НИЯУ МИФИ

Национальный исследовательский ядерный университет «МИФИ»

Третья Всероссийская
научно-техническая конференция

«Кибернетика и информационная безопасность»

КИБ-2025

КИБЕРНЕТИКА
И ИНФОРМАЦИОННАЯ
БЕЗОПАСНОСТЬ

3 – 4 декабря 2025 г.

Сборник
научных трудов

Том 1

Москва 2025

Национальный исследовательский ядерный университет «МИФИ»

Третья Всероссийская научно-техническая конференция

«Кибернетика и информационная безопасность»

«КИБ-2025»

Сборник научных трудов

В двух томах

Том 1

3–4 декабря 2025 г., Москва

Москва 2025

УДК 004.056 : 001(06)
ББК 32.973.202
В85

Третья Всероссийская научно-техническая конференция «Кибернетика и информационная безопасность «КИБ-2025». Сборник научных трудов. 3-4 декабря 2025 г., Москва. В 2-х т. Т. 1. М.: НИЯУ МИФИ, 2025. 224 с.

Настоящая книга содержит тезисы научных работ и докладов, предложенных специалистами на конференции «КИБ-2025».

Представленные материалы выполнены преподавателями, научными сотрудниками, молодыми учеными, аспирантами и студентами МИФИ и других вузов, специалистами академических научных и научно-производственных организаций Москвы и России, сотрудничающих с МИФИ. Работы отражают достижения и уровень исследований, тенденции и проблемы в развитии и обеспечении образования и научно-исследовательских работ по актуальным вопросам информационной безопасности, решению задач по защите информации, построения информационных и интеллектуальных систем управления в защищенном исполнении.

Книга предназначена читателям, интересующимся тематикой представленных научных направлений.

Статьи сборника издаются в авторской редакции.

Материалы получены 25.10.2025

ISBN 978-5-7262-3201-0
ISBN 978-5-7262-3199-0 (т. 1)

СОДЕРЖАНИЕ

ДОВЕРЕННАЯ ЭЛЕКТРОННАЯ КОМПОНЕНТНАЯ БАЗА И ПАК ДЛЯ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ. ТЕХНОЛОГИЧЕСКИЙ СУВЕРЕНИТЕТ

ЗАЩИЩЕННЫЕ КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

ИНТЕЛЛЕКТУАЛЬНОЕ УПРАВЛЕНИЕ СЕТЕВОЙ БЕЗОПАСНОСТЬЮ

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ СОЦИОТЕХНИЧЕСКИХ СИСТЕМ

КИБ-2025
КИБЕРНЕТИКА
И ИНФОРМАЦИОННАЯ
БЕЗОПАСНОСТЬ

Направление

Доверенная электронная компонентная база и ПАК для критической информационной инфраструктуры. Технологический суверенитет

Руководители секции:

ДУРАКОВСКИЙ А.П. – к.т.н., доцент, директор аттестационно-испытательного центра информационной безопасности и систем защиты информации НИЯУ МИФИ

КЕССАРИНСКИЙ Л.Н. – к.т.н., доцент, зам. директора аттестационно-испытательного центра информационной безопасности и систем защиты информации НИЯУ МИФИ

УДК 004.056

М.Ю. ТОЛСТЫХ

Московский государственный лингвистический университет

К ВОПРОСУ ОБЕСПЕЧЕНИЯ ТЕХНОЛОГИЧЕСКОГО СУВЕРЕНИТЕТА ДЛЯ ЗАЩИТЫ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ

Обоснован приоритет технологического суверенитета как фундамента национальной безопасности государства. Выделены и охарактеризованы ключевые препятствия для его реализации. Предложены некоторые способы достижения технологической автономии, обеспечивающей устойчивость стратегических систем и укрепляющей конкурентоспособность России в глобальном цифровом пространстве.

В эпоху цифровой трансформации и глобализации критическая информационная инфраструктура (КИИ) играет ключевую роль в обеспечении национальной безопасности, экономической стабильности и социального благополучия страны. Многие государства, в том числе Российская Федерация, сталкиваются с санкциями, ограничивающими доступ к импортным технологиям, кроме того, согласно национальному законодательству [1], обеспечение технологического суверенитета (ТС) становится приоритетом для минимизации рисков внешнего влияния и шпионажа через так называемые «закладки» и «бэкдоры» в иностранном программном обеспечении и оборудовании, также подчеркивается необходимость самостоятельного развития [2] и эксплуатации программно-аппаратных комплексов для защиты информации (ПАК) [3].

В данном контексте ТС выступает как инструмент национальной обороны, позволяющий избежать «технологической блокады» и обеспечить непрерывность функционирования стратегических систем. Кроме того, обеспечение ТС позволяет разрабатывать отечественные ПАК с встроенной защитой, соответствующие национальным стандартам безопасности и требованиям регуляторов, минимизируя риски компрометации.

На современном этапе существует ряд проблем для национальной самостоятельной разработки, производства и контроля ключевых технологий и систем, необходимых для экономической независимости и национальной безопасности страны. Среди основных можно выделить следующие.

1) Технологическая зависимость и уязвимости. Существующие защитные решения основаны на импортных компонентах, что создает риски цепочек поставок и потенциальных «бэкдоров», а также тормозит полноценную реализацию стратегии импортозамещения.

2) Регуляторные и экономические барьеры. Высокая стоимость процессов создания новых продуктов, услуг и технологий, отсутствие инвестиций и несовершенство, а иногда и противоречивость нормативной базы препятствуют масштабированию отечественных ПАК.

Для преодоления выявленных проблем предлагаются следующие меры, которые ориентированы на долгосрочный эффект и сочетают государственную поддержку с рыночными механизмами.

1) Разработка модульных отечественных платформ с открытым исходным кодом. Предлагается создание стандартизированных модульных ПАК на базе российских процессоров и операционных систем с интеграцией open-source компонентов для оперативной адаптации, что позволит минимизировать уязвимости через коллективный аудит кода.

2) Оптимизация регуляторных механизмов, финансовые стимулы и партнерства, привлечение акселераторов. Видится целесообразным совершенствование нормативной базы через гармонизацию стандартов, при которой будут устранены противоречия и упрощена сертификация ПАК. С экономической стороны предлагается внедрить целевые гранты и налоговые вычеты для организаций, разрабатывающих суверенные ПАК.

В заключение целесообразно отметить, что ТС в ПАК для КИИ должен восприниматься в первую очередь как стратегическая автономия, обеспечивающая конкурентоспособность в глобальном цифровом пространстве, способствующая обеспечению национальной безопасности, экономической стабильности и социального благополучия России.

Список литературы

1. Федеральный закон «О безопасности критической информационной инфраструктуры Российской Федерации» от 26.07.2017 № 187-ФЗ [Электронный ресурс] // URL: https://www.consultant.ru/document/cons_doc_LAW_220885/ (дата обращения: 10.09.2025).

2. Указ Президента РФ «О дополнительных мерах по обеспечению информационной безопасности Российской Федерации» от 1.05.2022 № 250 [Электронный ресурс] // URL: https://base.garant.ru/404561984/ (дата обращения: 10.09.2025).

3. Постановление Правительства РФ «О порядке перехода субъектов критической информационной инфраструктуры Российской Федерации на преимущественное применение доверенных программно-аппаратных комплексов на принадлежащих им значимых объектах критической информационной инфраструктуры Российской Федерации» от 14.11.2023 № 1912 [Электронный ресурс] // https://www.garant.ru/products/ipo/prime/doc/407912691/ (дата обращения: 09.09.2025).

УДК 004.056.53:004.891.3

Н.В. КОРНЕЕВ

РГУ нефти и газа (НИУ) имени И.М. Губкина, Москва
Финансовый университет, Москва

МИКРОСЕРВИСНАЯ АРХИТЕКТУРА ПАТТЕРНА БЕЗОПАСНОСТИ ПРИ УГРОЗЕ МОДИФИКАЦИИ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ

С использованием методов криптографии предложено архитектурное решение на базе микросервисов, включающее новый механизм защиты. В структуре микросервисной архитектуры он обладает принципиально новыми возможностями: целостность дата-сета, за счет выполнения операций хеширования и создания ЭЦП хэш-сумм. Экспериментальные исследования показали, что в процессе контейнеризации приложения в виртуальной среде, предложенная архитектура обеспечивает безопасность модели машинного обучения на тестовых данных при угрозе модификации модели.

Современная микросервисная архитектура в настоящее время включает ряд агентов для машинного обучения (ML). Сами агенты зависят от данных, используемых в модели машинного обучения. Циркулирующие в агентах машинного обучения данные напрямую влияют на качество модели. Это качество в первую очередь определяется подмножеством тренировочных данных (дата-сет). Далее на режимах реальной эксплуатации агент машинного обучения, и собственно сама модель, использует данные из генеральной совокупности. Эти данные могут быть подвержены изменению. Такое изменение является результатом модификации работы модели и появления угрозы безопасности информации [1]. В этой работе для решения данной проблемы мы предлагаем использовать архитектурное решение на базе микросервисов, включающее новый механизм защиты в виде паттерна безопасности.

Для реализации паттерна нами использован специальный стек технологий (рис. 1), построенный с использованием методов криптографии и архитектурного решения на базе микросервисов, включающего новый механизм защиты. В качестве языка программирования использовался язык Java и фреймворк Spring. Все входящие HTTP запросы проходят и обрабатываются в шлюзе, который создан на основе Spring Cloud Gateway. За аутентификацию и авторизацию пользователей отвечает пользовательский микросервис (Users

Mircoservice). Микросервис машинного обучения (Machine Learning Mircoservice) реализует модель ML.

