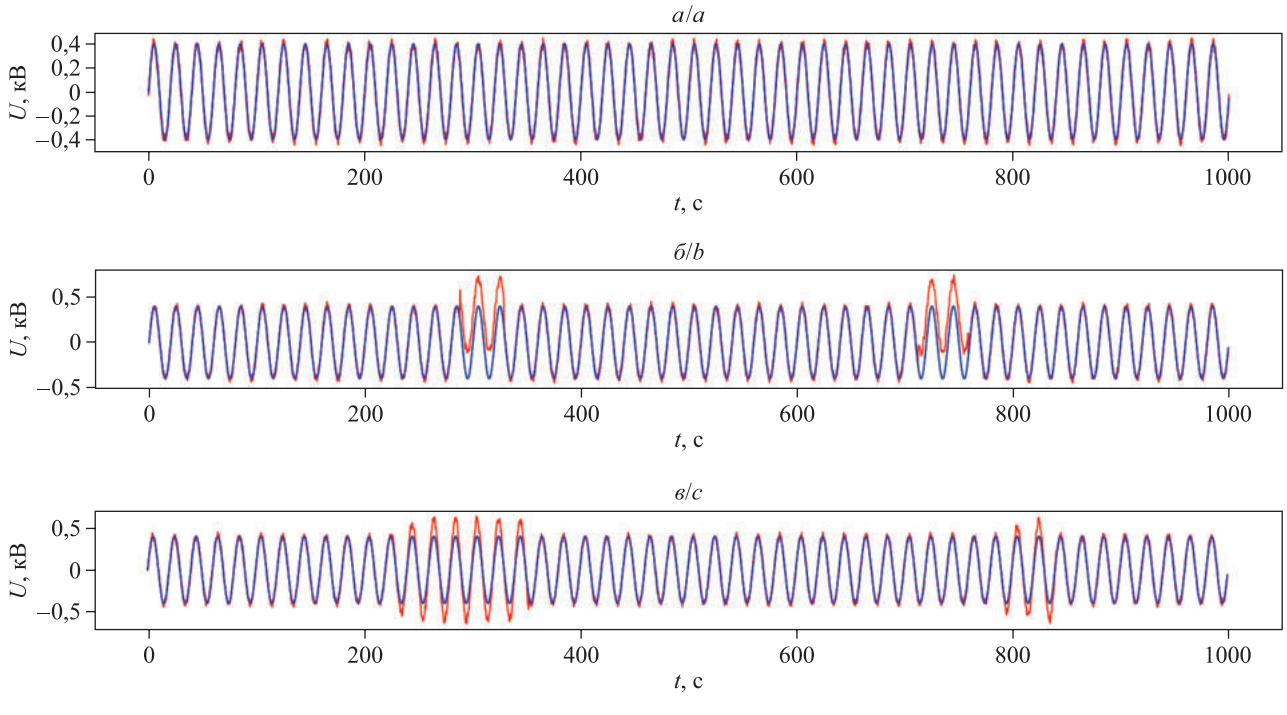

Рис. 3. Сгенерированный временной ряд напряжения с маркировкой аномалий (выделены красным цветом), соответствующих различным режимам работы системы:
а – нормальный режим; *б* – режим кибератаки; *в* – аварийный режим

Fig. 3. Generated voltage time series with anomalies (highlighted in red) corresponding to different system operating modes:
a – normal mode; *b* – cyberattack mode; *c* – emergency mode

Этап 2: предобработка и фильтрация данных. После генерации исходных данных в ЦД все временные ряды проходили этап предобработки. Для этой цели применялось двухуровневое дискретное вейвлет-преобразование с использованием различных типов вейвлетов, что позволило выбрать оптимальный метод фильтрации для последующего обнаружения аномалий. В ходе экспериментов были протестированы несколько типов вейвлетов, включая вейвлеты Добеши (db), симлеты (sym) и коифлеты (coif). Для иллюстрации результатов в статье выбран вейвлет Добеши 1 (db1), который показал хорошие результаты при обработке сигналов напряжения в энергосистеме. Вейвлет Добеши 1 особенно эффективен для обработки сигналов с резкими переходами и разрывами, что характерно для кибератак на умные энергосети.

Особое внимание уделялось аппроксимирующим коэффициентам второго уровня (A_2), которые представляют собой низкочастотную компоненту сигнала напряжения, содержащую информацию об основной тенденции работы энергосистемы. В отличие от исходного сигнала коэффициенты A_2 представляют собой «сглаженную» версию напряжения, где удалены высокочастотные шумы и кратковременные колебания, что позволяет фокусироваться на устойчивых характеристиках системы. В контексте энергосети коэффициент A_2 отражает базовую форму напряжения при нормальной работе и является чувствительным индикатором системных изменений, так как кибератаки и аварии часто влияют именно на основные характеристики сигнала, а не только на высокочастотные шумы.

