УДК 004.056 **DOI** 10.15622/ia.25.3.8

И.В. Котенко, И.Б. Саенко, Е.С. Митяков

ИСПОЛЬЗОВАНИЕ ЦИФРОВОГО ДВОЙНИКА ДЛЯ ОБНАРУЖЕНИЯ КИБЕРУГРОЗ В СИСТЕМАХ УПРАВЛЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫМИ ЭНЕРГОСЕТЯМИ

Котенко И.В., Саенко И.Б., Митяков Е.С. **Использование цифрового двойника для обнаружения киберугроз в системах управления интеллектуальными энергосетями.**

Аннотация. Цифровая трансформация энергетического сектора, сопровождающаяся повсеместным внедрением интеллектуальных сетей, расширяет поверхность кибератак и повышает уязвимость критически важной инфраструктуры. Традиционные системы безопасности, основанные на сигнатурах, статических правилах, а также на текущих реализациях методов машинного обучения, демонстрируют ограниченную эффективность против новых типов угроз, а их верификация на реальных объектах сопряжена с неприемлемыми рисками. В статье представлен и экспериментально апробирован интегрированный адаптивный фреймворк на основе цифрового двойника (ЦД), направленный на повышение результативности обнаружения киберугроз в системах управления интеллектуальными энергосетями. Фреймворк включает формализованную математическую постановку задачи обнаружения с учетом трех ключевых показателей: полноты (Recall), доли ложных срабатываний (FPR) и совокупного времени реакции (TTR), а также целевую функцию для их совместной оптимизации. Ключевым элементом фреймворка является контур адаптации с обратной связью по метрикам качества, управляемый трехуровневым механизмом (оперативный, тактический, стратегический режимы): результаты работы системы обнаружения угроз на реальном объекте используются для автоматического запуска процедур дообучения, оптимизации или генерации новых сценариев в безопасной виртуальной среде ЦД. Цифровой двойник формализован как кортеж взаимосвязанных моделей физических процессов, коммуникационной инфраструктуры, логики управления и базы знаний об угрозах, что позволяет учитывать многоуровневую архитектуру интеллектуальных энергосетей и специфические протоколы (МЭК 61850, Modbus). Экспериментальная валидация на синтетическом наборе данных, имитирующем работу АСУ ТП, подтвердила применимость концепции: использование синтетических данных, генерируемых в среде ЦД, позволило снизить частоту ложных срабатываний на 6,4% по сравнению с обучением на статичных данных, а активация адаптационного механизма способствовала сокращению совокупного модельного времени реакции на инциденты на 3,82%. Расчет комплексной целевой функции показал снижение обобщенного показателя качества на 16,8% после адаптации. Предложенный подход вносит вклад в решение проблем дефицита представительных данных и ограничений безопасного тестирования и может рассматриваться как шаг к реализации парадигмы «безопасность через проектирование» в энергетике. Полученные результаты представляют собой доказательство реализуемости концепции и обосновывают необходимость дальнейших исследований для промышленного внедрения, включая проверку на реальных промышленных сценариях.

Ключевые слова: интеллектуальные энергосети, цифровой двойник, адаптивное обнаружение угроз, киберустойчивость, системы управления.

1. Введение. Энергетический сектор, являясь фундаментом национальной безопасности и экономического благополучия,

переживает беспрецедентную цифровую трансформацию. На смену традиционным энергосистемам приходят интеллектуальные сети – сложные киберфизические системы (КФС), интегрирующие физические процессы генерации, передачи и распределения электроэнергии с цифровыми технологиями управления, мониторинга и связи [1]. Эта интеграция открывает огромные возможности для повышения эффективности, надежности и устойчивости энергоснабжения. Однако она же кардинально расширяет векторы киберугроз, превращая энергетическую инфраструктуру в привлекательную и критически важную мишень для злоумышленников [2, 3].

Объекты критической информационной инфраструктуры (КИИ), к которым относятся интеллектуальные энергосети, сталкиваются с эволюционирующими, целенаправленными и высоко скрытными атаками [4, 5]. Традиционные системы информационной безопасности, основанные на сигнатурном анализе и статических правилах, демонстрируют ограниченную эффективность против таких угроз, особенно в части своевременного выявления ранних признаков компрометации [6, 7]. Эти индикаторы часто статистически неотличимы от редких, но легитимных событий или технических сбоев, что приводит к высокому уровню ложных срабатываний или, что еще хуже, к пропуску реальных атак [8].

Системы управления интеллектуальными энергосетями представляют собой многоуровневые киберфизические комплексы, включающие полевой уровень (датчики, исполнительные механизмы, PMU), уровень контроллеров (ПЛК, RTU) и диспетчерский уровень (SCADA, EMS). На каждом уровне используются специфические протоколы (Modbus, IEC 61850, DNP3) и существуют характерные векторы кибератак (инъекции ложных данных, подмена управляющих команд на ПЛК, компрометация систем сбора и отображения данных и т.д.). Указанные особенности предъявляют повышенные требования к средствам обнаружения – необходимо учитывать временные ограничения реального времени и каскадный характер распространения атак и аномальных возмущений (сбоев, отказов и др.) [9].

Разработка и валидация современных методов обнаружения угроз ИБ для энергосистем сталкивается с рядом фундаментальных проблем [10, 11]:

1. Дефицит репрезентативных данных. Получение реальных, размеченных данных о кибератаках на действующих объектах КИИ крайне затруднено из-за конфиденциальности, этических и юридических ограничений, а также самой природы атак, которые стремятся остаться незамеченными.

2. Ограничения безопасного тестирования. Проведение экспериментов по моделированию атак на реальных энергетических объектах связано с неприемлемо высокими операционными и экономическими рисками, включая потенциальные сбои в энергоснабжении.

3. Динамичность угроз. Тактики, техники и процедуры (TTP) злоумышленников постоянно эволюционируют, требуя от систем защиты не статичности, а способности к непрерывной адаптации и самообучению.

Для операторов КИИ эффективность систем обнаружения угроз ИБ традиционно оценивается по трем ключевым показателям [12]:

1. Точность прогнозирования, т.е. способность системы корректно выявлять все реальные угрозы, минимизируя количество пропущенных атак (ложноотрицательных срабатываний). В терминах метрик классификации данному критерию соответствует полнота (Recall).

2. Снижение частоты ложных срабатываний, поскольку высокий уровень ложных тревог приводит к «усталости от оповещений» у персонала центров мониторинга безопасности, отвлекая их от реальных инцидентов и увеличивая операционные издержки. В количественном выражении это показатель доли ложных срабатываний (FPR).

3. Сокращение времени реагирования. Чем быстрее система позволяет обнаружить и классифицировать угрозу, тем быстрее могут быть инициированы контрмеры, что напрямую снижает потенциальный ущерб. В работе используется совокупное время реагирования (TTR), зависящее от количества истинных и ложных срабатываний.

Указанные показатели образуют компромиссную систему: стремление к максимальному Recall может увеличить FPR, а снижение FPR иногда достигается за счет потери Recall. TTR зависит от обоих. Поэтому в экспериментальной части работы оценивается изменение всех трех показателей совместно, что позволяет судить о комплексной эффективности адаптации.

В этом контексте концепция цифрового двойника (ЦД) – виртуальной, динамически обновляемой реплики физического объекта или системы – приобретает особую значимость как потенциальное решение данных проблем [13, 14]. Исследования последних лет подтверждают, что интеграция ЦД с технологиями искусственного интеллекта и машинного обучения позволяет создавать безопасные среды для генерации синтетических данных, моделирования сложных сценариев кибератак, обучения и тестирования моделей обнаружения аномалий без риска для реальной инфраструктуры [15 – 17]. Такой

подход способствует реализации парадигмы «безопасность через проектирование» (security-by-design) и обеспечивает возможность выявления даже ранее неизвестных (zero-day) угроз [18].

Несмотря на значительный прогресс, существующие решения на основе ЦД для энергетики часто фокусируются либо на моделировании физических процессов и атак [19, 20], либо на обнаружении аномалий [21, 22], но редко предлагают замкнутый, саморегулирующийся цикл, где результаты работы системы на реальном объекте автоматически используются для ее совершенствования в виртуальной среде. Особенно остро стоит проблема создания адаптивных механизмов, способных не только реагировать на новые угрозы, но и проактивно обновлять защитные модели на основе анализа их эффективности в реальном времени.

Цель данной статьи – представить и экспериментально валидировать интегрированный адаптивный фреймворк на основе ЦД для обнаружения киберугроз в системах управления интеллектуальными энергосетями, целенаправленно оптимизирующий три ключевых показателя результативности: полноту обнаружения (Recall), частоту ложных срабатываний (FPR) и совокупное время реагирования на инциденты (TTR). Этот фреймворк синергетически объединяет методологию адаптивного обнаружения угроз ИБ с архитектурой прототипа ЦД для АСУ энергосистемы.

Основной научный вклад работы заключается в следующем:

1. Предложение интегрированного фреймворка, объединяющего архитектуру ЦД для энергосистемы и метод адаптивного обнаружения угроз в единую систему, ориентированную на повышение Recall, снижение FPR и сокращение совокупного времени реагирования на инциденты.
2. Формализация и реализация трехуровневого адаптационного механизма (оперативный, тактический, стратегический режимы) как ключевого элемента фреймворка, обеспечивающего непрерывное поддержание актуальности и эффективности моделей защиты по всем трем критериям.
3. Экспериментальная валидация фреймворка, включающая:

– демонстрацию эффективности адаптационного механизма на модели АСУ ТП энергосистемы.

– подтверждение пользы использования синтетических данных, генерируемых в среде ЦД, для повышения качества обнаружения угроз ИБ, что косвенно способствует сокращению времени реагирования за счет повышения достоверности срабатываний.

– обсуждение практической применимости и ограничений предложенного подхода для промышленного внедрения.

2. Анализ релевантных работ. Научное сообщество активно исследует потенциал ЦД как инструмента для повышения киберустойчивости критически важных инфраструктур, включая энергетические системы. Обзор релевантных работ структурирован вокруг трех взаимосвязанных направлений: применение ЦД в энергетике, их использование для задач кибербезопасности и современные методы адаптивного обнаружения аномалий.

ЦД становятся ключевым инструментом цифровой трансформации энергетического сектора, обеспечивая виртуальное отражение физических активов и процессов в реальном времени. Эта технология позволяет не просто наблюдать за системой, но и прогнозировать ее поведение, оптимизировать работу и интегрировать возобновляемые источники энергии, что напрямую способствует повышению эффективности и устойчивости всей энергетической цепочки [23, 24].

Применение ЦД охватывает все этапы жизненного цикла энергии – от генерации и хранения до передачи, распределения и конечного потребления [25]. В ветро- и солнечной энергетике они повышают производительность и надежность установок [26, 27], в системах хранения – улучшают управление батареями [28], а в традиционных секторах, таких как нефтегаз, – обеспечивают контроль целостности активов и управление жизненным циклом [29]. Основные функции включают мониторинг в реальном времени, предиктивное обслуживание, оптимизацию сетей и прогнозирование отказов [30, 31].

Экономический и экологический эффект от внедрения ЦД подтверждается многочисленными исследованиями: снижение операционных затрат, экономия энергии до 30% и ускорение перехода к «чистой» энергетике за счет более точного и основанного на данных управления [32, 33]. ЦД становятся катализатором устойчивого развития, помогая минимизировать экологический след и двигаться к целям углеродной нейтральности [34].

Однако путь к массовому внедрению не лишен препятствий. Высокая стоимость реализации, сложности интеграции с унаследованными системами, отсутствие единых стандартов и, что особенно актуально для нашей работы, недостаточная проработка вопросов кибербезопасности остаются ключевыми барьерами [35, 36]. Будущее за созданием масштабируемых и совместимых архитектур,

а также за глубокой интеграцией с технологиями ИИ, 5G и IoT для построения автономных и самовосстанавливающихся энергосистем [37, 38].

ЦД все чаще рассматриваются не просто как инструмент оптимизации, но и как мощный актив в арсенале киберзащиты. Виртуальная реплика реальной системы превращается в безопасную «песочницу», где можно моделировать самые изощренные атаки, не рискуя нарушить работу КИИ [39, 40]. Это особенно ценно для энергетики, где любой сбой может иметь каскадные последствия. Ключевая ценность ЦД для ИБ проявляется в трех основных направлениях:

1. Моделирование и тестирование. ЦД позволяют создавать высокоточные модели киберфизических систем, на которых можно безопасно отрабатывать сценарии атак, тестировать защитные меры и оценивать уязвимости. Это позволяет выявлять слабые места до того, как они будут найдены злоумышленниками [41, 42].

2. Проактивная защита. Благодаря зеркальному отображению реальных процессов, системы на основе ЦД способны выявлять аномалии на ранних стадиях, что позволяет операторам оперативно реагировать и минимизировать потенциальный ущерб [43, 44].

3. Повышение эффективности обнаружения. ЦД могут использоваться как платформы для обмана, эффективно отвлекая злоумышленников и собирая ценные данные об их тактиках. В сочетании с методами объяснимого ИИ (XAI) это позволяет не просто обнаруживать угрозы, но и понимать их природу [45, 46].

Вместе с тем сама технология ЦД создает новые векторы атак. Интеграция с IoT и облачными сервисами расширяет поверхность атак, вызывая обеспокоенность по поводу целостности данных и возможности манипуляции системой [47, 48]. Отсутствие единых стандартов безопасности усложняет развертывание защищенных решений, особенно в сложных, гетерогенных средах [49]. Критически важным также является обеспечение точной и безопасной синхронизации между цифровой моделью и ее физическим прототипом, поскольку любые рассогласования могут быть использованы злоумышленниками [40, 45].

Перспективным направлением считается интеграция ЦД с передовыми технологиями. Например, сочетание с ИИ позволяет автоматизировать обнаружение угроз, а использование блокчейна – обеспечить прозрачность и контроль доступа [13, 50]. Не менее важен и принцип «безопасность через проектирование» (security-by-design), который предполагает внедрение механизмов защиты, таких как

модели нулевого доверия и непрерывный мониторинг, уже на этапе проектирования архитектуры ЦД [39, 47].

Таким образом, хотя ЦД открывают новые горизонты для кибербезопасности, их эффективное и безопасное применение требует комплексного подхода, учитывающего как возможности, так и присущие им риски.