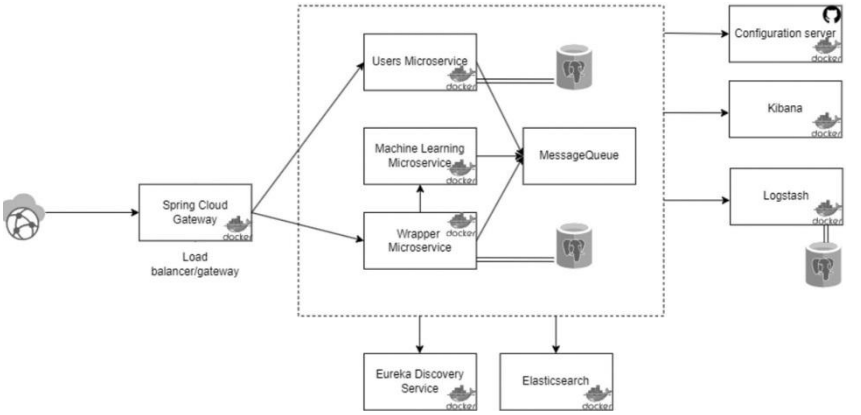

Рис. 1. Стек технологий

Безопасность модели машинного обучения обеспечивает микросервис Wrapper. Архитектурное решение механизма защиты включает концепцию вызовов, которые непосредственно проходят через него, таким образом реализуя эшелонированную оборону от угрозы порчи и/или подмены данных. Дополнительно все внешние данные проходят валидацию на стороне микросервиса. Целостность дата-сета, за счет выполнения операций хеширования и создания ЭЦП хэш-сумм реализуется с использованием БД обучающих данных. Этот специальный микросервис является доверенным сервисом между данными и моделью машинного обучения и реализует механизм защиты. Микросервис MessageQueue выполняет функционал по отправке асинхронных сообщений. Настроенная агрегация логов и мониторинг состояния системы функционирует на базе стека ELK (Elasticsearch, Logstash, Kibana). Микросервис Eureka Discovery Service комплексно выполняет основные функции управления всей системой, включая регистрацию микросервисов, формирование связей между ними, контроль и аудит состояние микросервисов.

Список литературы

1. Корнеев Н.В., Котрини Е.С. Паттерн для обеспечения безопасности приложения при угрозе модификации модели машинного обучения. Вопросы кибербезопасности. 2025, № 1(65), с. 117–127. DOI: http://dx.doi.org/10.21681/2311-3456-2025-1-117-127.

УДК 004.056

В.А. БУРКИН

РГУ нефти и газа (НИУ) имени И.М. Губкина, Москва

УТОЧНЕННАЯ МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ИНТЕГРАЛЬНОГО ПОКАЗАТЕЛЯ СОСТОЯНИЯ ЗАЩИЩЁННОСТИ ПРЕДПРИЯТИЯ ТЭК

Представлена уточнённая математическая модель интегрального показателя состояния защищённости (ПСЗ) для предприятий топливно-энергетического комплекса (ТЭК). Предлагается двухуровневая структура показателя с разделением на технический и организационный подиндексы.

Введение

Практика эксплуатации автоматизированных систем управления технологическими процессами (АСУ ТП) в ТЭК показывает, что формальное соответствие требованиям по защите информации само по себе не гарантирует фактической устойчивости технологического контура. Для принятия решений на уровне руководства необходим компактный, интерпретируемый и воспроизводимый показатель, агрегирующий разнородные частные метрики. Такой показатель должен быть прозрачен, сопоставим между объектами и чувствителен к критическим отклонениям [1–4]. В работе предлагается уточнённая математическая модель ПСЗ, учитывающая отраслевую специфику и отечественные методические подходы [5–7].

Цель – формализовать интегральный ПСЗ для оперативной оценки состояния защищённости и приоритизации мероприятий.

Математическая модель интегрального показателя

Пусть имеется набор нормированных метрик $K_1 \ldots K_{10} \in [0;1]$, из которых технические $K_1 \ldots K_7$ описывают состояние технологического контура АСУ ТП (сегментация, фильтрация трафика, контроль удалённого доступа, управление уязвимостями, мониторинг событий, резервирование/отказоустойчивость и т.п.):

$$I_{\text{тех}} = \sum_{i=1}^{7} w_i^{\text{тех}} K_i , \qquad \sum_{i=1}^{7} w_i^{\text{тех}} = 1, \tag{1}$$

а организационные $K_8 \ldots K_{10}$ – зрелость процессов (обучение, регламенты, исполнение предписаний):

$$I_{\text{орг}} = \sum_{i=8}^{10} w_i^{\text{орг}} K_i, \qquad \sum_{i=8}^{10} w_i^{\text{орг}} = 1. \tag{2}$$

Интегральное значение:

$$I = \alpha I_{\text{тех}} + (1-\alpha) I_{\text{орг}}, \;\; \alpha \in [0{,}6;\, 0{,}8]\,. \tag{3}$$

Стартовое значение α=0,7 целесообразно для отрасли, поскольку технические факторы непосредственно влияют на технологическую доступность и промышленную безопасность.

Зоны интерпретации

Применяется единая шкала:

– Зелёная зона: $I \geq 0{,}80$;

– Жёлтая зона: $0{,}60 \leq I < 0{,}80$;

– Красная зона: $I < 0{,}60$.

Цветовая оценка используется для выбора режима управления (плановый, корректирующий, кризисный) и запуска предписанных действий по каждому из просевших K_i.

Список литературы

1. ГОСТ Р ИСО/МЭК 27001-2021. Информационная безопасность, кибербезопасность и защита приватности. Системы менеджмента информационной безопасности. Требования.

2. МЭК 62443 (серия стандартов). Безопасность промышленных процессов и систем управления. Общие положения и требования к организациям, процессам и техническим средствам.

3. NIST SP 800-82. Руководство по защите автоматизированных систем управления технологическими процессами (последняя редакция). Национальный институт стандартов и технологий, США.

4. ФСТЭК России. Методические рекомендации по оценке показателя состояния технической защиты информации и обеспечения безопасности значимых объектов КИИ (актуальная редакция).

5. Милославская Н.Г., Сенаторов М.Ю. Проверка и оценка деятельности по управлению информационной безопасностью. 2-е изд. М.: Горячая линия – Телеком, 2022.

6. Милославская Н.Г., Толстой А.И. Управление информационной безопасностью. М.: НИЯУ МИФИ, 2020.

7. Правиков Д.И., Мурашкин В.А. Показатель состояния защищённости на объектах нефтегазовой отрасли // «Кибернетика и информационная безопасность – КИБ-2024». Сборник трудов. М.: НИЯУ МИФИ, 2024, с. 26–27.

УДК 004.056

И.С. ЛЕВЕНЕЦ

МИРЭА – Российский технологический университет, Москва

ПОДСИСТЕМА ПРЕДУПРЕЖДЕНИЯ КОМПЬЮТЕРНЫХ АТАК НА ОБЪЕКТЫ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ: АНАЛИЗ ФУНКЦИОНИРОВАНИЯ И СОВЕРШЕНСТВОВАНИЕ

В работе исследовались существующие требования и архитектурные решения подсистем предупреждения компьютерных атак в критической информационной инфраструктуре (КИИ) и предложены механизмы повышения их эффективности. Проведен анализ современных подходов к мониторингу и корреляции событий безопасности, а также структурам интеллектуальных сервисов защиты. В результате работы сформированы рекомендации по интеграции методов обнаружения инцидентов и повышения устойчивости функционирования КИИ. Основные выводы касаются оптимизации взаимодействия IDS, NTA и XDR-средств, а также внедрения адаптивных механизмов реагирования на угрозы.

Введение

В условиях роста числа целевых кибератак на объекты критической информационной инфраструктуры возникает потребность в эффективных подсистемах предупреждения инцидентов безопасности. Основными задачами таких подсистем являются своевременное обнаружение атакующих действий, корректная оценка их значимости и последующая автоматизированная реакция. Анализ функциональных требований представлен в [1], где подтверждается необходимость объединения данных различных средств мониторинга безопасности.

Постановка задачи

Цель исследования – выявить узкие места существующих подсистем предупреждения компьютерных атак в КИИ и предложить механизмы их совершенствования. Задачи включают:

1. Описание требований к подсистемам предупреждения на основе обзора архитектуры интеллектуальных сервисов защиты.

2. Анализ современных методов мониторинга инцидентов и средств IDS, NTA, XDR.

3. Оценка роли корреляции событий безопасности для повышения точности детектирования.

4. Разработка рекомендаций по повышению устойчивости функционирования КИИ при атакующих нагрузках.

5. Использование гибридных методов корреляции событий: статистического, поведенческого и сигнатурного анализов, что снижает уровень ложных срабатываний и повышает полноту обнаружения.

Способ решения

Для достижения поставленных задач предложены следующие подходы:

1. Внедрение многослойной архитектуры интеллектуальных сервисов, обеспечивающей разделение функций сбора, анализа и реагирования.

2. Интеграция XDR-платформы с IDS/IPS и с сетевым трафик-анализом (NTA) для единой видимости инцидентов на хосте и уровне сети.