На рис. 4 показан результат обработки сигнала напряжения, представленного на рис. 3, методом дискретного вейвлет-преобразования с использованием вейвлета Добеши 1. Визуально форма преобразованного сигнала близка к исходной, поскольку фильтрация сохраняет низкочастотную компоненту (основную волну), но удаляет высокочастотные колебания и случайные возмущения. Эти изменения не всегда заметны невооруженным глазом, однако они влияют на частотную структуру сигнала, что критично для автоматической классификации: сглаженный ряд позволяет алгоритмам надежно выделять аномальные отклонения, характерные для кибератак и аварийных состояний, при минимальном влиянии шумов.

Для последующего анализа и визуализации работы алгоритмов обнаружения аномалий были сформированы признаковые пространства, где по одной оси откладывались значения коэффициента A_2 в нормальном режиме работы системы, а по другой оси – значения коэффициента A_2 в аномальном состоянии (режим кибератаки или аварийный режим). Такой подход позволил четко визуализировать изменения в основных характеристиках сигнала и выявить моменты, когда поведение системы существенно отклонялось от нормы.

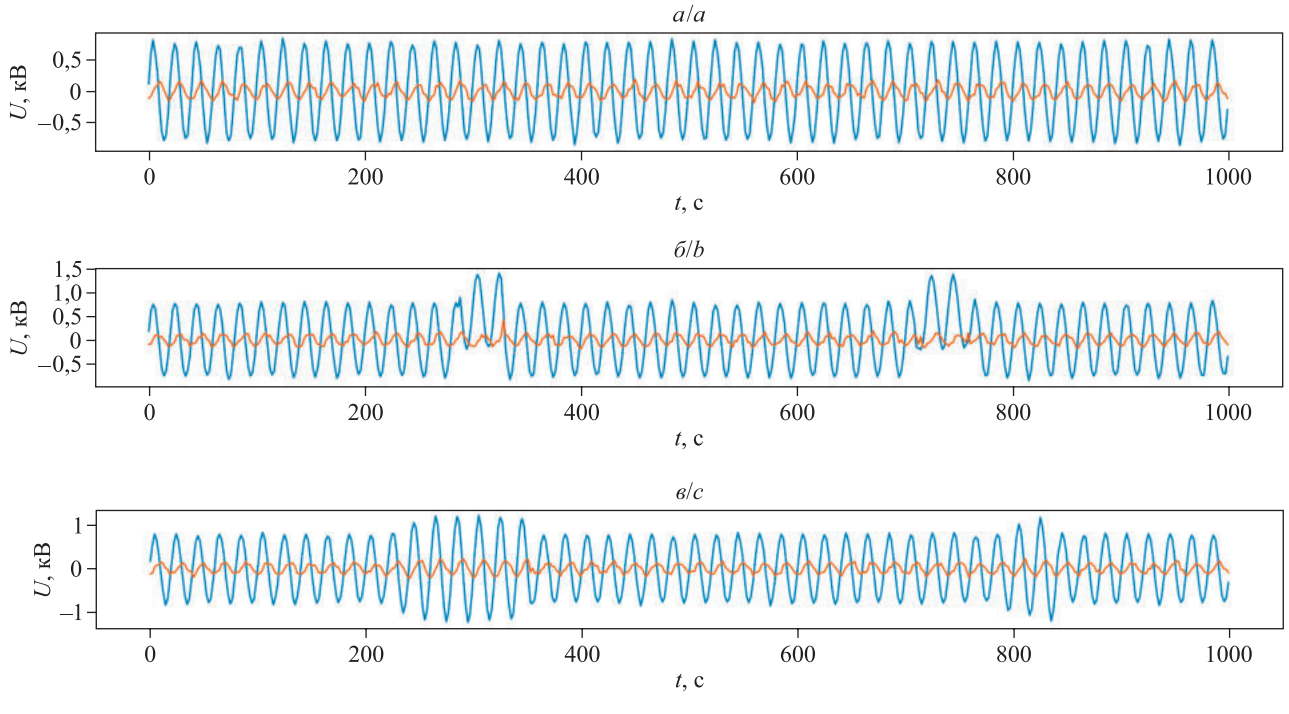

Рис. 4. Вейвлет-преобразование временного ряда напряжения, представленного на рис. 3:
а – нормальный режим; *б* – режим кибератаки; *в* – аварийный режим.
Синяя линия отражает исходный сигнал с шумами, оранжевая линия – сглаженное приближение (A_2), полученное из коэффициентов вейвлет-разложения

Fig. 4. Wavelet transform of the voltage time series shown in fig. 3:
a – normal mode; *b* – cyber attack mode; *c* – emergency mode.
The blue line represents the original signal with noise, the orange line represents the smoothed approximation (A_2) obtained from the wavelet decomposition coefficients

Этап 3: обучение алгоритмов. На этом этапе апробации методологии были проведены настройка и обучение алгоритмов обнаружения аномалий на объединенном датасете, включающем данные нормального режима и режима реализации угроз, смоделированных в ЦД. Алгоритмы обучались выявлять устойчивые поведенческие паттерны, характерные для системы при штатном функционировании, а также распознавать отклонения, возникающие в момент реализации угроз.