Быстрая цифровизация промышленных систем управления и интеллектуальных сетей неизбежно влечет за собой рост уязвимостей. В этих условиях статические системы защиты теряют актуальность, уступая место адаптивным методам обнаружения аномалий, способным учиться на новых данных и подстраиваться под меняющиеся условия работы и тактики злоумышленников. Такие методы становятся критически важными для обеспечения безопасности и надежности энергетической инфраструктуры.

Современный арсенал адаптивных подходов разнообразен и активно развивается. Рассмотрим некоторые из них:

1. Глубокое обучение показывает высокую точность в выявлении сложных, многоэтапных атак. Архитектуры, такие как адаптивные остаточные рекуррентные сети (ARRNN-DGRU) и сверточные LSTM (ConvLSTM), превосходят традиционные модели по скорости и надежности детектирования, особенно в условиях зашумленных данных [51 – 53].

2. Методы распределенного обучения и обучения с сохранением конфиденциальности решают проблему защиты данных в децентрализованных системах. Эти подходы позволяют обучать модели непосредственно на локальных устройствах – например, на «умных» счетчиках или контроллерах без необходимости передачи чувствительной информации на центральный сервер, сохраняя при этом высокую эффективность обнаружения [54 – 56]. Такие решения особенно актуальны для ресурсоограниченных сред и способствуют повышению доверия со стороны пользователей и операторов системы.

3. Модели, устойчивые к изменению в данных и интерпретируемые модели решают проблему ложных срабатываний, вызванных естественными изменениями в работе системы (например, сменой привычек потребления). Используя такие фреймворки, как LSTM или системы на основе правил (AI-BRB), эти методы не только адаптируются к новым условиям, но и объясняют причины своих решений, что критически важно для операторов [57, 58].

4. Статистические и физико-математические методы обеспечивают масштабируемое и быстрое обнаружение в реальном времени. Подходы, основанные на расширенных фильтрах Калмана

или гибридных AI-физических моделях, эффективно различают технические сбои, внешние возмущения и целенаправленные кибератаки, опираясь на фундаментальные законы физики энергосистем [59 – 61].

Таким образом, адаптивные методы обнаружения аномалий формируют важную парадигму для современных энергосистем. Их способность к самообучению и адаптации позволяет операторам не просто реагировать на угрозы, а опережать их. Однако для реализации этого потенциала требуется безопасная и управляемая среда для непрерывного обучения и верификации моделей – именно ту нишу и занимает технология ЦД.

Проведенный анализ подтверждает, что области применения ЦД в энергетике, использования ЦД для кибербезопасности и разработки адаптивных методов обнаружения аномалий развиваются стремительно, однако остаются в значительной степени изолированными. Существующие решения либо используют ЦД как пассивную среду для симуляции, либо применяют адаптивные алгоритмы без механизма непрерывного обновления на основе обратной связи от реальной системы. Ключевая научная новизна и практическая ценность нашей работы заключаются в создании интегрированного, саморегулирующегося фреймворка, в котором ЦД энергосистемы выступает не просто виртуальной копией, а динамической обучающей платформой. Предложенный замкнутый адаптивный цикл, управляемый трехуровневым механизмом (оперативный, тактический, стратегический), обеспечивает непрерывную верификацию и актуализацию моделей обнаружения угроз на основе реальных данных. Именно эта архитектура позволяет напрямую и системно оптимизировать три критических показателя эффективности: точность прогнозирования угроз, снижение частоты ложных срабатываний и сокращение времени реагирования на инциденты, что является конечной целью для операторов критически важной энергетической инфраструктуры.

3. Предлагаемый фреймворк. Фреймворк представляет собой интегрированную систему для повышения результативности обнаружения киберугроз в интеллектуальных энергосетях, основанную на концепции цифрового двойника и механизме адаптивного обнаружения угроз информационной безопасности. Отличительной особенностью фреймворка выступает формализация и автоматизация цикла адаптации, в котором обратная связь от реальной системы служит триггером для запуска конкретных

процедур в виртуальной среде: дообучения модели, оптимизации ее параметров или генерации новых сценариев.

Математическая постановка задачи. Пусть $X = \{x_t\}$ – временной ряд наблюдаемых параметров (например, напряжение, ток). Каждому наблюдению соответствует неизвестное состояние $y_t \in \{\text{норма}, \text{сбой}, \text{атака}\}$. Требуется построить классификатор $f(X, \Theta)$, где Θ – параметры модели, который по наблюдениям определяет состояние.

Качество классификации оценивается метриками *Recall* (полнота), *Precision* (точность), *FPR* (доля ложных срабатываний). Совокупное время реагирования *TTR* зависит от количества истинных и ложных срабатываний. Адаптация заключается в обновлении Θ на основе обратной связи от реальной системы. Цель адаптации – минимизация взвешенной суммы потерь:

$$L = \lambda_1 \cdot (1 - Recall) + \lambda_2 \cdot FPR + \lambda_3 \cdot \frac{TTR}{T_{\max}} \rightarrow \min, \quad (1)$$

где $\lambda_1, \lambda_2, \lambda_3$ – весовые коэффициенты, отражающие приоритеты при защите (например, $\lambda_1 > \lambda_2$ при критичности пропуска атаки), а $T_{\max}$ – максимально допустимое время реагирования.

Для решения сформулированной задачи разработан на основе ЦД, архитектура которого представлена далее.

Архитектура фреймворка построена на взаимодействии двух функциональных контуров: ЦД и реального объекта (АСУ энергосетью), связанных через центральный адаптационный механизм (Рис. 1) [62].

Логика работы фреймворка определяется следующим потоком данных и управления:

1. Инициализация в контуре ЦД. Процесс начинается с генерации синтетических наборов данных, отражающих нормальные, аварийные и атакующие сценарии. Эти данные используются для обучения и верификации моделей обнаружения аномалий.

2. Развертывание на реальном объекте. Верифицированные модели передаются и интегрируются в системы мониторинга и управления на всех уровнях АСУ энергосети.

3. Мониторинг и обнаружение. Модели анализируют потоки операционных данных в реальном времени, выявляя отклонения и классифицируя события.

4. Формирование обратной связи. Агрегированные данные о работе моделей, включая метрики качества (F1-score, *Recall*, *FPR*) и информацию о срабатываниях, передаются обратно в контур ЦД.

5. Адаптация. Адаптационный механизм анализирует поступившую обратную связь. При отклонении метрик за предустановленные пороги инициируется один из трех режимов адаптации (оперативный, тактический, стратегический), что приводит к обновлению моделей или сценариев в ЦД. После этого цикл повторяется.

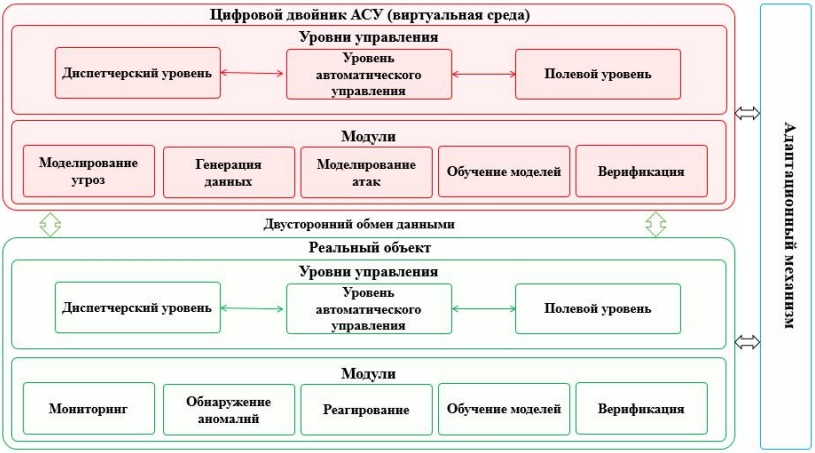

Рис. 1. Общая архитектура адаптивного фреймворка на основе ЦД

Такая архитектура создает предпосылки для поддержания актуальности механизмов защиты без необходимости прямого вмешательства в работу критически важной инфраструктуры.

Архитектура ЦД декомпозирована на восемь специализированных функциональных модулей, каждый из которых играет определенную роль в реализации замкнутого адаптивного цикла (таблица 1). В совокупности эти модули образуют единую платформу, способную не только имитировать работу энергосистемы, но и обеспечивать непрерывное обучение и адаптацию моделей обнаружения угроз, что способствует поддержанию их актуальности в условиях изменяющейся среды.

Для формализации предложенного подхода определим ЦД как кортеж взаимосвязанных моделей:

$$DT = \langle M_{phys}, M_{comm}, M_{cyber}, M_{threat}, Sync \rangle,$$

где M_{phys} – модель физических процессов энергосистемы, параметризуемая вектором настраиваемых параметров p_{phys} (в общем случае может быть задана системой дифференциальных уравнений, описывающих переходные процессы в линиях, генераторах, нагрузках); M_{comm} – модель коммуникационной инфраструктуры, включающая протоколы, топологию сети, задержки и нарушения связи; M_{cyber} – модель логики управления (алгоритмы ПЛК, SCADA, противоаварийной автоматики), представленная конечным автоматом или набором правил; M_{threat} – база знаний об уязвимостях, угрозах и сценариях атак (TTP), структурированная, например, в формате MITRE ATT&CK для промышленных систем; $Sync$ – механизм синхронизации, обеспечивающий согласование состояния ЦД с реальным объектом путем обновления параметров моделей по данным телеметрии.

В рамках экспериментального прототипа, описанного в разделе 4, реализована упрощенная версия ЦД, достаточная для доказательства реализуемости концепции. Модель M_{phys} представлена параметрической генерацией временного ряда напряжения; модели M_{comm} и на данном этапе не задействованы, а M_{threat} применяется на уровне генерации сценариев атак (импульсные искажения, фазовые сдвиги, низкочастотная модуляция). Предложенная архитектура ЦД (таблица 1) является расширяемой: компоненты могут заменяться на более сложные альтернативы (например, M_{phys} на нейросетевую модель или полную систему дифференциальных уравнений, M_{comm} на эмулятор сети) без изменения общей структуры фреймворка. Параметры p_{phys} калибруются по историческим данным реального объекта, их обновление осуществляется при поступлении новой информации в процессе адаптации.

Представленная архитектура ЦД реализована программно в соответствии с модульной структурой таблицы 1. Конкретные алгоритмы генерации синтетических данных (включая параметры сигналов нормального, аварийного и атакующего режимов), методы предобработки (вейвлет-преобразование с базисом Добеши db1 и нормализация Min-Max), а также процедуры обучения и адаптации модели обнаружения на основе Isolation Forest описаны далее. Таким образом, ЦД раскрыт не на уровне общих требований, а как конкретная реализация, параметры которой обеспечивают воспроизводимость экспериментов и позволяют рассматривать его как рабочий прототип.

Таблица 1. Роль модулей ЦД в реализации адаптивного фреймворка

№	Модуль	Описание
1	Модуль киберфизического моделирования	Разворачиваются все процессы моделирования, симуляции и адаптации. Его точность напрямую влияет на релевантность генерируемых сценариев и, следовательно, на качество обучения моделей.
2	Модуль моделирования угроз ИБ	Формирует и управляет базой знаний об уязвимостях, TTP-сценариях (например, на основе MITRE ATT&CK) и последствиях атак. В стратегическом режиме адаптации именно этот модуль обновляется на основе анализа новых угроз, выявленных на реальном объекте.
3	Модуль генерации синтетических данных	Генерирует обучающие и тестовые выборки, включая сложные, ранее не встречавшиеся паттерны, идентифицированные как недостаточно покрытые текущими моделями. Это ключевой компонент для оперативного и стратегического режимов адаптации.
4	Модуль симуляции атак	Внедряет сгенерированные атаки в киберфизическую модель для тестирования их реализуемости и оценки воздействия. Позволяет безопасно "проиграть" новые сценарии до их потенциального появления в реальной системе.
5	Модуль обнаружения аномалий	Выполняет детектирование угроз как в контуре ЦД (на этапе тестирования), так и на реальном объекте. Его выходные данные (метрики Precision, Recall, FPR) являются основными триггерами для запуска адаптационного механизма.
6	Модуль агрегации данных	Обеспечивает единый доступ ко всем данным – синтетическим, историческим и потоковым. Критически важен для процессов дообучения (оперативный режим) и переобучения (стратегический режим).
7	Модуль оценки угроз	Интерпретирует результаты обнаружения, оценивая угрозы по критериям вероятности, критичности и потенциального ущерба. Его выводы могут использоваться для приоритезации сценариев, которые должны быть сгенерированы в стратегическом режиме адаптации.
8	Модуль верификации	Оценивает корректность работы всей системы. Его отчеты о снижении эффективности моделей или недостаточной реалистичности сценариев являются формальным основанием для запуска адаптации, особенно в тактическом и стратегическом режимах.

Метод адаптивного обнаружения. Фреймворк реализует метод адаптивного обнаружения, который формализует процесс обновления моделей на основе анализа их эффективности в реальных условиях эксплуатации. Методика разделяет процессы, происходящие в контуре ЦД и на реальном объекте, и определяет условия для их взаимодействия через адаптационный механизм [63].

Этапы в контуре ЦД:

1. Моделирование сценариев и генерация данных. Создаются синтетические наборы, включающие нормальные, аварийные и атакующие режимы.

2. Предобработка данных. Применяются методы снижения размерности (PCA) и шумоподавления (вейвлет-преобразование, например, с базисом Добеши) для повышения качества входных данных.

3. Обучение моделей. Используются алгоритмы машинного обучения, такие как Isolation Forest, для построения моделей классификации событий на «норму», «сбой» или «угрозу».

4. Тестирование моделей. Оцениваются метрики качества (F1-score, *Recall, Precision, FPR*) на независимых тестовых выборках.

5. Внедрение моделей. Верифицированные модели передаются в контур реального объекта.

Этапы в контуре реального объекта:

1. Развертывание моделей. Алгоритмы интегрируются в программно-аппаратные средства контроля на объекте.

2. Мониторинг операционных данных. Осуществляется непрерывный сбор телеметрии, логов событий и параметров технологических процессов.

3. Обнаружение и классификация в реальном времени. Модели анализируют поступающие данные, выявляя аномалии и классифицируя события.

4. Реагирование. При обнаружении угрозы инициируются заранее определенные меры (оповещение, изоляция сегмента сети и т.д.).

5. Обратная связь и анализ. Агрегированные данные о работе моделей передаются обратно в контур ЦД для анализа.