3. Применение адаптивных алгоритмов машинного обучения для динамической настройки порогов тревог и выявления неизвестных угроз в режиме реального времени.

Результаты показали, что предложенная схема уменьшает ложные срабатывания на 25% по сравнению с базовой конфигурацией IDS+NTA, сохраняя при этом уровень обнаружения не менее 95% сложных атак.

Заключение

Внедрение предложенных механизмов корреляции событий и адаптивных алгоритмов машинного обучения позволяет существенно повысить эффективность подсистем предупреждения компьютерных атак в КИИ [2, 3]. Многослойная архитектура интеллектуальных сервисов обеспечивает гибкость и масштабируемость решения, а интеграция XDR-платформы с традиционными IDS/IPS и NTA-инструментами дает единую картину инцидентов и ускоряет время реакции.

Список литературы

1. Котенко И.В., Петров А.Б., Сидоров В.Г. Подсистема предупреждения компьютерных атак на объекты критической информационной инфраструктуры: анализ функционирования и реализации. Кибербезопасность и критическая инфраструктура, 2023, № 1, с. 13–27. DOI: 10.21681/2311-3456-2023-1-13-27.

2. ISO/IEC 27035:2016 Information security incident management. Международный стандарт, 2016. DOI: 10.3403/30253214.

3. ISO/IEC 27043:2015 Incident investigation principles and processes. Международный стандарт, 2015. DOI: 10.3403/30245543.

УДК 004.056

Д.И. ЕРМАЧКОВ, Д.И. ПРАВИКОВ

РГУ нефти и газа (НИУ) им. И.М. Губкина, Москва

РАЗРАБОТКА ВОСПРОИЗВОДИМОЙ МЕТОДИКИ ОЦЕНКИ ПЗИ ДЛЯ ОБЪЕКТОВ ТЭК

Аннотация. Представлена методика оценки показателя зрелости процессов информационной безопасности (ПЗИ) в организациях ТЭК в соответствии с Приказом ФСТЭК России № 117 от 11 апреля 2025 г. Методика обеспечивает воспроизводимую и количественно измеримую оценку зрелости с учётом отраслевой специфики, включая долю ОТ-контуров, уровень угроз и наличие внешних подрядчиков. Введено пороговое значение γ = 0,6, ограничивающее итоговый ПЗИ при недостижении пороговых уровней по ключевым доменам. Результаты могут использоваться для внутренних аудитов, планирования приоритетов и подготовки отраслевых бенчмарков зрелости.

Развитие нормативных требований в области информационной безопасности в критической инфраструктуре ТЭК требует не только выполнения мер, но и оценки их зрелости. Приказ ФСТЭК № 117 задаёт единые обязательные требования к субъектам КИИ, но не определяет методы измерения эффективности. Разработка показателя ПЗИ позволяет сформировать унифицированную систему оценки и повышения зрелости процессов ИБ.

Предложенная методика основана на десяти доменах зрелости, отражающих ключевые направления управления информационной безопасностью в ТЭК. Особое внимание уделяется доменам 3 (управление уязвимостями), 4 (управление доступом и привилегиями) и 6 (мониторинг и анализ событий ИБ), которые определяют устойчивость и непрерывность функционирования системы ИБ. Для этих доменов введено ограничение: если хотя бы один из них имеет уровень зрелости L < 0,6, итоговый ПЗИ не может превышать γ = 0,6.

Интегральный ПЗИ рассчитывается по формуле:

$$\mathrm{PZI} = \begin{cases} \sum_{i=1}^{n} W_i * L_i * M_i \ , & \text{если } L_3, L_4, L_6 \geq \gamma \\ \min\left(\gamma, \ \sum_{i=1}^{n} W_i * L_i * M_i\right), & \text{если } L_3, L_4, L_6 < \gamma \end{cases}$$

где Wi – вес домена, Li – нормированное значение уровня зрелости, Mi – модификатор контекста.

Показатель отражает баланс между нормативностью и риск-ориентированностью. Весовые коэффициенты Wi распределяются в зависимости от значимости процессов для устойчивости и управляемости ИБ. Модификаторы Mi учитывают особенности предприятия.

Методика предусматривает девять этапов оценки: от подготовки данных и сбора доказательств до проверки ключевых доменов, агрегирования и формирования отчёта для руководства. Результаты представляются в виде интегрального показателя (0–1), градации которого соответствуют уровням «низкая», «допустимая», «высокая» и «оптимальная» зрелость. Сводный расчёт ПЗИ по доменам представлен на рис. 1

Контур	IT
Класс защищённости (1–3)	2
Доля ОТ (0..1)	0,65
Внешние подключения (Да/Нет)	Да
Порог y (ограничение при провале критических доменов)	0,6
Дата оценки	05.10.2025

- ЗАПОЛНЯЕМЫЕ ЯЧЕЙКИ

Таблица расчёта по доменам (введите уровни 0–5, при необходимости KPI фактические значения для gate-доменов)

№	Домен	Описание	Баз. вес IT	Баз. вес OT	α (модификатор)	Вес до норм.	Норм. вес (W)	Уровень (0–5)	Ключевой Домен	Вклад
1	Управление процессами и управлением рисками	Политика ИБ; роли и ответственность; PDCA; ПЗИ/КЗИ; план совершенствования; отчётность ФСТЭК	0,1	0,08	0,95	0,095	0,096839959	4	НЕ ЯВЛЯЕТСЯ	0,077472
2	Учёт активов, конфигураций и изменений	Учёт IT/OT-активов; безопасные конфигурации; управление изменениями; объём и срок «неучтённого»	0,08	0,09	1,05	0,084	0,085626911	3	НЕ ЯВЛЯЕТСЯ	0,051376
3	Управление уязвимостями и обновлениями (ключевой)	Идентификация/устранение уязвимостей; соблюдение сроков; тестирование обновлений	0,14	0,16	1,2	0,168	0,171253823	3	ИСТИНА	0,102752
4	Управление доступом и привилегиями (ключевой)	Минимальные привилегии; МФА; PAM; переаттестации; задержка применения юр. изменений; длительность избыточных прав	0,13	0,15	1,1	0,143	0,145769623	4	ИСТИНА	0,116616
5	Удалённый/беспроводной доступ и мобильные устройства	Защищённый удалённый доступ; контроль Wi-Fi; MDM/BYOD; защищённость каналов	0,06	0,07	1	0,06	0,06116208	3	НЕ ЯВЛЯЕТСЯ	0,036697
6	Мониторинг и анализ событий ИБ (ключевой)	Сбор и корреляция событий; SIEM/SOC; охват IT/OT; отчётность; источники: сети, СКУД-логи	0,13	0,15	1,2	0,156	0,159021407	4	ИСТИНА	0,127217
7	Реагирование на инциденты и восстановление деятельности	RTO/RPO	0,1	0,12	1,1	0,11	0,112130479	4	НЕ ЯВЛЯЕТСЯ	0,089704
8	Безопасная разработка и внедрение ИС	Контроль ИБ в SDLC; SAST/DAST; security-gates; ввод в эксплуатацию	0,06	0,05	1	0,06	0,06116208	3	НЕ ЯВЛЯЕТСЯ	0,036697
9	Работа с подрядчиками и внешними организациями	ИБ-требования в договорах; аудит внешних подключений; контроль сторонних рисков	0,05	0,05	1,1	0,055	0,05606524	4	НЕ ЯВЛЯЕТСЯ	0,044852
10	Отраслевой (дополнительный) домен	Опциональный домен по специфике субъекта КИИ; применяется при наличии трассировки и доказательств	0,05	0,03	1	0,05	0,0509684	3	НЕ ЯВЛЯЕТСЯ	0,030581

Итог ПЗИ (без ограничений):	0,713965341
Ключевой домен провал?	ЛОЖЬ
Порог y:	0,6
ПЗИ (итог с учётом gate):	0,713965341

Класс зрелости:	**Высокая**

Рис. 1. Сводный расчёт ПЗИ по доменам

Таким образом, данная методика создаёт объективную основу для сопоставимой оценки зрелости субъектов КИИ. Её применение в организациях ТЭК позволяет повысить управляемость системы ИБ, снизить риск системных инцидентов и сформировать отраслевые бенчмарки зрелости.

Список литературы

1. ФСТЭК России. Приказ от 11 апреля 2025 г. № 117 «Об утверждении требований по обеспечению безопасности КИИ в ТЭК».

2. Милославская Н.Г., Толстой А.И. Управление информационной безопасностью. М.: НИЯУ МИФИ, 2020. – 536 с.

3. Когос К.Г., Финошин М.А. Перспективные подходы к обнаружению сетевых скрытых каналов. Безопасность информационных технологий, 2021, т. 28, № 2, с. 45–61.

УДК 004.056:615.47

А.П. ДУРАКОВСКИЙ, А.Л. ХЛОПОТНИКОВ

Национальный исследовательский ядерный университет «МИФИ», Москва

АНАЛИЗ УГРОЗ И ФАКТОРОВ, ВЛИЯЮЩИХ НА УСТОЙЧИВОСТЬ РОБОТИЗИРОВАННОЙ МЕДИЦИНСКОЙ СИСТЕМЫ

Работа посвящена анализу киберфизических угроз и разработке мер повышения устойчивости роботизированного медицинского комплекса (РМК) da Vinci Si. В ходе исследования решены три основные задачи: анализ архитектуры системы и идентификация угроз, оценка их критичности и разработка комплекса защитных мер. Выявлены неприемлемые риски, требующие немедленного устранения. Предложены технические и организационные меры противодействия угрозам.