Аварийные режимы, отражающие непреднамеренные сбои и технические неисправности, на этапе обучения не использовались, они были зарезервированы для этапа тестирования, что позволило сформулировать более строгую задачу для алгоритмов: научиться отличать реализации угроз от других видов отклонений, не связанных с враждебным воздействием.

Для апробации методологии были использованы как классические, так и более устойчивые к выбросам алгоритмы обнаружения аномалий и кластеризации. Их выбор объясняется следующими причинами:

- алгоритмы обнаружения аномалий *Isolation forest*, *One-class SVM*, *Local outlier factor (LOF)* хорошо зарекомендовали себя при работе с аномалиями в многомерных временных рядах, они способны фиксировать точечные или локализованные выбросы;
- алгоритмы кластеризации *Density-based spatial clustering of applications with noise (DBSCAN)*, *Ordering points to identify the clustering structure (OPTICS)*, *Spectral clustering* устойчивы к шуму и способны формировать сложные по форме кластеры, их применение оправдано тем, что аномалии в сложных системах могут проявляться не как отдельные выбросы, а как сдвиги между кластерами устойчивых режимов.

В качестве дополнительных методов использовались методы *K-means* и *Gaussian mixture models (GMM)*.

Этап 4: генерация новых реализаций угроз и аварийных состояний. Для проведения тестирования в ЦД были сгенерированы новые данные, включающие:

- новые реализации угроз, которые отличались от сценариев, использованных при обучении алгоритмов (устойчивые искажения формы сигнала, а также долговременные отклонения параметров энергосистемы, моделирующие новые возможные способы реализации потенциальных угроз);
- аварийные состояния, вызванные непреднамеренными факторами, такими как внутренние сбои, технические неисправности и прочие незлонамеренные аномалии, не связанные с вредоносной активностью.

Сформированная тестовая выборка позволила проверить способность обученных алгоритмов не только выявлять аномалии – индикаторы ранее неизвестных угроз, но и эффективно отличать их от незлонамеренных отклонений, связанных с аварийными состояниями.

Этап 5: тестирование. Обученные на данных этапа 1 алгоритмы были применены к тестовой выборке, сформированной на этапе 4. Основными задачами этого этапа являлись:

- обнаружение отклонений от нормального поведения в режиме, близком к реальному времени;
- корректная классификация выявленных аномалий с разделением их на две категории – индикаторы реализации угроз (кибератаки) и неугрожающие аномалии (аварии и технические сбои).

Результаты тестирования оценивались с помощью метрик *F1-score*, *Precision* и *False positive rate (FPR)*.

На рис. 5 и 6 представлены результаты сравнительного анализа работы шести алгоритмов обработки временных рядов, примененных в рамках эксперимента по обнаружению аномалий в умной энергосети. Были рассмотрены три алгоритма обнаружения аномалий (*Isolation forest*, *LOF* и *One-class SVM*) и три алгоритма кластеризации (*DBSCAN*, *OPTICS* и *Spectral clustering*), которые обрабатывали данные после дискретного вейвлет-преобразования сигнала напряжения. На рис. 5 приведены результаты для тестового сценария, имитирующего режим кибератаки, который характеризуется точечными аномалиями, резко отличающимися от нормального поведения системы. На рис. 6 показаны результаты для сценария, соответствующего аварийному режиму, при котором наблюдаются устойчивые отклонения, вызванные техническими сбоями или авариями в сети.

Для каждого алгоритма представлены два графика. Верхний график – визуализация точек в двумерном признаковом пространстве, сформированном на основе коэффициентов вейвлет-преобразования сигнала напряжения. По оси x отложено значение коэффициента A_2 вейвлет-преобразования сигнала напряжения в нормальном режиме, а по оси y – значение того же коэффициента, но в аномальном режиме (кибератака или авария). Таким образом, каждая точка соответствует одному временному окну и отражает изменение поведения сигнала по сравнению с нормой. Точки, расположенные вдоль диагонали, соответствуют участкам сигнала без аномалий, тогда как отклонения от диагонали указывают на моменты, когда основные характеристики сигнала изменились, – это и есть потенциальные аномалии. Цвет и маркеры отражают результат классификации: нормальные точки (фоновая цветовая заливка) и аномалии (красные точки или контуры). Нижний график – тепловая карта уровня аномальности во времени. По горизонтали отложено время, уровень аномальности визуализирован цветом – от синего (низкий уровень отклонения от нормы) до красного (высокий уровень отклонения от нормы).