Адаптация инициируется автоматически, когда качество работы модели на реальном объекте опускается ниже заданного порога. Формально, если модель вычисляет набор показателей качества $Q = \{q_1, q_2, \dots, q_n\}$, то адаптация запускается, если выполняется условие:

$$\exists q_i \in Q: q_i \notin [q_{imin}, q_{imax}]. \quad (2)$$

В зависимости от причины и характера отклонений выделяются три режима адаптации [63]:

1. Оперативный режим. В рамках данного режима осуществляется дообучение моделей на новых данных без изменения их архитектуры и параметров. Формально, при поступлении нового набора данных $\mathrm{D_{new}}$, обновляются параметры модели $\theta \to \theta'$ методом дообучения:

$$\theta' = \arg\min_{\theta} L(\theta;\ D_{train} \cup D_{new}), \quad (3)$$

где L – функция потерь, D_{train} – обучающая выборка, D_{new} – новые данные, θ – обучаемые параметры модели. Например, если модель начала ошибочно срабатывать на новую легитимную активность, ее дообучают на примерах такой активности, чтобы скорректировать ее реакцию.

2. Тактический режим. Включает пересмотр архитектуры модели, параметров и порогов классификации, что формализуется как оптимизация модели в более широком пространстве гиперпараметров:

$$h' = \arg\min_{h}\left[\min_{\theta} L(\theta,\ h;\ D_{train} \cup\ D_{new})\right],$$
$$\theta' = \arg\min_{\theta}(\theta,\ h';\ D_{train} \cup\ D_{new}), \quad (4)$$

где h – гиперпараметры модели (архитектура, пороги и др.). Например, при росте ложных срабатываний на определенный тип аварийных состояний производится подбор новых порогов классификации или увеличение сложности модели (например, глубины или числа деревьев в ансамбле).

3. Стратегический режим. В данном режиме происходит возврат к этапу моделирования сценариев и генерации обучающих данных с последующим повторным обучением моделей:

$$D_{synth} = G(S_{new}), \theta'' = \arg\min_{\theta} L\left(\theta; D_{train} \cup D_{synth}\right), \quad (5)$$

где S_{new} – новые сценарии, G – генератор данных (в ЦД объекта КИИ), D_{synth} – данные, сгенерированные по сценариям угроз. Режим обеспечивает обновление обучающих выборок с учетом новых или измененных сценариев угроз. Например, при появлении информации о новой тактике целевой атаки в ЦД моделируются соответствующие сценарии, и модели обучаются на расширенном наборе данных.

Стратегический режим адаптации предназначен для реагирования на принципиально новые типы угроз, которые не могут быть покрыты простым дообучением или оптимизацией гиперпараметров. Его запуск может быть инициирован автоматически при обнаружении аномалий, не соответствующих известным сценариям, либо по команде оператора после появления информации о новых тактиках атак из внешних источников. В первом случае модуль моделирования угроз (таблица 1) генерирует гипотетические сценарии, соответствующие выявленным отклонениям, и передает их в модуль генерации синтетических данных. Во втором случае новые сценарии разрабатываются экспертом и загружаются в ЦД.

В рамках текущей реализации фреймворка предусмотрена возможность как автоматической, так и полуавтоматической генерации, однако для обеспечения безопасности целесообразно включать этап верификации оператором перед обновлением (повторным обучением) моделей.

Процесс адаптации показан на рисунке 2.

Рис 2. Замкнутый цикл адаптации

Для автоматического выбора режима адаптации введем систему решающих правил, основанную на анализе отклонений метрик качества и характеристик новых данных.

Пусть в момент времени t после обработки очередного блока данных на реальном объекте получен вектор метрик $Q_t = (Recall_t, FPR_t, F1_t)$. Заданы допустимые интервалы: $Recall \in [R_{min}, 1], FPR \in [0, F_{max}]$, а также порог δ_{rec} минимального улучшения *Recall* при оперативной адаптации.

Оперативный режим (O) активируется при выполнении следующих условий: ($Recall_t < R_{min}$) и имеются новые достоверные данные D_{new}. Выполняется дообучение без изменения гиперпараметров.

Тактический режим (T) активируется, если после k последовательных оперативных адаптаций ($k \geq 2$) улучшение *Recall* не превысило δ_{rec}: ($Recall_t - Recall_{t-k} < \delta_{rec}$) и ($FPR_t > F_{max}$). Производится оптимизация гиперпараметров модели (например, поиск по сетке) на накопленных данных.

Стратегический режим (C) активируется при выполнении одного из следующих условий:

– обнаружение аномалии, не соответствующей ни одному из известных классов в M_threat (фиксируется модулем оценки угроз на основе кластеризации или экспертных правил);

– поступление из внешних источников информации о новой тактике атак (TTP), отсутствующей в текущей базе знаний.

Формально:

Значения R_{min}, F_{max}, δ_{rec} и k задаются на этапе настройки фреймворка исходя из требований к безопасности объекта и могут корректироваться в процессе эксплуатации.

4. Экспериментальная валидация и результаты. Целью экспериментальной валидации выступает количественная оценка эффективности предложенного адаптивного фреймворка на основе ЦД для повышения киберустойчивости интеллектуальных энергосистем. Оценка проводится по трем критериям:

– точность прогнозирования угроз ИБ;
– частота ложноположительных срабатываний;
– сокращение времени реагирования на инциденты ИБ.

Для обеспечения научной строгости и воспроизводимости результатов ниже представлено описание **набора данных**, на котором проводилась **проверка**, а также методика проведения серии из трех взаимосвязанных экспериментов.

Для апробации фреймворка был сгенерирован синтетический набор данных, имитирующий работу АСУ интеллектуальной энергосети. Набор данных моделирует поведение одного параметра (напряжения в линии электропередачи) в трех различных режимах функционирования системы:

1. Режим «Норма» (Normal). Основной сигнал – синусоида с частотой 50 Гц и амплитудой 220 В. Добавлена суточная модуляция амплитуды ±10% и гауссов шум со стандартным отклонением σ = 5 В.
2. Режим «Технический сбой» (Failure): Искаженная синусоида с плавным изменением частоты (45-55 Гц) и амплитуды (±25%).
3. Режим «Признаки киберугрозы» (Threat). Синусоида 50 Гц с наложенными атакующими воздействиями: импульсные искажения (падение амплитуды до -40%), фазовые сдвиги (0,1π-0,4π) и низкочастотная модуляция (1-5 Гц, амплитуда 10-20% от основной). Генерация синтетических данных выполняется модулем 3 ЦД. В стратегическом режиме адаптации генератор может формировать новые комбинации атакующих воздействий на основе анализа недостаточно покрытых областей пространства признаков.

Структура данных следующая:

– общее количество записей – 150000 временных отсчетов;
– частота дискретизации – 100 Гц (100 измерений напряжения в секунду);

– продолжительность моделирования – 25 минут реального времени;

– целевая переменная (метка класса). Каждая запись размечена одним из трех классов: Normal (Норма): 100000 записей (66,7%); Failure (Технический сбой): 25000 записей (16,7%); Threat (Признаки киберугрозы): 25000 записей (16,7%).

Разбиение набора данных было выполнено с учетом задачи обучения модели на "нормальном" поведении и последующей адаптации. Обучающая выборка была сформирована из первых 70000 записей класса Normal (70% от общего объема данных этого класса). Оставшиеся 30000 записей класса Normal, а также все записи классов Failure (25000) и Threat (25 000) были объединены в пул для формирования валидационной и тестовой выборок.

Валидационная выборка была сформирована путем случайной выборки 7500 записей из каждого из трех классов (всего 22500 записей, 15% от общего объема набора данных). Аналогично, тестовая выборка была сформирована из следующих 7500 записей каждого класса (всего 22500 записей, 15% от общего объема набора данных). Такое разбиение обеспечило сбалансированность валидационной и тестовой выборок по классам, что критически важно для объективной оценки способности модели различать нормальное поведение, технические сбои и признаки кибератак.

В целях повышения эффективности последующего машинного обучения, к сгенерированным временным рядам применяется двухэтапная предобработка данных: вейвлет-преобразование и нормализация. Вейвлет-преобразование применялось для подавления высокочастотного шума. В рамках эксперимента использовался базис Добеши первого порядка (db1) с уровнем декомпозиции 4. Коэффициенты детализации двух самых высокочастотных уровней были обнулены, что позволило снизить шумовую составляющую сигнала при сохранении информативных признаков аномалий. Нормализация (в частности, Min-Max нормализация) применялась для масштабирования всех признаков к единому диапазону [0; 1]. Нормализация обеспечивает равноправное участие всех признаков в формировании прогноза модели. Последовательное применение вейвлет-преобразования и нормализации позволяет подготовить данные к этапу обучения модели, повысив их информативность и обеспечив корректную работу алгоритмов машинного обучения.

Для обеспечения полной воспроизводимости экспериментов синтетический набор данных был сгенерирован с помощью детерминированного алгоритма, инициализированного фиксированным

случайным сидом (seed = 42). Алгоритм последовательно формирует временной ряд напряжения, имитирующий три режима функционирования энергосистемы – нормальный (Normal), аварийный (Failure) и признаки киберугрозы (Threat). Процедура генерации включает следующие шесть шагов:

1. Создание базового сигнала – синусоиды частотой 50 Гц и амплитудой 220 В.
2. Наложение суточной модуляции амплитуды в пределах ±10%.
3. Добавление гауссова шума со стандартным отклонением $\sigma = 5$ В.
4. Последовательная генерация 150000 временных отсчетов с шагом 10 мс (частота дискретизации 100 Гц).
5. Внедрение в случайные моменты времени заранее определенных шаблонов (паттернов):

– для режима «Технический сбой» – плавное изменение частоты (45-55 Гц) и амплитуды (±25%);

– для режима «Признаки киберугрозы» – импульсные искажения (падение амплитуды до -40%), фазовые сдвиги ($0{,}1\pi$-$0{,}4\pi$) и низкочастотная модуляция (1-5 Гц, амплитуда 10-20% от основной).

6. Присвоение каждому сгенерированному отсчету метки класса в зависимости от активного режима: «Normal», «Failure» или «Threat».

Эксперименты проводились на программно-аппаратной платформе, имитирующей архитектуру АСУ ТП интеллектуальной энергосети. Платформа реализована в виде изолированной виртуальной среды и включает эмуляторы промышленных контроллеров (ПЛК) и систем SCADA, что позволяет воспроизводить реалистичные сценарии функционирования и кибератак в контролируемых условиях, исключая риски для реальной инфраструктуры.

Все вычисления выполнялись на рабочей станции с процессором Intel Core i7, 32 ГБ ОЗУ, под управлением ОС Ubuntu 22.04 LTS. Использовались библиотеки Python: scikit-learn 1.4.2, PyWavelets 1.4.1, pandas 2.2.1, numpy 1.26.4. Для обеспечения воспроизводимости все случайные генераторы были зафиксированы с помощью np.random.seed(42).

Эксперимент 1. Валидация адаптационного механизма – повышение точности прогнозирования. Целью первого эксперимента являлась количественная оценка эффективности предложенного адаптационного механизма в повышении точности прогнозирования угроз информационной безопасности, с акцентом на минимизацию пропущенных атак через увеличение метрики Recall. В качестве тестовой среды использовалась синтетическая модель АСУ ТП

энергосистемы, построенная на основе описанного ранее набора данных. В качестве базового алгоритма обнаружения был выбран метод Isolation Forest [64]. Данный выбор обусловлен задачей доказательства концепции: этот алгоритм устойчив к несбалансированным выборкам, обладает высокой интерпретируемостью и позволяет изолированно оценить вклад адаптационного механизма без усложнения, связанного с глубокими нейросетями. Предложенный фреймворк является алгоритмически независимым и может быть расширен на другие модели (LSTM, графовые нейросети, гибридные физико-информационные модели) [65] без изменения логики адаптации.

Начальные параметры модели были установлены следующим образом: количество деревьев – 100, параметр ‘contamination’ («загрязнение») – 0,1, что соответствует ожидаемой доле аномалий в 10%. Обучение проводилось исключительно на данных класса «Норма» (70000 записей), что имитирует реальную ситуацию дефицита данных об атаках на этапе первоначального развертывания системы. Тестирование осуществлялось на сбалансированной выборке из 22500 записей, включающей по 7500 записей каждого из трех классов событий.

На первом этапе модель была обучена и протестирована в исходной конфигурации. Результаты показали, что метрика *Recall* составила 0,8253, что указывает на высокую базовую способность модели выявлять аномалии, при этом сохраняется потенциал для улучшения. Для повышения эффективности был активирован адаптационный механизм. Были последовательно применены два режима: оперативный (инкрементное дообучение на дополнительных 7500 записях класса Normal из валидационной выборки) и тактический (оптимизация гиперпараметров с помощью поиска по сетке, включавшего варианты n_estimators = [100; 150; 200] и contamination = [0,05; 0,1; 0,15]). Лучшей комбинацией, максимизирующей Recall на валидационной выборке, оказалась n_estimators =100, contamination = 0,15. После проведения этих процедур финальная адаптированная модель была протестирована на том же тестовом наборе.

Результаты, представленные в таблице 2, демонстрируют, что активация адаптационного механизма позволила стабилизировать и незначительно улучшить производительность модели. Метрика *Recall*, отражающая способность системы выявлять истинные угрозы, увеличилась на 0,7%, в то время как FPR снизилась на 0,7 процентных пункта. Это свидетельствует о том, что даже в условиях упрощенной модели, адаптация на основе обратной связи позволяет системе тонко

настраивать свои параметры для повышения общей сбалансированности (F1-score), что является важным шагом в направлении минимизации совокупного ущерба.

Таблица 2. Сравнение метрик модели до и после адаптации

Показатель	До адаптации	Послед адаптации	Изменение
Recall	0,8253	0,8325	+0,0072
F1-score	0,7583	0,7605	+0,0021
Precision	0,7014	0,6999	-0,0015
FPR (%)	17,47%	16,75%	-0,72%

Следует подчеркнуть, что данные результаты получены в контролируемой виртуальной среде на синтетических данных, моделирующих лишь один параметр энергосистемы. Они не могут быть напрямую экстраполированы на работу системы в реальных промышленных условиях, где данные многомерны, а атаки являются многокомпонентными, скрытными и эволюционирующими. Тем не менее, эксперимент подтверждает принципиальную работоспособность предложенного адаптационного механизма: система успешно прошла полный цикл – от получения обратной связи до корректировки модели – и продемонстрировала измеримое улучшение ключевых показателей. Это является важным доказательством реализуемости концепции для дальнейшего масштабирования фреймворка.