Введение

Актуальность исследования обусловлена прямой зависимостью между кибербезопасностью и физической безопасностью пациента при использовании РМК. Критическим фактором риска признана зависимость от проприетарного программного обеспечения и сервисного канала OnSite, что ограничивает возможности независимого аудита безопасности [1]. В работе решены три основные задачи, направленные на комплексный анализ угроз и разработку мер защиты.

Основные проблемы и предлагаемые решения

Три основные задачи. Во-первых, проведен анализ архитектуры РМК da Vinci Si и идентифицированы ключевые угрозы и уязвимости, среди которых доминируют киберфизические риски, связанные с манипуляцией управляющими данными и нарушением доступности каналов связи и визуализации. Критическим фактором риска признана зависимость от проприетарного программного обеспечения и сервисного канала OnSite, что ограничивает возможности независимого аудита безопасности [1].

Во-вторых, выполнена оценка критичности угроз на основе методологии ГОСТ ISO 14971-2021 [2]. Разработаны шкалы для оценки тяжести последствий (S), ориентированной на ущерб здоровью пациента и нарушение основных функциональных характеристик системы, и вероятности реализации угроз (P). Построенная карта рисков показала, что угрозы манипуляции управляющими данными (R=15) и DoS-атаки на

канал визуализации (R=16) относятся к категории неприемлемых рисков, требующих немедленного устранения.

В-третьих, разработан комплекс рекомендаций по минимизации рисков [3]. К техническим мерам отнесены: внедрение аппаратного контроля целостности и доверенной загрузки для защиты от компрометации прошивки; строгая сегментация и изоляция сетевого контура РМК от общебольничной сети для противодействия сетевым атакам; криптографическая защита внутренних коммуникационных шин. К организационным мерам относятся: внедрение строгой политики аутентификации, регламента обновлений и непрерывного обучения персонала. Сформулированы предложения по совершенствованию нормативно-правовой базы, включая интеграцию требований ФСТЭК России и введение обязанности производителей раскрывать спецификации протоколов для независимого аудита [4].

Заключение

Практическая значимость работы заключается в том, что ее результаты могут быть использованы производителями РМК для устранения уязвимостей, медицинскими учреждениями для повышения безопасности эксплуатации и регуляторами при обновлении стандартов кибербезопасности. Реализация предложенных мер позволит существенно повысить устойчивость роботизированных медицинских систем к современным киберфизическим угрозам.

Список литературы

1. Singh, S., Vyas, A., & Bansal, S. (2021). Securing robotic surgical systems: A review of security vulnerabilities and countermeasures. Medical Devices & Sensors, 4(1), e10102. DOI:10.18280/ijsse.130318.

2. ГОСТ ISO 14971-2021. Изделия медицинские. Применение менеджмента риска к медицинским изделиям.

3. Орысюк, Д.А. Робототехника. Биомедицинские приложения робототехнических систем и их проблемы / Д.А. Орысюк, Т.В. Луинда // Молодые ученые - развитию Национальной технологической инициативы (ПОИСК). – 2022. – № 1. – С. 1047-1048. – EDN OXUBSP. (дата обращения: 20.10.2025).

4. Левоневский, Д.К. Модели сценариев функционирования медицинской киберфизической системы в штатных и экстренных ситуациях / Д. К. Левоневский, А. И. Мотиенко // Программная инженерия. – 2022. – Т. 13, № 8. – С. 383-393. – DOI 10.17587/prin.13.383-393. – EDN JQVKXB.

УДК 003.26: 004.056

С.С. МИНАКОВ
Академия криптографии Российской Федерации, Москва

О ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ ПОСТКВАНТОВЫХ АЛГОРИТМОВ И КВАНТОВОЗАВИСИМЫХ КЛЮЧЕЙ В СХЕМЕ КРИПТОГРАФИЧЕСКОЙ ЗАЩИТЫ ОБЛАЧНЫХ ХРАНИЛИЩ

Работа посвящена отдельным аспектам расширения криптографической схемы серии «Утро-1» постквантовыми алгоритмами шифрования и\или квантовозависимыми ключами, что позволяет строить автоматизированные гибридные криптографические схемы обработки информации на стороне потребителей облачных услуг с использованием стандартизированных симметричных и асимметричных шифров, новых постквантовых алгоритмов, а также квантово-криптографических систем связи и распределения ключей.

Постановка задачи

Стремительное развитие систем информатизации и цифровая трансформация государственного управления привели к переводу на автоматизированную и интеллектуальную обработку большинства процессов управления и обработки информации в госуправлении, в различных организациях, на предприятиях промышленности и транспорта. Существенным аспектом, влияющим на уровень информационной безопасности, является то, что в рамках реализации государственных и национальных программ «Информационное общество», «Цифровая экономика Российской Федерации», «Экономика данных и цифровое государственное управление» Министерством цифрового развития, массовых коммуникаций и связи Российской Федерации предложены и Правительством Российской Федерации одобрены концепции перевода государственных информационных ресурсов в ЦОДы[1] и развёртывания государственной единой облачной платформы (платформа ГЕОП «Гособлако»), с 2019 г. начаты соответствующие эксперименты[2] и к 2021 г. создана единая цифровая платформа Российской Федерации «ГосТех» (платформа ГосТех[3,4]).

[1] Распоряжение Правительства Российской Федерации от 07.10.2015 г. № 1995-р (ред. от 18.10.2018) «Об утверждении Концепции перевода обработки и хранения государственных информационных ресурсов, не содержащих сведения, составляющие государственную тайну, в систему федеральных и региональных центров обработки данных».

[2] Постановление Правительства РФ от 28.08.2019 № 1114 «О проведении эксперимента по переводу информационных систем и информационных ресурсов федеральных органов

В контексте роста ИТ-преступности [1] и при использовании гражданами и организациями сторонних для защищаемых систем сервисов облачных вычислений во многом утрачивается подконтрольность обрабатываемой и хранимой информации, существенно изменяется модель угроз и нарушителя [1, 2], что зачастую не позволяет наделять сторону облачных сервисов требуемыми гарантиями доверия.

При этом рассматривая перевод инфраструктуры обработки и хранения информации в облачные сервисы, необходимо учитывать квантовую угрозу и при использования облачных хранилищ данных третьей стороны переходить от криптографического протокола к криптографической схеме [2, 3], предусмотрев её расширение использованием квантовозависимых ключей из систем квантово-криптографического распределения ключей (ККС ВРК) и новых алгоритмов шифрования постквантового класс.

В данной работе предложено расширить криптографическую схему «Утро-1» [2] использованием ключей из ККС ВРК и(или) постквантовых алгоритмов и взаимодействовать с облачными сервисами по стандартизированным интерфейсам CDMI[5,6] и по прикладными протоколам AWS Simple Storage Service (Amazon Web Services S3 protocol) и Calendaring Extensions to WebDAV (CalDAV protocol) для доступа к ресурсам облачных хранилищ с передачей защищённых (шифрованных) данных в облачные сервисы категории Data Storage as a Service.

Учитывая использование стандартизированных криптографических механизмов и свойств схемы «Утро-1» [2, 5] и публичные результаты её технологической реализации [3, 4], здесь и далее по тексту работы будет предполагаться, что они известны исследователям (читателям) [10]. Будет показано только два подхода к противодействию «квантовой угрозе» в

исполнительной власти и государственных внебюджетных фондов в государственную единую облачную платформу, а также по обеспечению федеральных органов исполнительной власти и государственных внебюджетных фондов автоматизированными рабочими местами и программным обеспечением».

[3] Распоряжение Правительства РФ от 21 октября 2022 г. № 3102-р «Об утверждении Концепции создания и функционирования единой цифровой платформы Российской Федерации "ГосТех" и плана мероприятий ("дорожной карты") по ее созданию».

[4] «Концепция обеспечения информационной безопасности единой цифровой платформы Российской Федерации «ГосТех» (утв. приказом Минцифры России от 12 января 2023 г. № 7).

[5] ГОСТ Р ИСО/МЭК 17826–2015 «Интерфейс управления облачными данными (CDMI)» (Документ утвержден и введен в действие Приказом Федерального агентства по техническому регулированию и метрологии от 29 мая 2015 г. N 467-ст).

[6] ISO/IEC 17826:2022 «Information technology Cloud Data Management Interface (CDMI) Version 2.0.0» (Документ утвержден Техническим комитетом ISO/IEC JTC 1).

рамках задачи о защищенном облачном хранении данных: использование квантовозависимой ключевой информации, полученной с использованием ККС ВРК и использование математических методов «постквантовой» криптографии наряду с классическими стандартизованными блочными шифрами.

Варианты расширения криптографической схемы постквантовыми алгоритмов шифрования и(или) квантовозависимыми ключами

Рассмотрим изменения криптографической схемы, предусматривающие использование квантовозависимых ключей шифрования из квантовых криптографических систем выработки и распределения ключей (ККС ВРК), а также технологические ограничения при использовании постквантовых алгоритмов шифрования.

Предлагаемая модификация криптографической схемы «Утро» состоит в использовании типовой ККС ВРК [9] на этапе выработки и распределения «базовой» ключевой информации: ключа KEK_{VKO} (вариант 1) или мастер-ключей K_E, K_M (вариант 2).