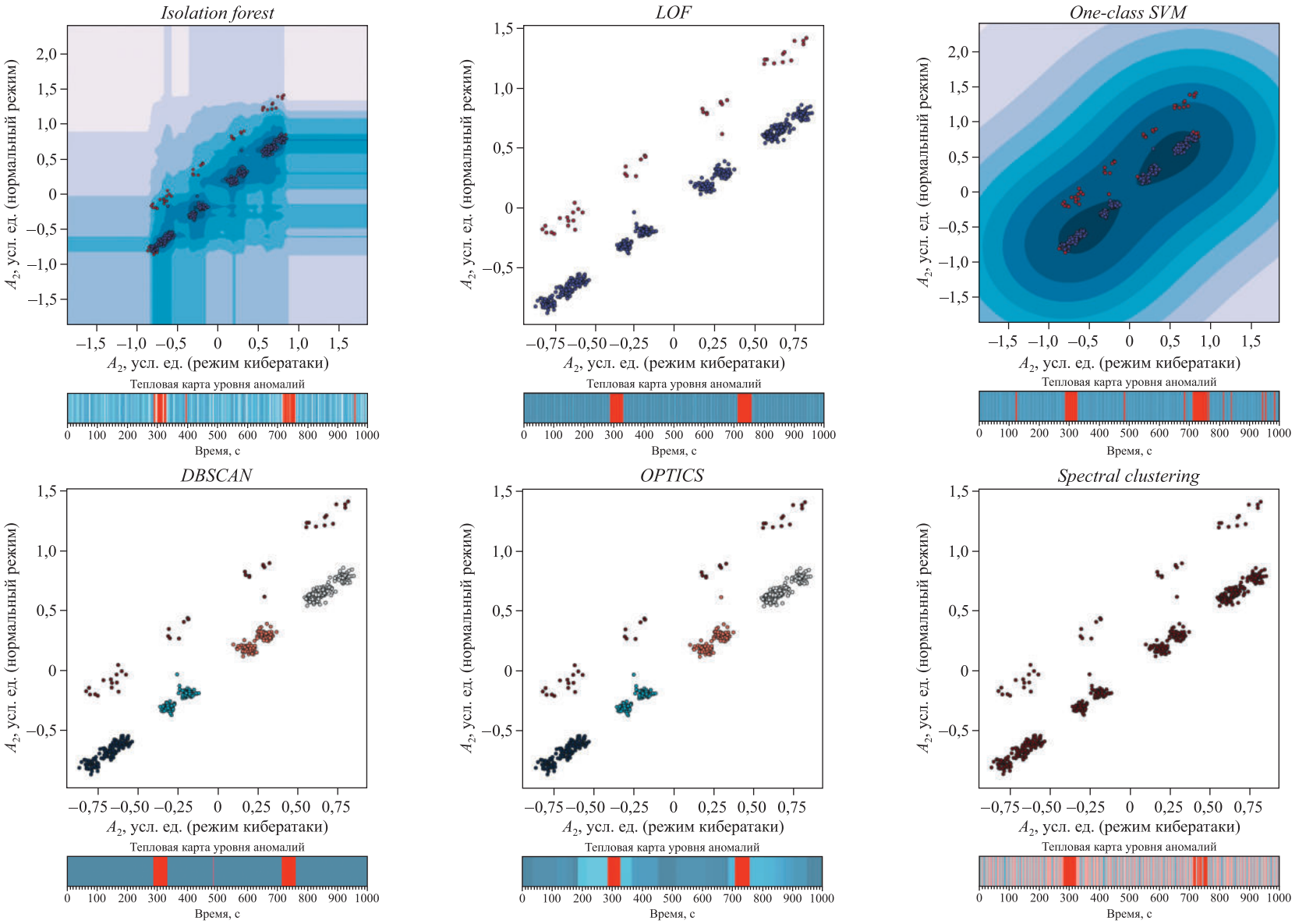

Рис. 5. Графики выделения аномальных данных в режиме кибератаки. Для каждого алгоритма представлены два графика. Верхний график – визуализация точек в двумерном признаковом пространстве, сформированном на основе коэффициентов A_2 вейвлет-преобразования сигнала напряжения. Нижний график – тепловая карта уровня аномальности во времени (синий цвет соответствует низкому уровню аномальности, красный цвет – высокому уровню аномальности)