Незначительный прирост метрик, зафиксированный в ходе эксперимента, объясняется рядом факторов, заложенных в методику исследования. Основной целью первого эксперимента было не достижение максимальных абсолютных значений метрик, а демонстрация работоспособности замкнутого цикла адаптации. Успешное прохождение данных через все этапы фреймворка (от генерации в ЦД до адаптации модели на основе обратной связи) и фиксация даже незначительного, но статистически устойчивого улучшения является важным результатом, подтверждающим жизнеспособность предложенной архитектуры.

Эксперимент 2. Валидация пользы синтетических данных – снижение частоты ложных срабатываний. Целью второго эксперимента является количественная оценка вклада синтетических данных, генерируемых в среде ЦД, в снижение FPR системы обнаружения угроз информационной безопасности. В отличие от первого эксперимента, где адаптация модели происходила за счет дообучения на новых «нормальных» данных и оптимизации гиперпараметров, здесь фокус смещен на оценку качества

и релевантности аномальных данных, используемых для обучения. Ключевой исследовательский вопрос: обеспечивают ли синтетические аномалии, сгенерированные по запросу адаптационного механизма ЦД, измеримое преимущество по сравнению с использованием эквивалентного объема «внешних» имитационных данных при прочих равных условиях?

Для обеспечения научной строгости и исключения влияния конфаундеров (в частности, объема обучающей выборки), эксперимент реализован в рамках методологии контролируемого сравнения при фиксированном объеме данных. Это означает, что обе сравниваемые модели обучаются на идентичном общем количестве примеров (85000 записей), но с принципиальным различием в происхождении данных для аномальных классов:

1. Контрольная конфигурация (Baseline – «Внешние имитационные данные»); модель обучается на комбинированном наборе, включающем:

– 70000 записей класса «Норма» (исходная обучающая выборка, идентичная используемой в Эксперименте 1);

– 15000 записей аномалий (по 7500 записей классов «Failure» и «Threat»), случайным образом выбранных из валидационной выборки. Эти данные представляют собой статичный, «внешний» пул, сформированный заранее и не адаптируемый под текущие потребности модели.

2. Экспериментальная конфигурация (Proposed – «Синтетические данные ЦД»); модель обучается на комбинированном наборе, включающем:

– 70000 записей класса «Норма» (та же исходная выборка);

– 15000 записей аномалий (по 7500 записей классов «Failure» и «Threat»), сгенерированных заново с помощью детерминированного алгоритма, имитирующего работу модуля генерации синтетических данных ЦД. Эти данные создаются в момент проведения эксперимента, что соответствует сценарию проактивной генерации данных по запросу адаптационного механизма.

Таким образом, единственной независимой переменной в эксперименте выступает источник данных для аномальных классов: статичный внешний пул vs. динамически сгенерированный в среде ЦД. Все остальные параметры (алгоритм обучения, процедура предобработки, валидационная и тестовая выборки) остаются идентичными для обеих конфигураций.

Эксперимент включает следующие этапы:

1. Формирование обучающих наборов. Для контрольной конфигурации из валидационной выборки извлекаются все доступные записи классов «Failure» и «Threat». Для экспериментальной конфигурации с помощью детерминированного алгоритма генерируются новые временные ряды для тех же классов и в том же количестве, с использованием физических и атакующих моделей, описанных в разделе 4.1. Оба обучающих набора формируются путем конкатенации 70 000 «нормальных» записей с соответствующими 15000 аномальных записей и последующего случайного перемешивания.

2. Предобработка данных. Ко всем данным (обучающим, валидационным, тестовым) применяется единый конвейер предобработки: дискретное вейвлет-преобразование с последующей нормализацией Min-Max. Нормализация выполняется с использованием скейлера, обученного исключительно на исходных 70000 записях класса «Норма».

3. Оптимизация гиперпараметров. Для каждой конфигурации независимо выполняется поиск оптимальных гиперпараметров алгоритма Isolation Forest (n_estimators $\in$ {100; 150; 200}, contamination $\in$ {0,05; 0,1; 0,15}) на валидационной выборке. Целевой функцией оптимизации является максимизация метрики Recall для аномалий (объединенные классы «Failure» и «Threat»), что соответствует задаче минимизации пропущенных угроз.

4. Финальное обучение и оценка. Модели с найденными оптимальными гиперпараметрами обучают на своих полных обучающих наборах и оценивают на единой, сбалансированной тестовой выборке (22500 записей: по 7500 на каждый из трех классов). Оценка проводится по ключевым метрикам: *Recall*, *Precision*, F1-score и *FPR*.

Результаты оценки на тестовой выборке представлены в таблице 3. Анализ данных демонстрирует статистически значимое преимущество экспериментальной конфигурации.

Основной результат эксперимента заключается в следующем. Использование синтетических данных, генерируемых в среде ЦД, позволило снизить FPR на 6,4% (или на 3,56 процентных пункта) при одновременном повышении F1-score на 4,7% и Recall на 8,0%. Это свидетельствует о том, что данные, созданные «по запросу» в рамках адаптивного контура ЦД, в данных условиях являются более информативными для задачи обучения модели, чем данные, взятые из статичного «внешнего» пула. Они позволяют модели лучше отделить истинные угрозы (класс «Threat») и технические сбои (класс «Failure») от нормального функционирования, что напрямую приводит к снижению числа ложных тревог.

Таблица 3. Сравнение эффективности моделей, обученных на «внешних» и «синтетических» аномалиях на тестовой выборке

Метрика	Контрольная конфигурация	Экспериментальная конфигурация	Абсолютное изменение	Относительное изменение
Recall	0,4469	0,4825	+0,0356	+8,0%
Precision	0,7155	0,7138	-0,0017	-0,2%
F1-score	0,5501	0,5758	+0,0257	+4,7%
FPR (%)	55,31%	51,75%	-3,56 п.п.	-6,4%

Результаты эксперимента подтверждают гипотезу о том, что ЦД выступает не просто как инструмент симуляции, а как генератор знаний. Способность ЦД проактивно создавать синтетические сценарии атак и сбоев, адаптированные под текущие потребности модели, является ключевым фактором для снижения операционной нагрузки на персонал за счет уменьшения ложных срабатываний.

Эксперимент 3. Количественная оценка сокращения совокупного времени реагирования на инциденты ИБ через оптимизацию операционной нагрузки. Третий эксперимент направлен на количественную оценку влияния адаптационного механизма на моделируемое сокращение совокупного времени реагирования (Total Time to Respond, TTR) на инциденты информационной безопасности.

В отличие от экспериментов 1 и 2, где аномалиями считались как кибератаки, так и технические сбои, в данном эксперименте мы фокусируемся исключительно на классе «Threat». Технические сбои (Failure) рассматриваются как часть нормального функционирования, поскольку они не требуют немедленного реагирования со стороны центра мониторинга безопасности. Таким образом, тестовая выборка содержит 15 000 нормальных записей (7500 класса «Normal» и 7500 класса «Failure») и 7500 записей класса «Threat». Используется та же выборка, что и в эксперименте 1, но с переопределенной целевой переменной.

Совокупное время реагирования определяется как суммарное время, затрачиваемое на обработку всех срабатываний (истинных и ложных) в течение фиксированного периода мониторинга. Оно зависит от количества истинных срабатываний (True Positives, TP) и ложных срабатываний (False Positives, FP):

$$TTR = (TP \times C_{tp}) + (FP \times C_{fp}), \quad (6)$$

где C_{tp} – среднее время (в секундах), затрачиваемое на обработку одного истинного срабатывания (например, верификация угрозы, инициация

контмер, документирование инцидента), C_{fp} – среднее время (в секундах), затрачиваемое на обработку одного ложного срабатывания (например, анализ логов, отклонение тревоги, сброс состояния).

Значения констант C_{tp} и C_{fp} основаны на эмпирических данных, приведенных в исследованиях по эффективности центров мониторинга безопасности. Согласно отраслевым оценкам, среднее время первичной верификации реального инцидента составляет от 30 до 90 минут, а обработка ложного срабатывания – от 5 до 30 минут [66 – 68]. Для целей данного эксперимента выбраны консервативные оценки нижней границы в секундах: C_{tp} = 1800 секунд (30 минут), C_{fp} = 300 секунд (5 минут). Это соответствует минимально возможному времени обработки при высокой степени автоматизации и зрелости процессов центра мониторинга и реагирования на инциденты. В реальных условиях с более длительными процедурами обработки выигрыш от снижения числа ложных срабатываний будет выше.

Процедура проведения эксперимента следующая:

1. Исходные данные. В качестве входных данных используются результаты прогнозирования двух моделей на единой тестовой выборке, полученные в ходе эксперимента 1:

– сценарий «До адаптации»: предсказания базовой модели (обучена только на данных класса «Норма»);

– сценарий «После адаптации»: предсказания финальной адаптированной модели (прошедшей оперативный и тактический режимы адаптации)».

Для расчета *TP* и *FP* метки классов переопределяются. Положительным классом считается «Threat», отрицательным – объединение классов «Normal» и «Failure».

2. Расчет метрик TP и FP. Для каждого сценария на основе вектора предсказаний и истинных меток рассчитывается матрица ошибок. Из нее извлекаются значения TP и FP:

– *TP* – количество записей, где истинный класс – «Threat» и модель предсказала «аномалия» (-1);

– *FP* – количество записей, где истинный класс – «Normal» или «Failure», но модель предсказала «аномалия» (-1).

3. Расчет TTR. По формуле (6) рассчитывается совокупное время реагирования для каждого сценария.

4. Сравнение и оценка сокращения. Рассчитывается абсолютное и относительное сокращение TTR:

$$\Delta TTR = TTR_{before} - TTR_{after}, \quad (7)$$

$$\Delta TTR\% = \left(\frac{\Delta TTR}{TTR_{before}}\right) \times 100\%. \quad (8)$$

Результаты расчета представлены в таблице 4. Расчет в таблице выполнен для бинарной классификации, где положительным классом являются только кибератаки («Threat»), а технические сбои отнесены к нормальному фону. Значения *TP* и *FP* получены непосредственно из матриц ошибок моделей до и после адаптации на тестовой выборке объемом 22500 записей (15000 нормальных и 7500 атак).

Таблица 4. Расчет совокупного времени реагирования (TTR) до и после адаптации.

Сценарий	*TP*	*FP*	*TTR*, сек.	Моделируемое сокращение *TTR*
До адаптации	2231	2621	4802100	-
После адаптации	2147	2513	4618500	-183 600 (-3,82%)

Анализ данных демонстрирует, что активация адаптационного механизма привела к сокращению совокупного времени реагирования на 3,82%. Сокращение обусловлено снижением количества ложных срабатываний на 108 единиц (с 2621 до 2513), что уменьшает нагрузку на персонал центра мониторинга безопасности. Небольшое снижение числа истинных срабатываний (на 84 события, с 2231 до 2147) связано с изменением порогов классификации в процессе тактической адаптации, однако вклад сокращения *FP* оказался доминирующим при выбранных весовых коэффициентах. Полученный результат подтверждает, что даже умеренное улучшение метрик, особенно снижение *FPR*, способно дать измеримый практический эффект в виде экономии времени аналитиков.

Для количественной оценки эффективности адаптации использовалась функция потерь (1). Приняв весовые коэффициенты $\lambda_1 = 0{,}5$ (приоритет полноты), $\lambda_2 = 0{,}3$ (чувствительность к ложным срабатываниям) и $\lambda_3 = 0{,}2$ (время реагирования), а также $T_{max} = 4802100$ с (базовое время до адаптации), получим снижение функции потерь (1) на 16,8%.

5. Обсуждение. Представленная работа предлагает интегрированный адаптивный фреймворк на основе ЦД для повышения результативности обнаружения киберугроз в интеллектуальных энергосистемах. Данный раздел посвящен обсуждению полученных результатов, их места в контексте существующих исследований, а также анализу сильных и слабых сторон предложенного подхода.

Научный вклад настоящего исследования заключается в интеграции архитектуры ЦД для АСУ энергосистемы, метода адаптивного обнаружения аномалий и механизма генерации синтетических данных в единый, управляемый цикл адаптации на основе обратной связи. В отличие от большинства существующих решений, фокусирующихся на одном аспекте (моделирование, обнаружение или генерация данных), предложенный фреймворк создает замкнутую систему, где результаты работы на виртуальном прототипе используются для непрерывного совершенствования моделей обнаружения.

Результаты серии экспериментов позволяют сделать следующие выводы:

1. Активация адаптационного механизма приводит к измеримому, хотя и незначительному, улучшению баланса между чувствительностью и специфичностью модели: *Recall* увеличился на 0,7%, *FPR* снизился на 0,7 п.п. Это подтверждает работоспособность концепции замкнутого цикла, где обратная связь от системы используется для тонкой настройки модели.

2. Использование синтетических данных, генерируемых в среде ЦД, демонстрирует статистически значимое преимущество: снижение *FPR* на 6,4% (3,56 п.п.) и рост F1-score на 4,7% по сравнению с обучением на статичном внешнем пуле данных. Это свидетельствует о том, что адаптивная генерация сценариев позволяет создавать более релевантные обучающие примеры, что напрямую снижает операционную нагрузку за счет уменьшения ложных срабатываний.

3. Моделируемое сокращение совокупного времени реагирования (*TTR*) на 3,82% при снижении числа ложных срабатываний на 108 единиц подчеркивает практическую значимость даже незначительных улучшений метрик. Это измеримый эффект, оказывающий прямое влияние на эффективность работы центров мониторинга безопасности.

В целом, результаты согласуются с гипотезой о том, что ЦД может служить не просто инструментом для симуляции, а динамической платформой для генерации данных и тестирования. Однако важно подчеркнуть: ЦД не является источником автономной функции обнаружения или принятия решений. Его роль строго ограничена предоставлением контролируемой среды для генерации сценариев и верификации моделей. Качество конечного результата полностью определяется двумя факторами:

- качеством алгоритма обнаружения аномалий;
- релевантностью данных.

Эксперименты демонстрируют, что при фиксированном алгоритме обнаружения, улучшение качества входных данных (через адаптивную генерацию в ЦД) приводит к измеримому повышению эффективности. Таким образом, ЦД выступает как инфраструктурный элемент, автоматизирующий процессы, которые вручную были бы крайне затратны или невозможны.

Результаты подтверждают возможность применения подхода, однако не доказывают его эффективность в промышленной эксплуатации.