Первый вариант предпочтителен с точки зрения скорости функционирования криптографической схемы в целом, так как выработка криптографически защищенных ключей в ККС осуществляется с невысокой скоростью относительно скорости шифрования. Такой вариант нивелирует и главную угрозу со стороны квантового вычислителя, предотвращая возможное раскрытие ключевой информации (секретных ключей x и y).

Второй вариант подразумевает выработку ключевой информации большего объема, но существенно упрощает криптографическую схему. При выработке ключей K_E, K_M с помощью ККС ВРК отпадает необходимость в шифровании, имитозащите этих ключей, а также в процедуре экспорта материала ключей. Таким образом, шаги 1–3 [2] при размещении и получении данных заменяются одной процедурой:

Шаг 1’–3’. Совместная выработка мастер-ключей K_E, K_M в ККС ВРК.

Выбор конкретных параметров ККС ВРК влияет на скорость выработки КИ, предельно допустимое расстояние ее передачи, криптографическую стойкость ключей, возможность противодействия атакам со стороны потенциального нарушителя. Перечислим основные характерные черты возможной ККС ВРК: реализация одно- или двухпроходной схемы передачи оптических сигналов; использование в оптической схеме одного или двух однофотонных детекторов; выбор сторонами номера оптического сигнала (в двухмодовом состоянии) для наложения фазового сдвига; выбор и, возможно, смена для разных сессий ВРК квантового протокола (BB84, SARG04, АКМ2017 [9], АКМ2021 [11], протоколов на симметричных когерентных состояниях).

Одним из недостатков ККС является ограниченность расстояния передачи КИ в силу наличия оптических потерь [8]. Выработка ключевой информации возможна на расстояниях от нескольких десятков до сотен километров, в зависимости от требований к стойкости вырабатываемых ключей [9]. Эта проблема может быть решена, например, построением сетей ККС, обеспечивающих передачу КИ на (теоретически) неограниченные расстояния, и дающих возможность создания дублирующих линий передачи КИ для защиты от несанкционированного доступа.

Российский опыт разработки систем криптографической защиты информации на основе ККС ВРК [9, 11] для организации защиты объектов критической информационной инфраструктуры позволил реализовать протоколы и интерфейсы взаимодействия КС ВРК и средств криптографической защиты информации, при этом разработанная квантовая криптографическая система обеспечила выработку криптостойкой ключевой информации со средней скоростью не менее 10 бит/с при дальности передачи квантовых сигналов не менее чем на 15 км, что достаточно для размещения в оптоволоконной инфраструктуре ЦОД. При этом взаимодействующие с ККС ВРК средства криптографической защиты информации могут обеспечивать обработку данных на скорости 10 Гбит/с и выше[7].

Рассмотрим изменения криптографической схемы [4], предусматривающие использование постквантовых алгоритмов; точнее – новых асимметричных криптосхем, способных противостоять угрозе использования квантовых алгоритмов и компьютеров [6]. Заметим, что современный уровень развития квантовых вычислений позволяет считать существующие симметричные алгоритмы шифрования и функции хеширования (в том, числе использующие в СНГ: ГОСТ Р 34.10-2018, ГОСТ Р 34.11-2018), защищенными от квантовых компьютеров.

До принятия российского стандарта по постквантовым криптографическим механизмам [7], рассмотрим в данной статье один из трех постквантовых федеральных стандартов обработки информации США (Federal Information Processing Standard, FIPS USA), выпущенных в 2024 году институтом NIST: это стандарт FIPS 203, описывающий алгоритм инкапсуляции ключей (ML-KEM) на основе модульных решеток.

[7] А.А. Карманов. Способ и устройство квантового распределения ключей с контролем параметров квантового канала // Патент на изобретение RU 2840355 C1, 21.05.2025. Заявка № 2024117861 от 27.06.2024.

В алгоритме ML-KEM[8] реализованы следующие процедуры:

– выработка открытого и закрытого ключей для, соответственно, инкапсуляции и декапсуляции общего ключа;

– выработка (инкапсуляция) общего ключа симметричного шифрования;

– восстановление (декапсуляция) общего ключа симметричного шифрования.

Отметим, что важным этапом алгоритма является вычисление шифртекста C (на этапе инкапсуляции) для последующей проверки декапсулированного ключа. В табл. 1 приведены значения длин (в байтах) шифртекста.

Таблица 1. Размеры шифртекста и ключей алгоритма ML-KEM

Вариант алгоритма	Длина ключа инкапсуляции	Длина ключа декапсуляции	Длина шифртекста
ML-KEM-512	800	1632	768
ML-KEM-768	1184	2400	1088
ML-KEM-1024	1568	3168	1568

Для сравнения приведем параметры еще нескольких других схем обмена ключевой информацией, (потенциально) также имеющих определённый уровень стойкости [6, 7], табл. 2.

Таблица 2. Размеры параметров (в байтах) некоторых схем обмена ключами

Схема	Длина секретного ключа	Длина открытого ключа	Длина шифртекста
Three Bears	40	1584	1697
LEDAcrypt	40	18016	9008
FrodoKEM	31272	15632	15768
RQC	3510	3510	3574
SIKE	826	726	766

Отметим, что независимо от выбора варианта алгоритма, длина вырабатываемого ключа фиксирована и составляет не менее 256 бит, что позволяет сделать вывод о потенциальной возможности использования ML-KEM для исследования процедур выработки ключевого материала в используемых российских стандартизированных алгоритмах блочного шифрования («Магма» и «Кузнечик»).

[8] National Institute of Standards and Technology (2024) Module-Lattice-Based KeyEncapsulation Mechanism Standard. (Department of Commerce, Washington, D.C.), Federal Information Processing Standards Publication (FIPS) NIST FIPS 203. DOI: 10.6028/NIST.FIPS.203.

Указанное определяет направление дальнейших исследований, в котором потребуется обеспечить стойкость как в классическом, так и в квантовом смысле (в условиях применения квантового вычислителя).

Список литературы

1. Минаков С.С., Закляков П.В. Информационные технологии и преступления: учеб. пособие – М.: ДМК Пресс, 2023. – 160 с. ISBN 978-5-93700-194-8.

2. Минаков С.С. Основные криптографические механизмы защиты данных, передаваемых в облачные сервисы и сети хранения данных / С. С. Минаков // Вопросы кибербезопасности. – 2020. – № 3(37). – С. 66–75. – DOI 10.21681/2311-3456-2020-03-66-75. – EDN HCRGIJ.

3. Тихов, С.В. Технология защиты облачных хранилищ данных / С.В. Тихов, И.В. Карпов // Системы и средства защиты информации: Сборник статей 15-й межведомственной научно-практической конференции имени Е. А. Матвеева, Пенза, 12–14 сентября 2023 года. – Пенза: Пензенский государственный университет, 2024. – С. 120–126. – ISBN: 978-5-907807-76-1 — EDN AZADHO

4. Вариант построения программного решения с гибридной криптографической системой защиты данных, хранящихся на облачном накопителе, и перспективными режимами работы блочных шифров // С. С. Минаков, И. В. Карпов, С. В. Тихов, И. В. Мартынов // Системы и средства защиты информации : Сборник статей 16-й межведомственной научно-практической конференции имени Е. А. Матвеева, Пенза, 10–13 сентября 2024 года. – Пенза: Изд-во ПГУ, 2025. – 188 с. – С. 94–105. – EDN PFWUSF.

5. Минаков, С. С. О подходах к оценке стойкости криптографических механизмов информационной безопасности для систем и устройств интернета вещей и искусственного интеллекта вещей в условиях ограничений вычислительных возможностей и объёма ключа / С. С. Минаков // Комплексная защита информации : Сборник материалов XXIX научно-практической конференции, Санкт-Петербург, 15–17 мая 2024 года. – Санкт-Петербург: Санкт-Петербургский политехнический университет Петра Великого, 2024. – 292 с. – С. 270–275. – ISBN 978-5-7422-8659-2

6. Панков К. Н., Миронов Ю. Б. Использование постквантовых алгоритмов в задачах защиты информации в телекоммуникационных системах. М.: Горячая линия, Телеком, 2023. 236 с.

7. Гребнев С.В. (2019) Тенденции развития постквантовой криптографии. // Материалы XXI научно-практическая конференция «РусКрипто'2019». URL: https://ruscrypto.ru/accociation/archive/rc2019.html (Дата доступа: 20.09.2025)

8. Wiring surface loss of a superconducting transmon qubit / N.S. Smirnov, E.A. Krivko, A.A. Solovyova [et al.] // Sci. Rep. 2024. V. 14. P. 7326. DOI: 10.1038/s41598-024-57248-y

9. Алиев Ф.К., Корольков А.В., Матвеев Е.А., Орлов С.С., Шеремет И.А. Квантовая криптографическая система АКМ2017 на основе ресурса несепарабельности состояния спиновый синглет //Системы высокой доступности, 2018. Т. 14. № 4. С. 61–72.

10. Минаков С.С., Тихов С.В. (2025) О механизмах криптографической защиты данных публичных облачных хранилищ и перспективах стандартизации технических спецификаций к прикладным протоколам облачных хранилищ данных. // Материалы XXVII научно-практической конференции «РусКрипто'2025», доклад URL – https://ruscrypto.ru /resource/ archive/rc2025/files/09_minakov_tikhov.pdf (Дата доступа: 20.09.2025)

11. Класс квантовых криптографических систем АКМ2021 на основе использования синглетных состояний многокубитовых квантовых систем / Ф.К. Алиев, А.В. Корольков, Е. А. Матвеев // Системы высокой доступности. 2022. № 3(18). С. 5–22.