Fig. 5. Graphs for identifying abnormal data in cyberattack mode. Two graphs are presented for each algorithm. The top graph is a visualisation of points in a two-dimensional feature space formed based on the A_2 coefficients of the wavelet transform of the voltage signal. The bottom graph is a heat map of the anomaly level over time (blue colour corresponds to low anomaly levels, red colour corresponds to high anomaly levels)

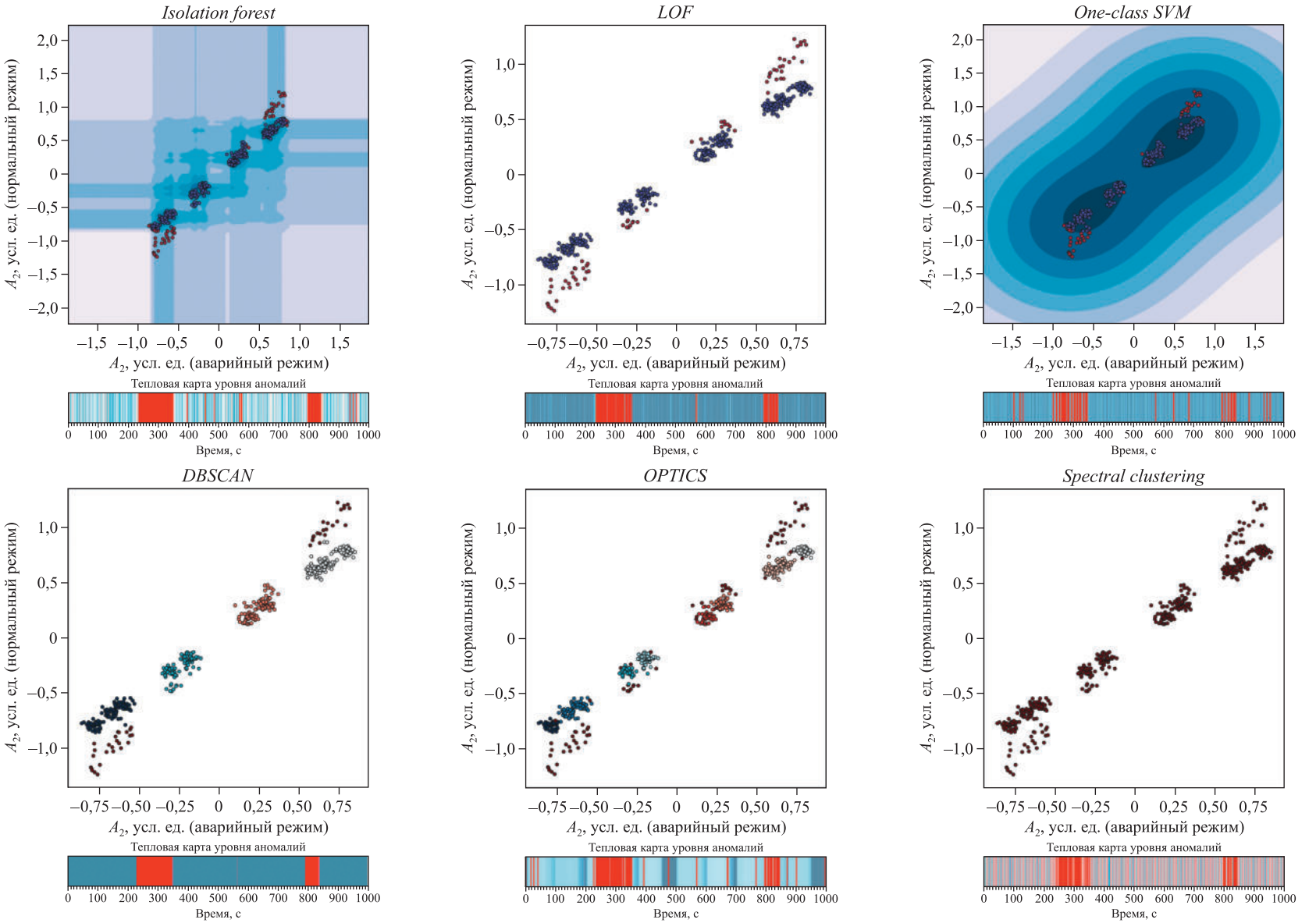

Рис. 6. Графики выделения аномальных данных в аварийном режиме. Для каждого алгоритма представлены два графика. Верхний график – визуализация точек в двумерном признаковом пространстве, сформированном на основе коэффициентов A_2 вейвлет-преобразования сигнала напряжения. Нижний график – тепловая карта уровня аномальности во времени (синий цвет соответствует низкому уровню аномальности, красный цвет – высокому уровню аномальности)

Fig. 6. Graphs for identifying abnormal data in emergency mode. Two graphs are presented for each algorithm. The top graph is a visualisation of points in a two-dimensional feature space formed based on the A_2 coefficients of the wavelet transform of the voltage signal. The bottom graph is a heat map of the anomaly level over time (blue colour corresponds to low anomaly levels, red colour corresponds to high anomaly levels)

Такой подход позволяет не только оценить поведение алгоритмов в признаковом пространстве, но и сопоставить обнаруженные аномалии с их временным расположением, что критически важно для диагностики и классификации режимов работы энергосистемы. Аналогичная обработка и анализ выполнялись для всего тестового датасета.