Основные преимущества фреймворка, вытекающие из его архитектуры и подтвержденные экспериментами, заключаются в следующем:

1. Все этапы, связанные с моделированием атак и обучением моделей, происходят в изолированной виртуальной среде ЦД. Это нивелирует риск нарушения работы реальной энергосистемы.

2. Трехуровневый адаптационный механизм (оперативный, тактический, стратегический) обеспечивает непрерывную обратную связь и корректировку моделей **с минимальным ручным вмешательством** (в стратегическом режиме возможна полуавтоматическая генерация сценариев с последующей верификацией оператором), позволяя системе адаптироваться к меняющимся угрозам.

3. Генерация синтетических данных в среде ЦД позволяет обходить фундаментальную проблему отрасли – дефицит размеченных данных о реальных инцидентах.

4. Интеграция задач кибербезопасности на этапе проектирования и моделирования в ЦД позволяет выявлять и устранять уязвимости еще до развертывания системы в промышленную эксплуатацию.

Несмотря на полученные положительные результаты, необходимо четко обозначить границы их применимости. Проведенное исследование имеет ряд существенных ограничений:

1. Ограниченная размерность данных. Все эксперименты проведены на синтетическом наборе данных, имитирующем поведение единственного параметра (напряжения). Реальные киберфизические атаки на АСУ ТП являются многомерными, затрагивая множество взаимосвязанных параметров (ток, частота, температура, давление, сетевой трафик, логи событий). Архитектура фреймворка поддерживает многомерные данные за счет модульной структуры и возможности интеграции алгоритмов обнаружения, работающих с векторами признаков произвольной размерности. Проверка

эффективности в условиях многомерной корреляции составляет одно из направлений дальнейших исследований.

2. Упрощенность алгоритмической базы. В качестве детектора аномалий в рамках исследования использовался алгоритм Isolation Forest. Выбор обусловлен задачей доказательства осуществимости концепции: алгоритм обладает высокой интерпретируемостью, устойчив к несбалансированным выборкам и требует незначительных вычислительных ресурсов, что позволяет изолировать и оценить вклад именно механизмов адаптации и генерации данных, минимизируя влияние сложности модели детектирования. Следует отметить, что Isolation Forest не предназначен для решения задач обнаружения сложных, многоэтапных или малозаметных атак, требующих учета временной динамики и скрытых зависимостей в данных. Результаты, полученные с его использованием, не могут быть напрямую экстраполированы на более сложные архитектуры, такие как рекуррентные или сверточно-рекуррентные нейронные сети, а также гибридные модели, сочетающие методы машинного обучения с физическими законами функционирования энергосистем, которые в современных исследованиях демонстрируют более высокие показатели точности в аналогичных задачах.

3. Отсутствие верификации на реальных данных. Все результаты получены на синтетических данных, сгенерированных детерминированным алгоритмом. Они не подтверждены на исторических данных реальных инцидентов или на данных, полученных с физических тестовых стендов. Это является критическим ограничением, поскольку синтетические данные, даже при высокой степени детализации, не способны в полной мере воспроизвести стохастическую природу, шум и непредсказуемость реальных технологических процессов и атак.

4. Недостаточная сложность моделирования атак. Сгенерированные атаки представляют собой относительно простые паттерны (импульсные искажения, фазовые сдвиги). Они не моделируют скрытность, адаптивность или многоэтапность реальных APT, которые специально разработаны для того, чтобы оставаться незамеченными в течение длительного времени. Эффективность фреймворка против таких изощренных угроз не оценивалась.

5. Отсутствие анализа операционных издержек. В работе не проводился анализ вычислительной сложности адаптационного механизма и затрат на создание и поддержку ЦД. Внедрение ЦД

требует значительных ресурсов на этапе проектирования, калибровки и синхронизации с реальной системой. Эти затраты должны быть сопоставлены с получаемой выгодой для принятия обоснованных решений о промышленном внедрении.

Несмотря на указанные ограничения, проведенное исследование может представлять практический интерес на текущем этапе развития технологий ЦД. Предложенный подход формирует методологическую основу и предоставляет инструментарий для безопасной генерации обучающих данных и верификации алгоритмов обнаружения аномалий в контролируемой среде. Это может быть полезно на этапах проектирования, предварительной оценки устойчивости и обучения персонала, когда доступ к реальным данным ограничен или их использование сопряжено с рисками. Интеграция подобного фреймворка в процессы жизненного цикла АСУ может способствовать повышению их киберустойчивости, однако окончательная оценка его практической применимости требует дальнейшей экспериментальной проверки в условиях, приближенных к промышленной эксплуатации.

Следует отметить потенциальную уязвимость предложенного механизма к атакам типа «отравление модели» (Model Poisoning) [69]. Если злоумышленник сможет контролировать часть данных обратной связи, используемых для дообучения (например, имитировать легитимные изменения или маскировать атаки под нормальные отклонения), это может привести к постепенному смещению модели и снижению качества обнаружения. Данная проблема характерна для систем с обратной связью и требует дополнительных мер защиты, таких как проверка источников данных, использование робастных методов обучения или периодический аудит обновлений оператором. В рамках данного исследования данный аспект не рассматривался и выступает предметом дальнейших исследований.

Для объективной оценки места предложенного фреймворка в существующем ландшафте, в таблице 5 приведено его сравнение с двумя основными классами альтернативных решений. Сравнительный анализ носит качественный характер. Он основан на результатах, полученных в рамках экспериментов с алгоритмом Isolation Forest. Количественная валидация указанных характеристик – в частности, точности против сложных APT, скорости адаптации и полной стоимости владения (TCO) – в условиях реальной промышленной эксплуатации выходит за рамки данного исследования и определена как направление будущей работы.

Важно подчеркнуть, что предлагаемый фреймворк не позиционируется как замена существующим методам обнаружения

аномалий. Скорее, он предлагает альтернативный подход к организации процессов обучения и верификации моделей через использование виртуальной среды ЦД. Его основное преимущество – возможность безопасного тестирования и генерации данных без зависимости от наличия реальных инцидентов – потенциально позволяет использовать его в качестве вспомогательного инструмента при работе с различными классами алгоритмов обнаружения. Однако заявлять о его универсальной применимости ко всем существующим или будущим архитектурам преждевременно. Эффективность интеграции с конкретными моделями (включая нейросетевые или физико-ориентированные) требует отдельной экспериментальной проверки и не может быть постулирована априори.

Таблица 5. Сравнительный анализ подходов к обнаружению угроз ИБ

Критерий	Традиционные методы (сигнатуры, правила)	Современные методы глубокого обучения	Предлагаемый фреймворк
Точность (против zero-day)	Низкая	Высокая (при наличии данных)	Высокая (адаптируется к новым угрозам)
Адаптивность	Отсутствует	Низкая (требует ручного переобучения)	Высокая (автоматическая адаптация)
Безопасность тестирования	Высокая	Низкая (риск при обучении на live-данных)	Очень высокая (тестирование в ЦД)
Зависимость от реальных данных об атаках	Низкая	Очень высокая	Низкая (используются синтетические данные)
Время и стоимость внедрения	Быстрое и дешевое	Долгое и дорогое (сбор и разметка данных)	Среднее (настройка ЦД и сценариев)

6. Заключение. В статье представлен и экспериментально апробирован прототип интегрированного адаптивного фреймворка на основе ЦД для обнаружения признаков киберугроз в системах управления интеллектуальными энергосетями. Ключевым элементом фреймворка выступает замкнутый цикл адаптации, управляемый метриками качества, в котором результаты работы системы на реальном объекте используются для формализованного совершенствования моделей в виртуальной среде ЦД.

Результаты экспериментов демонстрируют принципиальную работоспособность предложенного подхода. Активация адаптационного механизма позволяет улучшить баланс между

чувствительностью и специфичностью модели (рост *Recall* на 0,7%, снижение *FPR* на 0,7 процентных пункта). Использование синтетических данных, генерируемых в среде ЦД, обеспечивает статистически значимое снижение частоты ложных срабатываний (на 6,4%) и рост F1-score (на 4,7%). Даже небольшое снижение числа ложных срабатываний (на 108 единиц) приводит к моделируемому сокращению совокупного времени реагирования (на 3,82%), что подтверждает практическую значимость оптимизации метрики FPR для повышения результативности работы систем обнаружения.

Следует подчеркнуть, что все выводы, сделанные в настоящей работе, основаны на результатах, полученных в рамках ограниченной экспериментальной модели – синтетического набора данных и виртуальной платформы, имитирующей архитектуру АСУ ТП Smart-сети. Следовательно, представленные результаты следует рассматривать исключительно как доказательство реализуемости концепции, демонстрирующее принципиальную работоспособность и потенциал предложенного фреймворка. Они не являются окончательным подтверждением его эффективности в условиях реальной промышленной эксплуатации, но обосновывают целесообразность проведения дальнейших, более масштабных исследований.

Для трансформации прототипа в промышленно применимое решение необходимо решить следующие задачи:

1. Расширить фреймворк для работы с комплексными, многомерными потоками данных, включая параметры тока, частоты, температуры, сетевого трафика и логов событий.
2. Провести апробацию на физических тестовых стендах и с использованием исторических данных реальных инцидентов для подтверждения эффективности в условиях, близких к промышленным.
3. Провести серию экспериментов с заменой Isolation Forest на другие классы моделей для определения границ применимости и вклада адаптационного механизма при различных типах детекторов.
4. Создать API для бесшовной интеграции фреймворка с существующими корпоративными платформами (SIEM, SOAR).
5. Оценка эффективности фреймворка при интеграции с современными архитектурами глубокого обучения (например, LSTM, ConvLSTM) и физико-ориентированными моделями на одинаковых наборах данных для объективного определения его вклада в повышение точности обнаружения.

Реализация указанных направлений, на наш взгляд, позволит приблизить перевод фреймворка из стадии исследовательского прототипа в стадию зрелого инженерного решения.

В заключение отметим направления дальнейших фундаментальных и прикладных исследований:

1. Исследование многомерной адаптивной детекции в киберфизических системах. Разработка и верификация моделей, способных одновременно анализировать коррелированные потоки данных (напряжение, ток, частота, сетевой трафик, логи событий).

2. Разработка теоретико-игровых и активных методов генерации сценариев в ЦД, исследование механизмов проактивного моделирования адаптивных атакующих стратегий на основе принципов состязательного машинного обучения (adversarial machine learning) и теории игр. Это позволит генерировать синтетические данные, максимально приближенные к тактикам реальных APT.

3. Формализация метрик эффективности замкнутого адаптационного цикла. Создание математической модели, описывающей взаимосвязь между качеством синтетических данных, скоростью адаптации, изменением метрик обнаружения и совокупным временем реагирования.

4. Исследование устойчивости адаптивных моделей к целенаправленным атакам на процесс обучения.

5. Разработка методов защиты цифрового двойника как критического элемента инфраструктуры кибербезопасности.

6. Экспериментальная верификация на физических тестовых стендах и реальных промышленных данных.

Эти направления формируют основу для систематического развития предложенного подхода от исследовательского прототипа к теоретически обоснованной и практически применимой методологии обеспечения киберустойчивости интеллектуальных энергосистем.

Литература

1. Gehrmann C., Gunnarsson M. A Digital Twin Based Industrial Automation and Control System Security Architecture // IEEE Transactions on Industrial Informatics. 2020. vol. 16. no. 1. pp. 669–680. DOI: 10.1109/TII.2019.2938885.
2. Venkatachary S.K., Prasad J., Alagappan A., Andrews L.J.B., Raj R.A., Duraisamy S. Cybersecurity and cyber-terrorism challenges to energy-related infrastructures – Cybersecurity frameworks and economics – Comprehensive review // International Journal of Critical Infrastructure Protection. 2024. vol. 45. 100677 p. DOI: 10.1016/j.ijcip.2024.100677.
3. Diaba S.Y., Shafie-khah M., Elmusrati M. Cyber-physical attack and the future energy systems: A review // Energy Reports. 2024. vol. 12. pp. 2914–2932. DOI: 10.1016/j.egyr.2024.08.060.
4. Kotenko I., Saenko I., Lauta O., Kribel A. An Approach to Detecting Cyber Attacks against Smart Power Grids Based on the Analysis of Network Traffic Self-Similarity // Energies. 2020. vol. 13. no. 19. 5031 p. DOI: 10.3390/en13195031.
5. Izrailov K., Buinevich M., Kotenko I. Intelligent Detection of Cyber Attacks on Electrical Power Systems Based on Simulation and Graph-Based Modeling //