УДК 004.051

Н.Г. ФЕТЕСКУ

РГУ нефти и газа (НИУ) им. И.М. Губкина, Москва

МЕТОДИКА СРАВНЕНИЯ СРЕДСТВ ЗАЩИТЫ ИНФОРМАЦИИ ПУТЕМ ПРИОРЕТИЗАЦИИ ИХ ВНЕДРЕНИЯ В ОБЪЕКТЫ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ

В работе приведено описание алгоритма, реализующего сравнение СЗИ, используя ранжированные списки ряда распределения вероятностей угроз на объекте КИИ. На основании алгоритма разработана методика сравнения, апробация которой была осуществлена на конкретном объекте КИИ – АСУ ТП нефтеперекачивающей станции. Полученные результаты подтвердили применимость разработанной методики.

В исследованиях [1–3], посвященных сравнению СЗИ, описаны методы, которые включают многокритериальную качественную оценку СЗИ одного типа, оценку возможностей противодействия средства защиты с учетом каналов утечек и количественную оценку показателей качества эффективности функционирования средства защиты.

Применение метода приоритезации внедрения СЗИ в объекты КИИ при сравнении позволяет решить следующие ключевые проблемы: невозможность сравнения средств разной направленности, создание многокритериальной количественной оценки эффективности СЗИ и создание подхода, учитывающего угрозы из нормативной базы ФСТЭК.

Использование данного метода требует построения и предварительного ранжирования списков актуальных угроз в объекте защиты. Путем применения метрик информационной и функциональной безопасности (методика оценки угроз, построение дерева событий) и экспертной оценки (метод анализа иерархий), можно получить ряды распределения вероятностей реализации угроз безопасности информации для объектов воздействия в объекте КИИ, вне зависимости от наличия или отсутствия технологического процесса.

Исходя из данных ранжированных списков угроз, далее выстраивается алгоритм сравнения СЗИ, который использует три критерия при сравнении СЗИ и выборе наиболее эффективного из них:

1) СЗИ способно парировать наиболее вероятную угрозу из списка, в противном случае оно неэффективно.

2) Эффективное СЗИ парирует наибольшее количество угроз.

3) Эффективное СЗИ в случае паритета в количестве парируемых угроз будет парировать наиболее вероятные угрозы.

Алгоритм сравнения СЗИ представлен блок-схемой на рис. 1. Применение данного алгоритма на выходе позволяет получить список наиболее эффективных средств защиты.

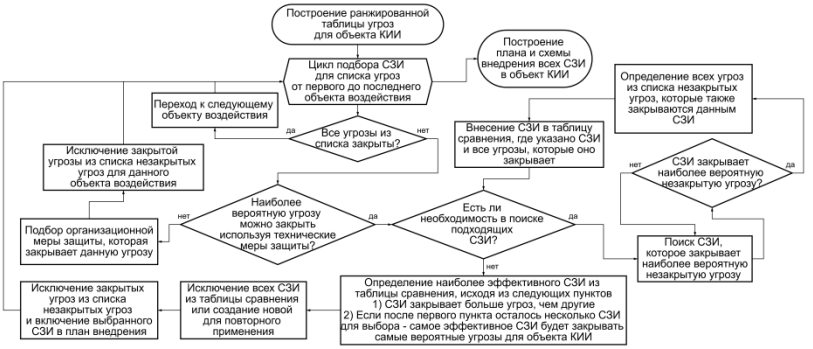

Рис. 1. Блок-схема алгоритма сравнения средств защиты для объекта КИИ

Предложенный алгоритм был рассмотрен для макета АСУ ТП нефтеперекачивающей станции. Его применение позволило выбрать три самых эффективных средства защиты для каждого объекта воздействия (АРМ, SCADA, ПЛК) в АСУ ТП и, в последствии, выстроить схему внедрения СЗИ на объекте КИИ, получив эффективную систему защиты.

Таким образом, разработанная методика позволяет выбирать наиболее эффективные решения, независимо от задач применения рассматриваемых СЗИ, учитывает нормативную базу, имеет гибкую и расширяемую структуру, а также позволяет учитывать дополнительные критерии при сравнении, что делает её полезной для практического применения в реальных условиях.

Список литературы

1. Габова К.А. и др. Сравнение средств защиты информации для включения в учебный процесс //Современное образование: интеграция образования, науки, бизнеса и власти. – 2022. – С. 13–18.

2. Дубровская В.А. Анализ состояния используемых средств защиты информации в отечественных системах защиты и их сравнение //Региональная информатика и информационная безопасность. – 2015. – С. 180–185.

3. Славнов К.В., Темченко А.Н. Особенности применения метода парных сравнений для оценки эффективности средств защиты информации от несанкционированного доступа. //Актуальные проблемы деятельности подразделений УИС. – 2014. – С. 140–146.

УДК 004.056

Д.А. ДЯТЛОВ, Е.А. НЕЩЕРЕТНЯЯ

Национальный исследовательский ядерный университет «МИФИ», Москва

ПРОАКТИВНАЯ БЕЗОПАСНОСТЬ КАК ОСНОВА ТЕХНОЛОГИЧЕСКОГО СУВЕРЕНИТЕТА КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ

Анализируется задача обеспечения технологической независимости объектов критической информационной инфраструктуры. Рассматривается процесс замены импортных компонентов на отечественные. Предлагается подход, при котором служба безопасности выступает в роли интегратора, формирующего сквозные требования и обеспечивающего целостность всей многоуровневой экосистемы – от производства электронной компонентной базы до эксплуатации сложных программно-аппаратных комплексов.

Введение

Курс на технологическую независимость объектов критической информационной инфраструктуры (КИИ) [1] – это не просто переход на отечественную электронную компонентную базу (ЭКБ) и программно-аппаратные комплексы (ПАК), а возможность построить принципиально новую, целостную и доверенную экосистему. Традиционная модель, где безопасность «добавляется» к готовому изделию, упускает суть проблемы [2]. Безопасность – это системное свойство, зависящее от всех ее уровней и их взаимосвязи. В этой новой парадигме специалист по безопасности должен трансформироваться из контролера в интегратора, обеспечивающего сквозную защищенность на всех этапах жизненного цикла КИИ.

От требований к архитектуре безопасности

Рассматривая КИИ «по вертикали» [3]: объект КИИ – система – ПАК – программное обеспечение – ЭКБ, становится очевидной недостаточность защиты только на одном (как правило, канальном) уровне [4]. Безопасность системы определяется прочностью самого слабого звена. Задача интегратора – выстроить эту цепь таким образом, чтобы не было слабых звеньев.

В рамках предлагаемого подхода роль специалиста по безопасности КИИ становится проактивной. Ключевыми функциями с точки зрения интегратора являются:

- формирование требований – разработка единого свода правил,

обязательных для всех участников экосистемы;

- контроль цепочки поставок – аудит производственных и логистических процессов, внедрение практик, исключающих риски подмены компонентов на этапе сборки и доставки;
- проектирование и внедрение единой системы управления всей экосистемы;
- координация взаимодействия между испытательными лабораториями, разработчиками и заказчиками;
- создание типовых, заранее протестированных и сертифицированных схем построения защищенных систем для объектов КИИ.

Заключение

Технологический суверенитет – это, в первую очередь, суверенитет над архитектурой, стандартами и процессами, обеспечивающими безопасность. Служба безопасности объектов КИИ готова выступить драйвером этих изменений, взяв на себя роль интегратора и архитектора доверия. Построив целостную, прозрачную и управляемую экосистему, можно гарантировать, что каждый отечественный чип, каждая плата и каждая программа будут не просто «аналогами», а элементами новой, действительно защищенной основы для КИИ.

Список литературы

1. Агафонов Н.Ю. Процесс импортозамещения в критической информационной инфраструктуре. Актуальные исследования. 2024, № 30 (212), с. 18–21. URL: https://apni.ru/article/9841-process-importozamesheniya-v-kriticheskoj-informacionnoj-infrastrukture. EDN: PDHUCP.

2. Попов А.В., Сплюхин Д.В., Сысоев В.Н., Хлыстов М.А., Цыгунька А.И. Цифровая устойчивость промышленности: почему встроенные системы информационной безопасности нельзя отодвигать на второй план. Энергоэксперт. 2025, № 2 (94), с. 16–20. URL: https://www.elibrary.ru/download/elibrary_82363953_26457192.pdf. EDN: GWKZIC.

3. Кессаринский Л.Н., Никифоров А.Ю. Подход к заданию общих требований к доверенной электронной компонентной базе для регулируемого рынка критической информационной инфраструктуры в вопросах и ответах. Безопасность информационных технологий. 2025, № 1, т. 32, с. 8-16. URL: https://bit.spels.ru/index.php/bit/article/view/1761/1457. EDN: EENLAL.

4. Лапсарь А.П., Назарян С.А., Владимирова А.И. Повышение устойчивости объектов критической информационной инфраструктуры к целевым компьютерным атакам. Вопросы кибербезопасности. 2022, № 2 (48), с. 39-51. DOI: 10.21681/2311-3456-2022-2-39-51.