Оценка эффективности протестированных алгоритмов была проведена на тестовом датасете. В табл. 1 и 2 приведены усредненные метрики качества работы алгоритмов, рассчитанные по всем обработанным данным в рамках эксперимента.

Таблица 1

Метрики оценки эффективности алгоритмов (режим кибератаки)

Table 1

Algorithm evaluation metrics (cyberattack mode)

Алгоритм	*Precision*, %	*F1-score*	*FPR*, %	Время отклика, мс
Isolation forest	98	0,96	2	20
LOF	89	0,82	15	150
One-class SVM	92	0,88	5	80
DBSCAN	98	0,95	3	130
OPTICS	96	0,92	8	160
Spectral clustering	92	0,89	12	180
K-means	85	0,78	10	50
GMM	84	0,82	17	100

Таблица 2

Метрики оценки эффективности алгоритмов (аварийный режим)

Table 2

Algorithm evaluation metrics (emergency mode)

Алгоритм	*Precision*, %	*F1-score*	*FPR*, %	Время отклика, мс
Isolation forest	93	0,88	5	20
LOF	82	0,80	20	150
One-class SVM	83	0,82	10	80
DBSCAN	95	0,86	8	130
OPTICS	90	0,83	11	160
Spectral clustering	86	0,80	18	180
K-means	93	0,88	5	20
GMM	82	0,80	20	150

Проведенная оценка эффективности алгоритмов по всему тестовому датасету выявила значительные различия в способах обнаружения аномалий для двух принципиально различных сценариев – режима кибератаки и аварийного режима. Главное отличие между ними заключается в динамике проявления аномалий: кибератаки сопровождаются резкими кратковременными скачками сигнала, тогда как аварийные состояния характеризуются плавным, но устойчивым изменением параметров системы. Это отличие отражается в показателях эффективности примененных алгоритмов.

В сценарии кибератаки (см. табл. 1) алгоритм *Isolation forest* демонстрирует максимальную точность (98 %) и минимальный уровень ложных срабатываний (2 %), что объясняется его высокой чувствительностью к точечным выбросам. В то же время при анализе аварийных режимов (см. табл. 2) наблюдается общее снижение эффективности большинства алгоритмов, особенно по метрикам *F1-score* и *FPR*, что связано с постепенным и менее выраженным характером изменений сигнала. Алгоритмы кластеризации при этом показывают более стабильные результаты в обоих сценариях, поскольку они ориентированы на выявление пространственно-временных кластеров аномалий, а не на детектирование отдельных выбросов. Следует отметить, что время отклика алгоритмов остается практически неизменным независимо от типа аномалий.

По результатам апробации методологии можно сделать следующие основные выводы:

• получены подтверждающие результаты, свидетельствующие о возможности применения синтетических данных, сгенерированных в ЦД, для обучения алгоритмов обнаружения аномалий, что отражается в стабильных значениях метрики *F1-score* при выявлении ранее неизвестных угроз;

• установлена способность предложенной системы эффективно различать аномалии, связанные с реализацией угроз ИБ, и незлонамеренные отклонения, вызванные аварийными состояниями и техническими неисправностями, что обеспечивается относительно низким уровнем ложных срабатываний;

• зафиксирована тенденция к снижению времени обнаружения аномалий по сравнению с традиционными сигнатурными методами, что потенциально способствует повышению оперативности реагирования в условиях реального времени;

• установлено, что ЦД показал себя как перспективная платформа для моделирования разнообразных сценариев атак и генерации репрезентативных обучающих и тестовых данных, обеспечивая безопасную среду для апробации алгоритмов;

• получены результаты, указывающие на целесообразность применения комбинированного подхода, включающего методы детектирования точечных выбросов и кластерного анализа, для повышения надежности выявления аномалий различной природы в умных энергосетях.