Proceedings of the International Conference Intelligent Systems (INTELS'24). Communications in Computer and Information Science (CCIS). Springer. 2026. vol. 2603. pp. 231–245. DOI: 10.1007/978-3-032-04758-8_18.
6. Lin H. Computer network information security monitoring system // Proceedings of Second International Conference on Big Data, Computational Intelligence, and Applications (BDCIA 2024). 2025. vol. 13550. DOI: 10.1117/12.3059895.
7. Villegas-Ch W., Govea J., Gutierrez R., Navarro A.M., Mera-Navarrete A. Effectiveness of an Adaptive Deep Learning-Based Intrusion Detection System // IEEE Access. 2024. vol. 12. pp. 184010–184027. DOI: 10.1109/ACCESS.2024.3512363.
8. Kotenko I., Saenko I., Lauta O., Kribel A. A Proactive Protection of Smart Power Grids against Cyberattacks on Service Data Transfer Protocols by Computational Intelligence Methods // Sensors. 2022. vol. 22. no. 19. 7506 p. DOI: 10.3390/s22197506.
9. Desnitsky V.A., Kotenko I.V., Nogin S.B. Detection of Anomalies in Data for Monitoring of Security Components in the Internet of Things // Proceedings of the XVIII IEEE International Conference on Soft Computing and Measurements (SCM). 2015. pp. 189–192. DOI: 10.1109/SCM.2015.7190452.
10. Aydin Z. Detecting Cybersecurity Threats in Digital Energy Systems Using Deep learning for Imbalanced Datasets // International Journal of Energy Economics and Policy. 2025. vol. 15. no. 3. pp. 614–628. DOI: 10.32479/ijeep.19649.
11. Cremer F., Sheehan B., Fortmann M., et al. Cyber risk and cybersecurity: a systematic review of data availability // The Geneva Papers on Risk and Insurance - Issues and Practice. 2022. vol. 47. pp. 698–736. DOI: 10.1057/s41288-022-00266-6.
12. Kotenko I., Chechulin A. Computer attack modeling and security evaluation based on attack graphs // Proceedings of the IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). 2013. pp. 614–619. DOI: 10.1109/IDAACS.2013.6662998.
13. Homaei M., Mogollon-Gutierrez O., Sancho J.C., et al. A review of digital twins and their application in cybersecurity based on artificial intelligence // Artificial Intelligence Review. 2024. vol. 57. 201 p. DOI: 10.1007/s10462-024-10805-3.
14. McLaughlin K.L. The power of digital twins in the cybersecurity mesh // EDPACS. 2023. vol. 68. no. 6. pp. 35–39. DOI: 10.1080/07366981.2023.2263214.
15. Krishnaveni S., Chen T.M., Sathiyanarayanan M., et al. CyberDefender: an integrated intelligent defense framework for digital-twin-based industrial cyber-physical systems // Cluster Computing. 2024. vol. 27. pp. 7273–7306. DOI: 10.1007/s10586-024-04320-x.
16. Salim M.M., Camacho D., Park J.H. Digital Twin and federated learning enabled cyberthreat detection system for IoT networks // Future Generation Computer Systems. 2024. vol. 161. pp. 701–713. DOI: 10.1016/j.future.2024.07.017.
17. Sousa B., Arieiro M., Pereira V., Correia J., Lourenco N., Cruz T. ELEGANT: Security of Critical Infrastructures With Digital Twins // IEEE Access. 2021. vol. 9. pp. 107574–107588. DOI: 10.1109/ACCESS.2021.3100708.
18. De Hoz D.J., Temperekidis A., Katsaros P., Konstantinou C. An IoT Digital Twin for Cyber-Security Defence Based on Runtime Verification // Lecture Notes in Computer Science. Springer. 2022. vol. 13701. pp. 556–574. DOI: 10.1007/978-3-031-19849-6_31.
19. Erkek I., Irmak E. Enhancing Cybersecurity of a Hydroelectric Power Plant Through Digital Twin Modeling and Explainable AI // IEEE Access. 2025. vol. 13. pp. 41887–41908. DOI: 10.1109/ACCESS.2025.3547672.
20. Li Y., Guan P., Li T., Larsen K.G., Aiello M., Pedersen T.B., Huang T., Zhang Y. Digital Twin for Secure Peer-to-Peer Trading in Cyber-Physical Energy Systems // IEEE Transactions on Network Science and Engineering. 2025. vol. 12. no. 2. pp. 669–683. DOI: 10.1109/TNSE.2024.3507956.
21. Patel T., Jadav N.K., Rathod T., Tanwar S., Garg D., Shahinzadeh H. AI-based Secure Intrusion Detection Framework for Digital Twin-enabled Critical Infrastructure //

Proceedings of the 14th International Conference on Information and Knowledge Technology (IKT). 2023. pp. 24–29. DOI: 10.1109/IKT62039.2023.10433057.
22. Ma J., Guo Y., Fang C., Zhang Q. Digital-Twin-Based CPS Anomaly Diagnosis and Security Defense Countermeasure Recommendation // IEEE Internet of Things Journal. 2024. vol. 11. no. 10. pp. 18726–18738. DOI: 10.1109/JIOT.2024.3366904.
23. Ghenai C., Husein L.A., Nahlawi M.A., Hamid A.K., Bettayeb M. Recent trends of digital twin technologies in the energy sector: A comprehensive review // Sustainable Energy Technologies and Assessments. 2022. vol. 54. 102837 p. DOI: 10.1016/j.seta.2022.102837.
24. Yu W., Patros P., Young B., Klinac E., Walmsley T.G. Energy digital twin technology for industrial energy management: Classification, challenges and future // Renewable and Sustainable Energy Reviews. 2022. vol. 161. 112407 p. DOI: 10.1016/j.rser.2022.112407.
25. Ismail F.B., Al-Faiz H., Hasini H., Al-Bazi A., Kazem H.A. A comprehensive review of the dynamic applications of the digital twin technology across diverse energy sectors // Energy Strategy Reviews. 2024. vol. 52. 101334 p. DOI: 10.1016/j.esr.2024.101334.
26. Hashmi R., Liu H., Yavari A. Digital Twins for Enhancing Efficiency and Assuring Safety in Renewable Energy Systems: A Systematic Literature Review // Energies. 2024. vol. 17. no. 11. 2456 p. DOI: 10.3390/en17112456.
27. Stadtmann F., Rasheed A., Kvamsdal T., et al. Digital Twins in Wind Energy: Emerging Technologies and Industry-Informed Future Directions // IEEE Access. 2023. vol. 11. pp. 110762–110795. DOI: 10.1109/ACCESS.2023.3321320.
28. Semeraro C., Aljaghoub H., Abdelkareem M.A., Alami A.H., Olabi A.G. Digital twin in battery energy storage systems: Trends and gaps detection through association rule mining // Energy. 2023. vol. 273. 127086 p. DOI: 10.1016/j.energy.2023.127086.
29. Wanasinghe T.R., Wroblewski L., Petersen B.K., et al. Digital Twin for the Oil and Gas Industry: Overview, Research Trends, Opportunities, and Challenges // IEEE Access. 2020. vol. 8. pp. 104175–104197. DOI: 10.1109/ACCESS.2020.2998723.
30. Amaral J.V.S.D., Santos C.H.D., Montevechi J.A.B., De Queiroz A.R. Energy Digital Twin applications: A review // Renewable and Sustainable Energy Reviews. 2023. vol. 188. 113891 p. DOI: 10.1016/j.rser.2023.113891.
31. Heluany J.B., Gkioulos V. A review on digital twins for power generation and distribution // International Journal of Information Security. 2023. vol. 23. pp. 1171–1195. DOI: 10.1007/s10207-023-00784-x.
32. Das O., Zafar M.H., Sanfilippo F., Rudra S., Kolhe M.L. Advancements in digital twin technology and machine learning for energy systems: A comprehensive review of applications in smart grids, renewable energy, and electric vehicle optimisation // Energy Conversion and Management: X. 2024. vol. 24. 100715 p. DOI: 10.1016/j.ecmx.2024.100715.
33. Ba L., Tangour F., Abbassi I.E., Absi R. Analysis of Digital Twin Applications in Energy Efficiency: A Systematic Review // Sustainability. 2025. vol. 17. no. 8. 3560 p. DOI: 10.3390/su17083560.
34. Koirala B., Cai H., Khayatian F., Munoz E., An J.G., Mutschler R., et al. Digitalization of Urban Multi-Energy Systems – Advances in Digital Twin Applications across Life-Cycle Phases // Advances in Applied Energy. 2024. vol. 16. 100196 p. DOI: 10.1016/j.adapen.2024.100196.
35. Gourisetti S.N.G., Bhadra S., Sebastian-Cardenas D.J., Touhiduzzaman Md., Ahmed O. A Theoretical Open Architecture Framework and Technology Stack for Digital Twins in Energy Sector Applications // Energies. 2023. vol. 16. no. 13. 4853 p. DOI: 10.3390/en16134853.

36. Boukaf M., Fadli F., Meskin N. A Comprehensive Review of Digital Twin Technology in Building Energy Consumption Forecasting // IEEE Access. 2024. vol. 12. pp. 187778–187799. DOI: 10.1109/ACCESS.2024.3498107.
37. Maksimovic M., Jokic S., Boskovic M.C. Innovative Horizons for Sustainable Smart Energy: Exploring the Synergy of 5G and Digital Twin Technologies // Process Integration and Optimization for Sustainability. 2025. vol. 9. pp. 431–470. DOI: 10.1007/s41660-024-00478-4.
38. Wang Y., Kang X., Chen Z. A survey of Digital Twin techniques in smart manufacturing and management of energy applications // Green Energy and Intelligent Transportation. 2022. vol. 1. no. 2. 100014 p. DOI: 10.1016/j.geits.2022.100014.
39. Masi M., Sellitto G.P., Aranha H., Pavleska T. Securing critical infrastructures with a cybersecurity digital twin // Software and Systems Modeling. 2023. vol. 22. pp. 689–707. DOI: 10.1007/s10270-022-01075-0.
40. Gulyamov S., Akhmedov A., Bazarov S., Ubaydullaeva A., Musaev S., Rodionov A., Odilkhujaev I. Using Digital Twins for Modeling and Testing Cybersecurity Scenarios in Smart Cities // Proceedings of the 6th International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA). 2024. pp. 655–658. DOI: 10.1109/SUMMA64428.2024.10803689.
41. Alhamam N., Rahman M.M.H., Aljughaiman A. A Comprehensive Review on Cybersecurity of Digital Twins Issues, Challenges, and Future Research Directions // IEEE Access. 2023. vol. 13. pp. 45106–45124. DOI: 10.1109/ACCESS.2025.3545004.
42. Coppolino L., Nardone R., Petruolo A., Romano L., Souvent A. Exploiting Digital Twin technology for Cybersecurity Monitoring in Smart Grids // Proceedings of the 18th International Conference on Availability, Reliability and Security. 2023. pp. 1–10. DOI: 10.1145/3600160.3605043.
43. Srivastava A., Liu C.-C., Ştefanov A., Basumallik S., et al. Digital Twins Serving Cybersecurity: More Than a Model: Cybersecurity as a Future Benefit of Digital Twins 2 // IEEE Power and Energy Magazine. 2024. vol. 22. no. 1. pp. 61–71. DOI: 10.1109/MPE.2023.3325196.
44. Olivares-Rojas J.C., Reyes-Archundia E., Gutierrez-Gnecchi J.A., et al. Towards Cybersecurity of the Smart Grid Using Digital Twins // IEEE Internet Computing. 2021. vol. 26. pp. 52–57. DOI: 10.1109/MIC.2021.3063674.
45. Sarker I.H., Janicke H., Mohsin A., Gill A.Q., Maglaras L. Explainable AI for cybersecurity automation, intelligence and trustworthiness in digital twin: Methods, taxonomy, challenges and prospects // ICT Express. 2024. vol. 10. pp. 935–958. DOI: 10.1016/j.icte.2024.05.007.
46. Suhail S., Iqbal M., Mclaughlin K. Digital-twin-driven deception platform: vision and way forward // IEEE Internet Computing. 2024. vol. 28. no. 4. pp. 40–47. DOI: 10.1109/MIC.2024.3406188.
47. Alcaraz C., Lopez J. Digital Twin: A Comprehensive Survey of Security Threats // IEEE Communications Surveys and Tutorials. 2022. vol. 24. no. 3. pp. 1475–1503. DOI: 10.1109/COMST.2022.3171465.
48. Holmes D., Papathanasaki M., Maglaras L., Ferrag M.A., Nepal S., Janicke H. Digital Twins and Cyber Security – solution or challenge? // Proceedings of the 2021 6th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM). 2021. pp. 1–8. DOI: 10.1109/SEEDA-CECNSM53056.2021.9566277.
49. Singh A. Enhancing Cybersecurity for Digital Twins: Challenges and Solutions // International Journal of Science and Technology. 2024. vol. 15. no. 4. pp. 1–9. DOI: 10.71097/ijsat.v15.i4.1651.
50. Kumar R., Aljuhani A., Javeed D., Kumar P., Islam S., Islam A.K.M.N. Digital Twins-enabled Zero Touch Network: A smart contract and explainable AI integrated

cybersecurity framework // Future Generation Computer Systems. 2024. vol. 156. pp. 191–205. DOI: 10.1016/j.future.2024.02.015.
51. Ravinder M., Kulkarni V. Smart Grid Anomaly Detection Using MFDA and Dilated GRU-based Neural Networks // Smart Grids and Sustainable Energy. 2025. vol. 10. 9 p. DOI: 10.1007/s40866-024-00216-2.
52. Alkuwari A.N., Al-Kuwari S., Albaseer A., Qaraqe M. Anomaly Detection on Smart Grids with Optimized Convolutional Long Short-Term Memory Model // IEEE Access. 2025. vol. 13. pp. 40399–40412. DOI: 10.1109/ACCESS.2025.3547037.
53. Yu L., Zhang X., Du L., Yue L. Anomaly Detection of Cyber Attacks in Smart Grid Communications Based on Residual Recurrent Neural Networks // Security and Privacy. 2025. vol. 8. no. 1. DOI: 10.1002/spy2.498.
54. Jithish J., Alangot B., Mahalingam N., Yeo K.S. Distributed Anomaly Detection in Smart Grids: A Federated Learning-Based Approach // IEEE Access. 2016. vol. 11. pp. 7157–7179. DOI: 10.1109/ACCESS.2023.3237554.
55. Abdel-Basset M., Moustafa N., Hawash H. Privacy-Preserved Generative Network for Trustworthy Anomaly Detection in Smart Grids: A Federated Semisupervised Approach // IEEE Transactions on Industrial Informatics. 2023. vol. 19. no. 1. pp. 995–1005. DOI: 10.1109/TII.2022.3165869.
56. Zhang Y., Fang X., Hordiichuk-Bublivska O., Beshley H., Beshley M. Modified Masking-Based Federated Singular Value Decomposition Method for Fast Anomaly Detection in Smart Grid Systems // Energies. 2023. vol. 16. no. 16. 5996 p. DOI: 10.3390/en16165996.
57. Fenza G., Gallo M., Loia V. Drift-Aware Methodology for Anomaly Detection in Smart Grid // IEEE Access. 2019. vol. 7. pp. 9645–9657. DOI: 10.1109/ACCESS.2019.2891315.
58. Li Y., Bai Y., Yang R., Feng Z., He W. Interpretable adaptive fault detection method for smart grid based on belief rule base // Scientific Reports. 2025. vol. 15. 7646 p. DOI: 10.1038/s41598-025-91897-x.
59. Hu P., Gao W., Li Y., Guo X., Hua F., Qiao L. Anomaly Detection and State Correction in Smart Grid Using EKF and Data Compensation Techniques // IEEE Sensors Journal. 2024. vol. 24. no. 8. pp. 12995–13009. DOI: 10.1109/JSEN.2024.3372973.
60. Gaggero G.B., Girdinio P., Marchese M. Artificial Intelligence and Physics-Based Anomaly Detection in the Smart Grid: A Survey // IEEE Access. 2025. vol. 13. pp. 23597–23606. DOI: 10.1109/ACCESS.2025.3537410.
61. Karimipour H., Geris S., Dehghantanha A., Leung H. Intelligent Anomaly Detection for Large-scale Smart Grids // Proceedings of the IEEE Canadian Conference of Electrical and Computer Engineering (CCECE). 2019. pp. 1–4. DOI: 10.1109/CCECE.2019.8861995.
62. Митяков Е.С. Разработка прототипа цифрового двойника автоматизированной системы управления интеллектуальной энергосетью для анализа угроз информационной безопасности // Computational Nanotechnology. 2025. Т. 12. №4. С. 116–123. DOI: 10.33693/2313-223X-2025-12-4-116-123.
63. Митяков Е.С. Метод обнаружения признаков угроз информационной безопасности объектов критической информационной инфраструктуры на основе цифровых двойников // Computational Nanotechnology. 2025. Т. 12. №3. С. 115–122. DOI: 10.33693/2313-223X-2025-12-3-115-122.
64. Kochergin S.V., Artemova S.V., Bakaev A.A., Mityakov E.S., Vegera Zh.G., Maksimova E.A. Cybersecurity of smart grids: Comparison of machine learning approaches training for anomaly detection // Russian Technological Journal. 2024. vol. 12. no. 6. pp. 7–19. DOI: 10.32362/2500-316X-2024-12-6-7-19.