УДК 004.056

А.И. ЗЫКОВ
Научный руководитель – к.т.н., доцент С.А. РЕЗНИЧЕНКО
Национальный исследовательский ядерный университет «МИФИ», Москва

ПРИМЕНЕНИЕ МЕТОДА КЛАСТЕРНОГО АНАЛИЗА В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ УПРАВЛЕНИЯ ТЕХНОЛОГИЧЕСКИМИ ПРОЦЕССАМИ

Целью работы является повышение безопасности исполнительных устройств автоматизированных систем управления технологическими процессами (АСУ ТП). Метод основан на алгоритмах K-Means и DBSCAN и направлен на выявление аномального поведения исполнительных устройств на основе кластерного анализа характеристик их работы. Тестирование, проведенное на синтетических данных, демонстрирует эффективность предложенного подхода для выявления аномалий. Результаты могут быть использованы для совершенствования систем защиты на объектах критической информационной инфраструктуры (КИИ).

В состав современных АСУ ТП входит большое количество исполнительных устройств (датчики, контроллеры и т.д.), обеспечивающих функционирование технологических процессов. Нарушения их нормальной работы, вызванные как сбоями, так и преднамеренными воздействиями, могут привести к серьезным последствиям для безопасности критической информационной инфраструктуры. Следовательно, актуальной задачей является разработка методов интеллектуального анализа данных, способных автоматически выявлять аномальное поведение исполнительных устройств [1].

Для решения задачи предложено использовать методы кластерного анализа, позволяющие группировать устройства по сходству их поведенческих признаков и, вследствие, наиболее эффективно выбирать методы противодействия [2]. В качестве исходных данных рассматриваются показатели, характеризующие работу устройств: среднее количество команд в минуту, среднее время отклика и энтропия полезной нагрузки передаваемых данных. Энтропия вычисляется на основании моделирования полезной нагрузки по формуле

$$H = -\sum_i p_i \log_2 p_i ,$$

где p_i – вероятность появления i-го символа/байта в передаваемом сообщении.

Для выделения типичных профилей использовался алгоритм K-Means, а для обнаружения отклонений – DBSCAN, эффективно выявляющий редкие и изолированные точки данных, соответствующие аномалиям.

С использованием синтетических данных, моделирующих нормальное и аномальное поведение исполнительных устройств, передающих данные по промышленным протоколам, результаты показали, что сочетания алгоритмов K-Means и DBSCAN позволяет эффективно выделять нормальные группы устройств и автоматически обнаруживать аномальные режимы работы (рис. 1). В частности, устройства с пониженной энтропией полезной нагрузки и увеличенным временем отклика корректно определялись как потенциально опасные.

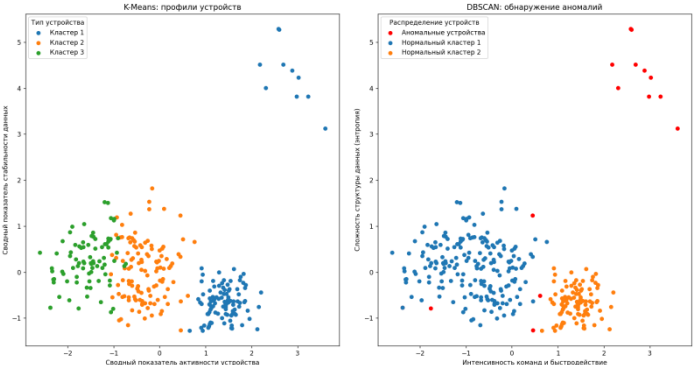

Рис. 2. Выделенные кластеры с идентифицированными аномалиями

Кластерный анализ позволяет автоматически выявлять закономерности в неразмеченных данных, что позволяет оптимизировать процесс ранней идентификации угроз. В отличие от ручного анализа, он обрабатывает большие объемы данных без участия человека и способен обнаруживать новые типы отклонений, уменьшая антропогенный фактор и, тем самым, повышая эффективность защиты. Данный подход может быть интегрирован в систему мониторинга безопасности промышленных сетей для раннего выявления нарушений.

Список литературы

1. Власов С.Л., Рыжкова Е.А. Применение технологий кластеризации для решения задач промышленной автоматики. Применение искусственного интеллекта в промышленном производстве. 2025, с. 74-79.

2. Гилеев И.В., Канавин С.В. Кластеризация угроз информационной безопасности. Общественная безопасность, законность и правопорядок в III тысячелетии. 2023, № 9-2, с. 90-94.

УДК 004.8:004.056

Д.Н. ЛИСОВ

Научно-технологический университет «Сириус», Федеральная территория «Сириус»

ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОБЕСПЕЧЕНИЯ КИБЕРУСТОЙЧИВОСТИ БЛОКЧЕЙН-ЭКОСИСТЕМ

Цель – анализ эффективности методов искусственного интеллекта для обеспечения киберустойчивости блокчейн-систем. Проведено исследование алгоритмов классического и глубокого обучения, а также объяснимого ИИ. Показано, что применение интеллектуальных методов позволяет достигать точности обнаружения аномалий 90–99.8%. При этом для корпоративных блокчейн-экосистем с приватными механизмами консенсуса требуется адаптация интеллектуальных методов к специфике криптостандартов и архитектурных решений.

Растущая сложность киберугроз для блокчейн-систем обусловливает необходимость поиска адаптивных методов защиты, способных выявлять аномалии в реальном времени. Искусственный интеллект демонстрирует высокую эффективность в обнаружении вторжений и классификации киберугроз. Однако его адаптация к корпоративным системам с приватными консенсусами (PoA, RAFT, Hyperledger Fabric и др.) остается недостаточно исследованной областью.

Алгоритмы классического машинного обучения демонстрируют устойчивую эффективность в задачах классификации. Support Vector Machines обеспечивают оптимальное разделение поведения узлов, Random Forest снижает переобучение через ансамблевый подход. Легковесные блокчейн-решения в сочетании с нейронными сетями достигают точности обнаружения атак 99.8% при снижении вычислительной сложности консенсуса [1, 2]. Методы нечетких множеств оценивают риски недостоверных данных с точностью 98% [3].

Методы глубокого обучения качественно расширяют аналитические возможности за счет автоматического извлечения признаков из сложных паттернов. CNN анализируют временные ряды транзакций, LSTM, моделируют долговременные зависимости между событиями, автоэнкодеры выявляют отклонения без предварительной разметки [4].

Однако высокая точность недостаточна для критической инфраструктуры, где требуется прозрачность решений. Объяснимый искусственный интеллект решает эту задачу через интерпретацию

моделей. Методы SHAP определяют вклад признаков, минимизируя ложные срабатывания. ICE визуализирует зависимости между параметрами и предсказаниями. Применение XAI обеспечивает соответствие требованиям регуляторов к верифицируемости решений [4].

Методы искусственного интеллекта демонстрируют эффективность 90–99.8% для киберустойчивости. При этом для корпоративных экосистем с приватными консенсусами применение ИИ находится на начальном этапе. Работы по квантовой устойчивости [5, 6], моделированию поведения в системах безопасности [7, 8] и общей киберустойчивости [9] создают основу, однако требуется адаптация методов к специфике криптостандартов ГОСТ, изучение федеративного обучения на узлах приватных сетей и интеграция XAI с механизмами консенсуса PoA, RAFT и Hyperledger Fabric.

Исследования проведены при финансовой поддержке проекта «Технологии противодействия ранее неизвестным квантовым киберугрозам», реализуемого в рамках государственной программы федеральной территории «Сириус» «Научно-технологическое развитие федеральной территории «Сириус» (Соглашение №23-03 от 27.09.2024).

Список литературы

1. Ghadi Y.Y. et al. A hybrid AI-Blockchain security framework for smart grids // Sci Rep. 2025. Т. 15, № 1. С. 20882.
2. Selvarajan S. et al. An artificial intelligence lightweight blockchain security model for security and privacy in IIoT systems // J Cloud Comp. 2023. Т. 12, № 1. С. 38.
3. Козин И.С. Метод обеспечения безопасности данных при их обработке в блокчейн-системе за счет применения искусственных нейронных сетей: PhD Thesis. СПб.: 2021.
4. Cifci A. Interpretable Prediction of a Decentralized Smart Grid Based on Machine Learning and Explainable Artificial Intelligence // IEEE Access. 2025. Т. 13. С. 36285–36305.
5. Petrenko S., Balyabin A. Модель квантовых угроз безопасности информации для национальных блокчейн-экосистем и платформ // Cybersecurity Issues. 2025. Т. 1, № 65. С. 7–17.
6. Skiba V.Y. и др. Concept of ensuring the resilience of operation of national digital platforms and blockchain ecosystems under the new quantum threat to security // Computing. 2025. Т. 18, № 2. С. 56–73.
7. Гнидко К.О. Метод управления топологией социальной сети с целью предотвращения неконтролируемого эпидемиологического распространения информации между узлами // Наукоемкие технологии в космических исследованиях Земли. Общество с ограниченной ответственностью Издательский дом Медиа паблишер, 2016. Т. 8, № 4. С. 64–69.
8. Гнидко К.О. Modeling of Individual and Group Behavior in P-Adic Coordinate Systems in Order to Solve Information Security Problems // Spiiras Proceedings. 2016. № 1 (44). С. 65–82.
9. Лавренева Е.В. Сравнительная характеристика платформ ethereum, hyperledger fabric и r3 corda. от ethereum к" мастерчейн" // Сборник трудов VII Конгресса молодых ученых. 2018. С. 137–143.