Вместе с тем можно отметить следующие допущения и ограничения методологии и ее апробации:

• результаты основаны на синтетических данных, сгенерированных в ЦД, что накладывает ограничения на прямую экстраполяцию выводов на реальные производственные условия из-за возможных отличий в характере и разнообразии реальных аномалий;

• в ходе обучения алгоритмов не были использованы аварийные режимы, что создает определенное ограничение на обобщающую способность моделей при классификации неугрожающих отклонений;

• тестирование проведено в контролируемой среде с ограниченным набором сценариев, что не исключает необходимости дополнительной валидации в условиях реального времени и на реальных данных;

• используемые методы кластеризации и обнаружения выбросов предполагают определенные статистические свойства данных и могут требовать адаптации под специфику конкретных систем.

Заключение

Таким образом, в статье изложена методология анализа угроз ИБ на основе ЦД, включающая формализацию КФС и пространства угроз ИБ, безопасное моделирование атак, генерацию синтетических данных и обучение систем обнаружения аномалий – индикаторов реализации угроз. Экспериментальная проверка методологии, проведенная на модели умной энергосети, показала, что алгоритмы, обученные исключительно на данных, сгенерированных в ЦД, демонстрируют высокую точность (значение метрики *F1-score* достигает 0,96) в выявлении редких и ранее неизвестных угроз.

Ключевым преимуществом методологии является ее проактивный характер, позволяющий готовиться к угрозам до их фактического проявления в реальных системах за счет моделирования разнообразных сценариев в ЦД. Это обеспечивает безопасность (обучение без риска для инфраструктуры), эффективность (сокращение времени реакции, повышение полноты и точности детектирования) и универсальность (применимость к объектам критической инфраструктуры, интернету вещей, облачным средам).

Перспективы дальнейших исследований включают интеграцию с SIEM/SOAR-платформами для автоматизации реагирования и использование генеративного искусственного интеллекта для создания более репрезентативных сценариев атак в ЦД. Таким образом, предложенная методология открывает путь к созданию адаптивных систем безопасности, способных противостоять эволюционирующим угрозам в сложных КФС.

Библиографические ссылки

1. Sharma A, Kosasih E, Zhang J, Brintrup A, Calinescu A. Digital twins: state of the art theory and practice, challenges, and open research questions. *Journal of Industrial Information Integration*. 2022;30:100383. DOI: 10.1016/j.jii.2022.100383.

2. Mezzour G, Benhadou S, Benhadou M, Haddout A. Unleashing the potential of digital twins: a new era with aeronautics 4.0. *F1000Research*. 2024;13:193. DOI: 10.12688/f1000research.144038.1.

3. Котенко ИВ, Саенко ИБ, Скоробогатов СЮ, Лаута ОС, Кочин ВП. Методика оценки устойчивости программно-конфигурируемых сетей в условиях компьютерных атак. *Журнал Белорусского государственного университета. Математика. Информатика*. 2024;3:90–102. EDN: WFDUZG.

4. Siddique S, Haque MA, Rifat RH, George R, Shujaee K, Gupta KD. Cyber security issues in the industrial applications of digital twins. In: Institute of Electrical and Electronics Engineers. *2023 IEEE symposium series on computational intelligence (SSCI); 2023 December 5–8; Mexico City, Mexico.* [S. l.]: Institute of Electrical and Electronics Engineers; 2024. p. 873–878. DOI: 10.1109/SSCI52147.2023.10371850.

5. Abdullahi SM, Zare A, Lazarova-Molnar S. Cybersecurity in distributed industrial digital twins: threats, defenses, and key takeaways. In: Degeler V, Dustegor D, Groefsema H, Lazovik E, editors. *DiDiT-2024. 1st International workshop on distributed digital twins; 2024 June 17; Groningen, the Netherlands*. [S. l.]: CEUR-WS; 2024. Paper 2 (CEUR workshop proceedings; volume 3755). DOI: 10.5445/IR/1000174715.

6. Kotenko I. Active vulnerability assessment of computer networks by simulation of complex remote attacks. In: Institute of Electrical and Electronics Engineers. *2003 International conference on computer networks and mobile computing (ICCNMC-2003); 2003 October 20–23; Shanghai, China*. [S. l.]: Institute of Electrical and Electronics Engineers; 2003. p. 40–47. DOI: 10.1109/ICCNMC.2003.1243025.

7. Homaei M, Mogollón-Gutiérrez O, Sancho JC, Ávila M, Caro A. A review of digital twins and their application in cybersecurity based on artificial intelligence. *Artificial Intelligence Review*. 2024;57(8):201. DOI: 10.1007/s10462-024-10805-3.

8. Jones D, Snider C, Nassehi A, Yon J, Hicks B. Characterising the digital twin: a systematic literature review. *CIRP Journal of Manufacturing Science and Technology*. 2020;29(part A):36–52. DOI: 10.1016/j.cirpj.2020.02.002.