65. Dong H., Kotenko I. Cybersecurity in the AI era: analyzing the impact of machine learning on intrusion detection // Knowledge and Information Systems. 2025. vol. 67. pp. 3915–3966. DOI: 10.1007/s10115-025-02366-w.
66. Abdi A., Bennouri H., Keane A. Cyber Resilience, Risk Management, and Security Challenges in Enterprise-Scale Cloud Systems: Comprehensive Review // Proceedings of the 13th Mediterranean Conference on Embedded Computing (MECO). 2024. pp. 1–8. DOI: 10.1109/MECO62516.2024.10577956.
67. Guide for Cybersecurity Event Recovery: NIST Special Publication 800-184. Gaithersburg, MD, USA: NIST, 2023. 53 p. URL: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-184.pdf (дата обращения: 06.02.2026).
68. Vielberth M., Bohm F., Fichtinger I., Pernul G. Security Operations Center: A Systematic Study and Open Challenges // IEEE Access. 2020. vol. 8. pp. 227756–227779. DOI: 10.1109/ACCESS.2020.3045514.
69. Kotenko I., Saenko I., Lauta O., Ichetovkin E., Sadovnikov V., Li W. Analysis of Modern Research on Protection against Adversarial Attacks in Energy Systems // Информатика и автоматизация. 2025. Т. 24. №6. С. 1751–1809. DOI: 10.15622/ia.24.6.8.

Котенко Игорь Витальевич — д-р техн. наук, профессор, заслуженный деятель науки российской федерации, главный научный сотрудник, руководитель лаборатории, лаборатория проблем компьютерной безопасности, Санкт-Петербургский Федеральный исследовательский цент Российской академии наук (СПб ФИЦ РАН). Область научных интересов: безопасность компьютерных сетей, управление политиками безопасности, контроль доступа, аутентификация, анализ защищенности, обнаружение компьютерных атак, защита от вирусов и сетевых червей, анализ и верификация протоколов безопасности, анализ и верификация систем информационной безопасности, защита программного обеспечения от взлома, управление цифровыми правами, технологии моделирования и визуализации для противодействия кибертерроризму. Число научных публикаций — 1000. ivkote@comsec.spb.ru; 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7(812)328-7181; факс: +7(812)328-4450.

Саенко Игорь Борисович — д-р техн. наук, профессор, главный научный сотрудник, лаборатория проблем компьютерной безопасности, Федеральное государственное бюджетное учреждение науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПБ ФИЦ РАН). Область научных интересов: автоматизированные информационные системы, информационная безопасность, искусственный интеллект, машинное обучение, теория моделирования и математическая статистика, теория информации. Число научных публикаций — 500. ibsaen@comsec.spb.ru; 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7(812)328-7181; факс: +7(812)328-4450.

Митяков Евгений Сергеевич — д-р экон. наук, профессор, заведующий кафедрой, кафедра предметно-ориентированные информационные системы, Федеральное государственное бюджетное образовательное учреждение высшего образования "МИРЭА - Российский технологический университет" (РТУ МИРЭА). Область научных интересов: информационная безопасность, безопасность критической информационной инфраструктуры, моделирование процессов управления инцидентами, экономическая безопасность, математическое моделирование. Число научных публикаций — 250. mityakov@mirea.ru; ул. Стромынка, 20, 107996, Москва, Россия; р.т.: +7(499)600-8080.

Поддержка исследований. Работа выполнена при частичной финансовой поддержке бюджетной темы FFZF-2025-0016.

DOI 10.15622/ia.25.3.8

I. KOTENKO, I. SAENKO, E. MITYAKOV

USING A DIGITAL TWIN FOR CYBER THREAT DETECTION IN INTELLIGENT POWER GRID CONTROL SYSTEMS

Kotenko I., Saenko I., Mityakov E. **Using a Digital Twin for Cyber Threat Detection in Intelligent Power Grid Control Systems.**

Abstract. The digital transformation of the energy sector, accompanied by the widespread deployment of smart grids, expands the cyber attack surface and increases the vulnerability of critical infrastructure. Traditional signature-based security systems demonstrate limited effectiveness against emerging threat types, while their verification on operational facilities entails unacceptable risks. This paper presents and experimentally validates an integrated adaptive framework based on a Digital Twin (DT), aimed at enhancing the effectiveness of cyber threat detection in intelligent power grid control systems. The framework includes a formal mathematical formulation of the detection problem considering three key metrics: Recall, False Positive Rate (FPR), and Total Time to Respond (TTR), together with a loss function for their joint optimization. The key element of the framework is a closed-loop adaptation governed by quality metrics, driven by a three-level mechanism (operational, tactical, and strategic modes): the performance results of the threat detection system on a real-world facility are used to automatically trigger incremental training, optimization, or generation of new scenarios within the secure virtual DT environment. The digital twin is formalized as a tuple of interconnected models of physical processes, communication infrastructure, control logic, and a threat knowledge base, enabling the framework to account for the multi-layer architecture of intelligent power grids and specific protocols (IEC 61850, Modbus). Experimental validation on a synthetic dataset simulating the operation of an industrial control system confirmed the applicability of the concept: the use of synthetic data generated in the DT environment reduced the false positive rate by 6.4% compared to training on static data, while activation of the adaptation mechanism contributed to a reduction in the aggregate model response time to incidents by 3.82%. The computation of the composite loss function showed a 16.8% reduction in the overall quality metric after adaptation. The proposed approach contributes to addressing the challenges of representative data scarcity and constraints on safe testing, and may be considered a step towards implementing the “security by design” paradigm in the energy sector. The obtained results constitute a proof of concept and substantiate the need for further research towards industrial deployment, including validation on real-world industrial scenarios.

Keywords: intelligent power grids, digital twin, adaptive threat detection, cyber-resilience, control systems.

References

1. Gehrmann C., Gunnarsson M. A Digital Twin Based Industrial Automation and Control System Security Architecture. IEEE Transactions on Industrial Informatics. 2020. vol. 16. no. 1. pp. 669–680. DOI: 10.1109/TII.2019.2938885.
2. Venkatachary S.K., Prasad J., Alagappan A., Andrews L.J.B., Raj R.A., Duraisamy S. Cybersecurity and cyber-terrorism challenges to energy-related infrastructures – Cybersecurity frameworks and economics – Comprehensive review. International Journal of Critical Infrastructure Protection. 2024. vol. 45. 100677 p. DOI: 10.1016/j.ijcip.2024.100677.
3. Diaba S.Y., Shafie-khah M., Elmusrati M. Cyber-physical attack and the future energy systems: A review. Energy Reports. 2024. vol. 12. pp. 2914–2932. DOI: 10.1016/j.egyr.2024.08.060.

4. Kotenko I., Saenko I., Lauta O., Kribel A. An Approach to Detecting Cyber Attacks against Smart Power Grids Based on the Analysis of Network Traffic Self-Similarity. Energies. 2020. vol. 13. no. 19. 5031 p. DOI: 10.3390/en13195031.
5. Izrailov K., Buinevich M., Kotenko I. Intelligent Detection of Cyber Attacks on Electrical Power Systems Based on Simulation and Graph-Based Modeling. Proceedings of the International Conference Intelligent Systems (INTELS'24). Communications in Computer and Information Science (CCIS). Springer. 2026. vol. 2603. pp. 231–245. DOI: 10.1007/978-3-032-04758-8_18.
6. Lin H. Computer network information security monitoring system. Proceedings of Second International Conference on Big Data, Computational Intelligence, and Applications (BDCIA 2024). 2025. vol. 13550. DOI: 10.1117/12.3059895.
7. Villegas-Ch W., Govea J., Gutierrez R., Navarro A.M., Mera-Navarrete A. Effectiveness of an Adaptive Deep Learning-Based Intrusion Detection System. IEEE Access. 2024. vol. 12. pp. 184010–184027. DOI: 10.1109/ACCESS.2024.3512363.
8. Kotenko I., Saenko I., Lauta O., Kribel A. A Proactive Protection of Smart Power Grids against Cyberattacks on Service Data Transfer Protocols by Computational Intelligence Methods. Sensors. 2022. vol. 22. no. 19. 7506 p. DOI: 10.3390/s22197506.
9. Desnitsky V.A., Kotenko I.V., Nogin S.B. Detection of Anomalies in Data for Monitoring of Security Components in the Internet of Things. Proceedings of the XVIII IEEE International Conference on Soft Computing and Measurements (SCM). 2015. pp. 189–192. DOI: 10.1109/SCM.2015.7190452.
10. Aydin Z. Detecting Cybersecurity Threats in Digital Energy Systems Using Deep learning for Imbalanced Datasets. International Journal of Energy Economics and Policy. 2025. vol. 15. no. 3. pp. 614–628. DOI: 10.32479/ijeep.19649.
11. Cremer F., Sheehan B., Fortmann M., et al. Cyber risk and cybersecurity: a systematic review of data availability. The Geneva Papers on Risk and Insurance - Issues and Practice. 2022. vol. 47. pp. 698–736. DOI: 10.1057/s41288-022-00266-6.
12. Kotenko I., Chechulin A. Computer attack modeling and security evaluation based on attack graphs. Proceedings of the IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). 2013. pp. 614–619. DOI: 10.1109/IDAACS.2013.6662998.
13. Homaei M., Mogollon-Gutierrez O., Sancho J.C., et al. A review of digital twins and their application in cybersecurity based on artificial intelligence. Artificial Intelligence Review. 2024. vol. 57. 201 p. DOI: 10.1007/s10462-024-10805-3.
14. McLaughlin K.L. The power of digital twins in the cybersecurity mesh. EDPACS. 2023. vol. 68. no. 6. pp. 35–39. DOI: 10.1080/07366981.2023.2263214.
15. Krishnaveni S., Chen T.M., Sathiyanarayanan M., et al. CyberDefender: an integrated intelligent defense framework for digital-twin-based industrial cyber-physical systems. Cluster Computing. 2024. vol. 27. pp. 7273–7306. DOI: 10.1007/s10586-024-04320-x.
16. Salim M.M., Camacho D., Park J.H. Digital Twin and federated learning enabled cyberthreat detection system for IoT networks. Future Generation Computer Systems. 2024. vol. 161. pp. 701–713. DOI: 10.1016/j.future.2024.07.017.
17. Sousa B., Arieiro M., Pereira V., Correia J., Lourenco N., Cruz T. ELEGANT: Security of Critical Infrastructures With Digital Twins. IEEE Access. 2021. vol. 9. pp. 107574–107588. DOI: 10.1109/ACCESS.2021.3100708.
18. De Hoz D.J., Temperekidis A., Katsaros P., Konstantinou C. An IoT Digital Twin for Cyber-Security Defence Based on Runtime Verification. Lecture Notes in Computer Science. Springer. 2022. vol. 13701. pp. 556–574. DOI: 10.1007/978-3-031-19849-6_31.
19. Erkek I., Irmak E. Enhancing Cybersecurity of a Hydroelectric Power Plant Through Digital Twin Modeling and Explainable AI. IEEE Access. 2025. vol. 13. pp. 41887–41908. DOI: 10.1109/ACCESS.2025.3547672.

20. Li Y., Guan P., Li T., Larsen K.G., Aiello M., Pedersen T.B., Huang T., Zhang Y. Digital Twin for Secure Peer-to-Peer Trading in Cyber-Physical Energy Systems. IEEE Transactions on Network Science and Engineering. 2025. vol. 12. no. 2. pp. 669–683. DOI: 10.1109/TNSE.2024.3507956.
21. Patel T., Jadav N.K., Rathod T., Tanwar S., Garg D., Shahinzadeh H. AI-based Secure Intrusion Detection Framework for Digital Twin-enabled Critical Infrastructure. Proceedings of the 14th International Conference on Information and Knowledge Technology (IKT). 2023. pp. 24–29. DOI: 10.1109/IKT62039.2023.10433057.
22. Ma J., Guo Y., Fang C., Zhang Q. Digital-Twin-Based CPS Anomaly Diagnosis and Security Defense Countermeasure Recommendation. IEEE Internet of Things Journal. 2024. vol. 11. no. 10. pp. 18726–18738. DOI: 10.1109/JIOT.2024.3366904.
23. Ghenai C., Husein L.A., Nahlawi M.A., Hamid A.K., Bettayeb M. Recent trends of digital twin technologies in the energy sector: A comprehensive review. Sustainable Energy Technologies and Assessments. 2022. vol. 54. 102837 p. DOI: 10.1016/j.seta.2022.102837.
24. Yu W., Patros P., Young B., Klinac E., Walmsley T.G. Energy digital twin technology for industrial energy management: Classification, challenges and future. Renewable and Sustainable Energy Reviews. 2022. vol. 161. 112407 p. DOI: 10.1016/j.rser.2022.112407.
25. Ismail F.B., Al-Faiz H., Hasini H., Al-Bazi A., Kazem H.A. A comprehensive review of the dynamic applications of the digital twin technology across diverse energy sectors. Energy Strategy Reviews. 2024. vol. 52. 101334 p. DOI: 10.1016/j.esr.2024.101334.
26. Hashmi R., Liu H., Yavari A. Digital Twins for Enhancing Efficiency and Assuring Safety in Renewable Energy Systems: A Systematic Literature Review. Energies. 2024. vol. 17. no. 11. 2456 p. DOI: 10.3390/en17112456.
27. Stadtmann F., Rasheed A., Kvamsdal T., et al. Digital Twins in Wind Energy: Emerging Technologies and Industry-Informed Future Directions. IEEE Access. 2023. vol. 11. pp. 110762–110795. DOI: 10.1109/ACCESS.2023.3321320.
28. Semeraro C., Aljaghoub H., Abdelkareem M.A., Alami A.H., Olabi A.G. Digital twin in battery energy storage systems: Trends and gaps detection through association rule mining. Energy. 2023. vol. 273. 127086 p. DOI: 10.1016/j.energy.2023.127086.
29. Wanasinghe T.R., Wroblewski L., Petersen B.K., et al. Digital Twin for the Oil and Gas Industry: Overview, Research Trends, Opportunities, and Challenges. IEEE Access. 2020. vol. 8. pp. 104175–104197. DOI: 10.1109/ACCESS.2020.2998723.
30. Amaral J.V.S.D., Santos C.H.D., Montevechi J.A.B., De Queiroz A.R. Energy Digital Twin applications: A review. Renewable and Sustainable Energy Reviews. 2023. vol. 188. 113891 p. DOI: 10.1016/j.rser.2023.113891.
31. Heluany J.B., Gkioulos V. A review on digital twins for power generation and distribution. International Journal of Information Security. 2023. vol. 23. pp. 1171–1195. DOI: 10.1007/s10207-023-00784-x.
32. Das O., Zafar M.H., Sanfilippo F., Rudra S., Kolhe M.L. Advancements in digital twin technology and machine learning for energy systems: A comprehensive review of applications in smart grids, renewable energy, and electric vehicle optimization. Energy Conversion and Management: X. 2024. vol. 24. 100715 p. DOI: 10.1016/j.ecmx.2024.100715.
33. Ba L., Tangour F., Abbassi I.E., Absi R. Analysis of Digital Twin Applications in Energy Efficiency: A Systematic Review. Sustainability. 2025. vol. 17. no. 8. 3560 p. DOI: 10.3390/su17083560.
34. Koirala B., Cai H., Khayatian F., Munoz E., An J.G., Mutschler R., et al. Digitalization of Urban Multi-Energy Systems – Advances in Digital Twin