9. Stavropoulos P, Mourtzis D. Digital twins in industry 4.0. In: Mourtzis D, editor. *Design and operation of production networks for mass personalization in the era of cloud technology*. [S. l.]: Elsevier; 2022. p. 277–316. DOI: 10.1016/B978-0-12-823657-4.00010-5.

10. Novikova E, Kotenko I. Analytical visualization techniques for security information and event management. In: Kilpatrick P, Milligan P, Stotzka R, editors. *Proceedings of the 2013 21st Euromicro International conference on parallel, distributed, and network-based processing (PDP-2013); 2013 February 27 – March 1; Belfast, United Kingdom*. [S. l.]: Institute of Electrical and Electronics Engineers; 2013. p. 519–525. DOI: 10.1109/PDP.2013.84.

11. Suhail S, Iqbal M, Hussain R, Jurdak R. ENIGMA: an explainable digital twin security solution for cyber-physical systems. *Computers in Industry*. 2023;151:103961. DOI: 10.1016/j.compind.2023.103961.

12. Mustofa R, Rafiquzzaman M, Ibne Hossain NU. Analyzing the impact of cyber-attacks on the performance of digital twin-based industrial organizations. *Journal of Industrial Information Integration*. 2024;41:100633. DOI: 10.1016/j.jii.2024.100633.

13. Lucchese M, Salerno G, Pugliese A. A digital twin-based approach for detecting cyber-physical attacks in ICS using knowledge discovery. *Applied Sciences*. 2024;14(19):8665. DOI: 10.3390/app14198665.

14. Nguyen TN. Toward human digital twins for cybersecurity simulations on the metaverse: ontological and network science approach. *JMIRx Med*. 2022;3(2):e33502. DOI: 10.2196/33502.

15. Lopes PV, Silveira L, Guimaraes Aquino RD, Ribeiro CH, Skoogh A, Verri FAN. Synthetic data generation for digital twins: enabling production systems analysis in the absence of data. *International Journal of Computer Integrated Manufacturing*. 2024; 37(10–11):1252–1269. DOI: 10.1080/0951192X.2024.2322981.

16. Pärn E, Ghadiminia N, García de Soto B, Oti-Sarpong K. A perfect storm: digital twins, cybersecurity, and general contracting firms. *Developments in the Built Environment*. 2024;18:100466. DOI: 10.1016/j.dibe.2024.100466.

17. Alshammari K, Beach T, Rezgui Y. Cybersecurity for digital twins in the built environment: current research and future directions. *Journal of Information Technology in Construction (ITcon)*. 2021;26:159–173. DOI: 10.36680/j.itcon.2021.010.

18. Rahman MH, Hamedani EY, Son Y-J, Shafae M. Taxonomy-driven graph-theoretic framework for manufacturing cybersecurity risk modeling and assessment. *Journal of Computing and Information Science in Engineering*. 2024;24(7):071003. DOI: 10.1115/1.4063729.

19. Qureshi AR, Asensio A, Imran M, Garcia J, Masip-Bruin X. A survey on security enhancing digital twins: models, applications and tools. *Computer Communications*. 2025;238:108158. DOI: 10.1016/j.comcom.2025.108158.

20. Acharya S, Khan AA, Päivärinta T. Interoperability levels and challenges of digital twins in cyber-physical systems. *Journal of Industrial Information Integration*. 2024;42:100714. DOI: 10.1016/j.jii.2024.100714.

21. David I, Shao G, Gomes C, Tilbury D, Zarkout B. Interoperability of digital twins: challenges, success factors, and future research directions. In: Margaria T, Steffen B, editors. *Leveraging applications of formal methods, verification and validation. Application areas. Proceedings of the 12th International symposium, ISoLA-2024; 2024 October 27–31; Crete, Greece. Part 5*. Cham: Springer; 2025. p. 27–46 (Lecture notes in computer science; volume 15223). DOI: 10.1007/978-3-031-75390-9_3.

22. Voas J, Mell P, Laplante P, Piroumian V. *Security and trust considerations for digital twin technology*. Gaithersburg: National Institute of Standards and Technology; 2025 February. Report No.: NIST IR 8356. DOI: 10.6028/NIST.IR.8356.

23. Кочергин СВ, Артемова СВ, Бакаев АА, Митяков ЕС, Вегера ЖГ, Максимова ЕА. Повышение безопасности смарт-сетей: спектральный и фрактальный анализ как инструменты выявления кибератак. *Russian Technological Journal*. 2025;13(1):7–15. DOI: 10.32362/2500-316X-2025-13-1-7-15.

Получена 07.07.2025 / исправлена 21.08.2025 / принята 11.11.2025.
Received 07.07.2025 / revised 21.08.2025 / accepted 11.11.2025.