Applications across Life-Cycle Phases. Advances in Applied Energy. 2024. vol. 16. 100196 p. DOI: 10.1016/j.adapen.2024.100196.
35. Gourisetti S.N.G., Bhadra S., Sebastian-Cardenas D.J., Touhiduzzaman Md., Ahmed O. A Theoretical Open Architecture Framework and Technology Stack for Digital Twins in Energy Sector Applications. Energies. 2023. vol. 16. no. 13. 4853 p. DOI: 10.3390/en16134853.
36. Boukaf M., Fadli F., Meskin N. A Comprehensive Review of Digital Twin Technology in Building Energy Consumption Forecasting. IEEE Access. 2024. vol. 12. pp. 187778–187799. DOI: 10.1109/ACCESS.2024.3498107.
37. Maksimovic M., Jokic S., Boskovic M.C. Innovative Horizons for Sustainable Smart Energy: Exploring the Synergy of 5G and Digital Twin Technologies. Process Integration and Optimization for Sustainability. 2025. vol. 9. pp. 431–470. DOI: 10.1007/s41660-024-00478-4.
38. Wang Y., Kang X., Chen Z. A survey of Digital Twin techniques in smart manufacturing and management of energy applications. Green Energy and Intelligent Transportation. 2022. vol. 1. no. 2. 100014 p. DOI: 10.1016/j.geits.2022.100014.
39. Masi M., Sellitto G.P., Aranha H., Pavleska T. Securing critical infrastructures with a cybersecurity digital twin. Software and Systems Modeling. 2023. vol. 22. pp. 689–707. DOI: 10.1007/s10270-022-01075-0.
40. Gulyamov S., Akhmedov A., Bazarov S., Ubaydullaeva A., Musaev S., Rodionov A., Odilkhujaev I. Using Digital Twins for Modeling and Testing Cybersecurity Scenarios in Smart Cities. Proceedings of the 6th International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA). 2024. pp. 655–658. DOI: 10.1109/SUMMA64428.2024.10803689.
41. Alhamam N., Rahman M.M.H., Aljughaiman A. A Comprehensive Review on Cybersecurity of Digital Twins Issues, Challenges, and Future Research Directions. IEEE Access. 2023. vol. 13. pp. 45106–45124. DOI: 10.1109/ACCESS.2025.3545004.
42. Coppolino L., Nardone R., Petruolo A., Romano L., Souvent A. Exploiting Digital Twin technology for Cybersecurity Monitoring in Smart Grids. Proceedings of the 18th International Conference on Availability, Reliability and Security. 2023. pp. 1–10. DOI: 10.1145/3600160.3605043.
43. Srivastava A., Liu C.-C., Ştefanov A., Basumallik S., et al. Digital Twins Serving Cybersecurity: More Than a Model: Cybersecurity as a Future Benefit of Digital Twins 2. IEEE Power and Energy Magazine. 2024. vol. 22. no. 1. pp. 61–71. DOI: 10.1109/MPE.2023.3325196.
44. Olivares-Rojas J.C., Reyes-Archundia E., Gutierrez-Gnecchi J.A., et al. Towards Cybersecurity of the Smart Grid Using Digital Twins. IEEE Internet Computing. 2021. vol. 26. pp. 52–57. DOI: 10.1109/MIC.2021.3063674.
45. Sarker I.H., Janicke H., Mohsin A., Gill A.Q., Maglaras L. Explainable AI for cybersecurity automation, intelligence and trustworthiness in digital twin: Methods, taxonomy, challenges and prospects. ICT Express. 2024. vol. 10. pp. 935–958. DOI: 10.1016/j.icte.2024.05.007.
46. Suhail S., Iqbal M., Mclaughlin K. Digital-twin-driven deception platform: vision and way forward. IEEE Internet Computing. 2024. vol. 28. no. 4. pp. 40–47. DOI: 10.1109/MIC.2024.3406188.
47. Alcaraz C., Lopez J. Digital Twin: A Comprehensive Survey of Security Threats. IEEE Communications Surveys and Tutorials. 2022. vol. 24. no. 3. pp. 1475–1503. DOI: 10.1109/COMST.2022.3171465.
48. Holmes D., Papathanasaki M., Maglaras L., Ferrag M.A., Nepal S., Janicke H. Digital Twins and Cyber Security – solution or challenge?. Proceedings of the 2021 6th South-East Europe Design Automation, Computer Engineering, Computer Networks

and Social Media Conference (SEEDA-CECNSM). 2021. pp. 1–8. DOI: 10.1109/SEEDA-CECNSM53056.2021.9566277.
49. Singh A. Enhancing Cybersecurity for Digital Twins: Challenges and Solutions. International Journal of Science and Technology. 2024. vol. 15. no. 4. pp. 1–9. DOI: 10.71097/ijsat.v15.i4.1651.
50. Kumar R., Aljuhani A., Javeed D., Kumar P., Islam S., Islam A.K.M.N. Digital Twins-enabled Zero Touch Network: A smart contract and explainable AI integrated cybersecurity framework. Future Generation Computer Systems. 2024. vol. 156. pp. 191–205. DOI: 10.1016/j.future.2024.02.015.
51. Ravinder M., Kulkarni V. Smart Grid Anomaly Detection Using MFDA and Dilated GRU-based Neural Networks. Smart Grids and Sustainable Energy. 2025. vol. 10. 9 p. DOI: 10.1007/s40866-024-00216-2.
52. Alkuwari A.N., Al-Kuwari S., Albaseer A., Qaraqe M. Anomaly Detection on Smart Grids with Optimized Convolutional Long Short-Term Memory Model. IEEE Access. 2025. vol. 13. pp. 40399–40412. DOI: 10.1109/ACCESS.2025.3547037.
53. Yu L., Zhang X., Du L., Yue L. Anomaly Detection of Cyber Attacks in Smart Grid Communications Based on Residual Recurrent Neural Networks. Security and Privacy. 2025. vol. 8. no. 1. DOI: 10.1002/spy2.498.
54. Jithish J., Alangot B., Mahalingam N., Yeo K.S. Distributed Anomaly Detection in Smart Grids: A Federated Learning-Based Approach. IEEE Access. 2016. vol. 11. pp. 7157–7179. DOI: 10.1109/ACCESS.2023.3237554.
55. Abdel-Basset M., Moustafa N., Hawash H. Privacy-Preserved Generative Network for Trustworthy Anomaly Detection in Smart Grids: A Federated Semisupervised Approach. IEEE Transactions on Industrial Informatics. 2023. vol. 19. no. 1. pp. 995–1005. DOI: 10.1109/TII.2022.3165869.
56. Zhang Y., Fang X., Hordiichuk-Bublivska O., Beshley H., Beshley M. Modified Masking-Based Federated Singular Value Decomposition Method for Fast Anomaly Detection in Smart Grid Systems. Energies. 2023. vol. 16. no. 16. 5996 p. DOI: 10.3390/en16165996.
57. Fenza G., Gallo M., Loia V. Drift-Aware Methodology for Anomaly Detection in Smart Grid. IEEE Access. 2019. vol. 7. pp. 9645–9657. DOI: 10.1109/ACCESS.2019.2891315.
58. Li Y., Bai Y., Yang R., Feng Z., He W. Interpretable adaptive fault detection method for smart grid based on belief rule base. Scientific Reports. 2025. vol. 15. 7646 p. DOI: 10.1038/s41598-025-91897-x.
59. Hu P., Gao W., Li Y., Guo X., Hua F., Qiao L. Anomaly Detection and State Correction in Smart Grid Using EKF and Data Compensation Techniques. IEEE Sensors Journal. 2024. vol. 24. no. 8. pp. 12995–13009. DOI: 10.1109/JSEN.2024.3372973.
60. Gaggero G.B., Girdinio P., Marchese M. Artificial Intelligence and Physics-Based Anomaly Detection in the Smart Grid: A Survey. IEEE Access. 2025. vol. 13. pp. 23597–23606. DOI: 10.1109/ACCESS.2025.3537410.
61. Karimipour H., Geris S., Dehghantanha A., Leung H. Intelligent Anomaly Detection for Large-scale Smart Grids. Proceedings of the IEEE Canadian Conference of Electrical and Computer Engineering (CCECE). 2019. pp. 1–4. DOI: 10.1109/CCECE.2019.8861995.
62. Mityakov E.S. [Development of a prototype of a digital twin of an automated smart grid control system for analyzing information security threats]. Komp'yuternye nanotekhnologii – Computational Nanotechnology. 2025. vol. 12. no. 4. pp. 116–123. DOI: 10.33693/2313-223X-2025-12-4-116-123. (In Russ.).
63. Mityakov E.S. [Method for detecting signs of information security threats to critical information infrastructure objects based on digital twins]. Komp'yuternye

nanotekhnologii – Computational Nanotechnology. 2025. vol. 12. no. 3. pp. 115–122. DOI: 10.33693/2313-223X-2025-12-3-115-122. (In Russ.).
64. Kochergin S.V., Artemova S.V., Bakaev A.A., Mityakov E.S., Vegera Zh.G., Maksimova E.A. Cybersecurity of smart grids: Comparison of machine learning approaches training for anomaly detection. Russian Technological Journal. 2024. vol. 12. no. 6. pp. 7–19. DOI: 10.32362/2500-316X-2024-12-6-7-19.
65. Dong H., Kotenko I. Cybersecurity in the AI era: analyzing the impact of machine learning on intrusion detection. Knowledge and Information Systems. 2025. vol. 67. pp. 3915–3966. DOI: 10.1007/s10115-025-02366-w.
66. Abdi A., Bennouri H., Keane A. Cyber Resilience, Risk Management, and Security Challenges in Enterprise-Scale Cloud Systems: Comprehensive Review. Proceedings of the 13th Mediterranean Conference on Embedded Computing (MECO). 2024. pp. 1–8. DOI: 10.1109/MECO62516.2024.10577956.
67. Guide for Cybersecurity Event Recovery: NIST Special Publication 800-184. Gaithersburg, MD, USA: NIST, 2023. 53 p. Available at: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-184.pdf (accessed 06.02.2026).
68. Vielberth M., Bohm F., Fichtinger I., Pernul G. Security Operations Center: A Systematic Study and Open Challenges. IEEE Access. 2020. vol. 8. pp. 227756–227779. DOI: 10.1109/ACCESS.2020.3045514.
69. Kotenko I., Saenko I., Lauta O., Ichetovkin E., Sadovnikov V., Li W. [Analysis of Modern Research on Protection against Adversarial Attacks in Energy Systems]. Informatika i avtomatizatsiya – Informatics and Automation. 2025. vol. 24. no. 6. pp. 1751–1809. DOI: 10.15622/ia.24.6.8. (In Russ.).

Kotenko Igor — Ph.D., Dr. Sci., Professor, Honored Scientist of the Russian Federation, Chief researcher, Head of the Laboratory, Laboratory of Computer Security Problems, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS). Research interests: computer network security, security policy management, access control, authentication, security analysis, detection of computer attacks, protection against viruses and network worms, analysis and verification of security protocols, analysis and verification of information security systems, software protection against hacking, digital rights management, modeling and visualization technologies for countering cyberterrorism. The number of publications — 1000. ivkote@comsec.spb.ru; 39, 14th line V.O., 199178, Saint-Petersburg, Russia; office phone: +7(812)328-7181; fax: +7(812)328-4450.

Saenko Igor — Ph.D., Dr. Sci., Professor, Chief scientific officer, Laboratory of Computer Security Problems, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS). Research interests: automated information systems, information security, artificial intelligence, machine learning, modeling theory and mathematical statistics, information theory. The number of publications — 500. ibsaen@comsec.spb.ru; 39, 14th line V.O., 199178, Saint-Petersburg, Russia; office phone: +7(812)328-7181; fax: +7(812)328-4450.

Mityakov Evgeniy — Ph.D., Dr. Sci., Professor, Head of the Department, Department of Subject-Oriented Information Systems, Federal State Budgetary Educational Institution of Higher Education "MIREA - Russian Technological University" (RTU MIREA). Research interests: information security, security of critical information infrastructure, modeling of incident management processes, economic security, mathematical modeling. The number of publications — 250. mityakov@mirea.ru; 20, Stromynka St., 107996, Moscow, Russia; office phone: +7(499)600-8080.

Acknowledgements. The reported study was partially funded by the budget project FFZF-2025-0016.