УДК 004.056

Д.Н. ЛИСОВ

Научно-технологический университет «Сириус», Федеральная территория «Сириус»

НАЦИОНАЛЬНЫЕ БЛОКЧЕЙН-ЭКОСИСТЕМЫ: СРАВНИТЕЛЬНЫЙ АНАЛИЗ АРХИТЕКТУР И УГРОЗ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Цель – сравнительный анализ архитектурных решений национальных блокчейн-экосистем РФ и идентификация векторов киберугроз. Выполнен анализ трех платформ различного назначения, что позволило выявить специфику применяемых технических решений. Установлено, что все платформы используют отечественные или проверенные криптопримитивы. При этом прогнозируемое появление квантовых компьютеров к 2027–2031 годам создает стратегическую угрозу, требующую превентивной оценки квантовой устойчивости криптостандартов ГОСТ.

Обеспечение киберустойчивости цифровой инфраструктуры является стратегическим приоритетом. Блокчейн-экосистемы Мастерчейн, Waves Enterprise и РЦИС.РФ формируют ключевые элементы цифровой экономики. При этом различие в архитектурных подходах определяет специфику угроз информационной безопасности, что обусловливает необходимость их детального анализа.

Сравнительный анализ выявляет три принципиально различных архитектурных стратегии. Мастерчейн реализует подход технологического суверенитета через сертифицированное ФСТЭК СКЗИ на криптостандартах ГОСТ 34.10-2012/2018, ГОСТ 34.11, ГОСТ 34.12, обеспечивая высокую производительность до 60 млрд транзакций в год [1, 2]. Waves Enterprise выбирает гибридный подход с приватными сайдчейнами (PoA) и механизмом анкоринга, что объясняет успешное внедрение в авиационной отрасли и банковском секторе [3, 4]. РЦИС.РФ строит общественно-государственную инфраструктуру на Hyperledger Fabric с более 30 узлами государственных органов и IT-компаний [5].

Выявленные архитектурные различия определяют спектр угроз информационной безопасности. Спектр включает классические атаки: 51%, Man-in-the-Middle, DDoS, Eclipse-атаки и уязвимости смарт-контрактов, затрагивающие все типы платформ [6–8]. Однако стратегическую угрозу представляют квантовые технологии. Алгоритм Шора компрометирует ECDSA и криптосхемы на дискретном логарифмировании. Появление криптографически значимых квантовых

компьютеров к 2027–2031 годам актуализирует переход на постквантовую криптографию [9, 10].

Национальные блокчейн-экосистемы демонстрируют различные архитектурные стратегии, обусловленные спецификой применения. При этом все платформы функционируют в условиях растущего ландшафта угроз, где квантовые технологии создают качественно новый вызов. Это требует разработки превентивных стратегий перехода на постквантовую криптографию для обеспечения долгосрочной киберустойчивости национальной цифровой инфраструктуры.

Результаты получены при финансовой поддержке проекта «Технологии противодействия ранее неизвестным квантовым киберугрозам», реализуемого в рамках государственной программы федеральной территории «Сириус» «Научно-технологическое развитие федеральной территории «Сириус» (Соглашение №23-03 от 27.09.2024.).

Список литературы

1. СКЗИ Мастерчейн. Портал документации Мастерчейн [Электронный ресурс]. URL: https://docs.dltru.org/docs/22.3/ (дата обращения: 22.10.2025).
2. Зверева А.С., Половинкина Я.В. Технология распределенных реестров-мастерчейн. Пинск: Полесский государственный университет, 2022.
3. Waves Whitepaper [Электронный ресурс]. URL: https://www.allcryptowhitepapers.com/waves-whitepaper/ (дата обращения: 22.10.2025).
4. Федосенко М.Ю. Особенности и перспективы реализации банковских систем противодействия мошенничеству на технологии блокчейн // Экономика и качество систем связи. Общество с ограниченной ответственностью «НПО» 2024. № 1 (31). С. 107–113.
5. Гаврикова М.М., Каталевский А.А. Цифровые платформы для мониторинга и регистрации интеллектуальных прав // XXIX региональная конференция молодых ученых и исследователей. 2025. С. 53.
6. Petrenko S., Balyabin A. Модель квантовых угроз безопасности информации для национальных блокчейн-экосистем и платформ // Cybersecurity Issues. 2025. Т. 1, № 65. С. 7–17.
7. Гнидко К.О. Метод управления топологией социальной сети с целью предотвращения неконтролируемого эпидемиологического распространения информации между узлами // Наукоемкие технологии в космических исследованиях Земли. ООО Издательский дом Медиа паблишер, 2016. Т. 8, № 4. С. 64–69.
8. Гнидко К.О. Modeling of Individual and Group Behavior in P-Adic Coordinate Systems in Order to Solve Information Security Problems // Spiiras Proceedings. 2016. № 1 (44). С. 65–82.
9. Skiba V.Y. et al. Concept of ensuring the resilience of operation of national digital platforms and blockchain ecosystems under the new quantum threat to security // Computing. 2025. Т. 18, № 2. С. 56–73.
10. Петренко А.С., Петренко С.А., Маковейчук Я.Т. Концептуальная Модель Квантово-Устойчивого Блокчейн. ООО «Издательство Типография «Ариал», 2023. С. 176–179.

УДК 004.056

В.В. БУНИН, А.П. ДУРАКОВСКИЙ

Национальный исследовательский ядерный университет «МИФИ», Москва

РАЗРЕШИТЕЛЬНАЯ СИСТЕМА ДОСТУПА К ФАЙЛОВЫМ СИСТЕМАМ ОС ASTRA LINUX

В работе проведён анализ механизмов разграничения доступа к файловым системам в ОС Astra Linux. Исследована эффективность мандатного контроля доступа (МКД) в сочетании с дискреционной моделью. Рассмотрены проблемы конфигурирования политик Security-Enhanced Linux (SELinux). Разработаны практические рекомендации по повышению защищённости файловых объектов. Предложены новые подходы к автоматизации управления политиками безопасности, расширению аудита и визуализации процессов контроля доступа.

Введение

ОС Astra Linux сертифицирована для применения в государственных и силовых структурах. Её назначение – обеспечение безопасности информации в системах критической информационной инфраструктуры (КИИ) [1]. Это требует использования многоуровневых механизмов регулирования потоков информации, центральное место среди которых занимает разрешительная система доступа к файловым системам. Astra Linux реализует мандатный контроль доступа (МКД) и принудительный контроль на основе политик SELinux [2]. Интеграция дискреционной модели с мандатным контролем на основе меток безопасности позволяет эффективно разграничивать права субъектов [3]. Практическое внедрение этих механизмов сопряжено со сложностями. Настройка МКД и управление политиками SELinux требуют высокой квалификации, ошибки конфигурации могут создавать критические уязвимости [4]. Существующие средства аудита не всегда обеспечивают необходимую детализацию для выявления и расследования инцидентов.

Основные проблемы и предлагаемые решения

Анализ механизмов контроля доступа в Astra Linux выявил комплекс проблем. Основная сложность заключается в конфигурировании политик SELinux, т.к. ручное редактирование конфигурационных файлов требует глубоких специальных знаний и увеличивает вероятность ошибок [4]. Одновременно наблюдается недостаточная гибкость МКД, где жесткая привязка прав к меткам безопасности ограничивает оперативное

перераспределение полномочий [3]. Существенным ограничением являются и возможности системы аудита, фиксирующей лишь базовый набор событий безопасности [2]. Для решения этих проблем предлагается внедрение системы шаблонов политик безопасности для типовых сценариев работы, что упростит развертывание защиты. Одновременно рекомендуется интеграция ролеориентированной модели (RBAC) с существующей системой МКД для реализации механизма временного делегирования полномочий [5]. Совершенствование системы аудита должно включать детальное протоколирование операций с метками безопасности и автоматическое уведомление о подозрительных действиях. Для упрощения работы администраторов предлагается разработка графического интерфейса визуализации политик [4], а для гибридных инфраструктур – модули централизованного управления доступом и инструменты тестирования настроек SELinux [5].

Заключение

Проведённое исследование показало, что Astra Linux обладает развитым аппаратом разграничения доступа на основе интеграции мандатного и дискреционного контроля. Однако механизмы требуют совершенствования в области упрощения управления, повышения гибкости и расширения мониторинга. Предложенные решения – автоматизация управления SELinux, внедрение RBAC, расширение аудита и разработка графических инструментов – направлены на устранение выявленных ограничений. Их реализация позволит снизить нагрузку на администраторов, минимизировать влияние человеческого фактора и повысить устойчивость инфраструктуры к современным угрозам.

Список литературы

1. Иванов Д.В., Гурулев Д.В. Самая безопасная операционная система Astra Linux // Журнал «Трибуна ученого». 2020. Вып. 07. С. 19–27.

2. Девянин П.Н. и др. Интеграция мандатного и ролевого управления доступом в верифицированной модели безопасности ОС // Труды ИСП РАН. 2020. Т. 32. Вып. 1. С. 7–26.

3. Евдокимов О.Г., Гавдан Г.П., Резниченко С.А. Подход к оценке эффективности системы обеспечения ИБ распределённой системы передачи данных // БИТ. 2022. Т. 29. Вып. 2. С. 57–70.

4. Задорожный П.В. Систематизация и анализ средств защиты информации на базе ОС Astra Linux // Наука и просвещение. 2022. С. 54–58.

5. Ильина О.Б., Купчиненко О.П., Скоропад А.В. Изменения в системе защиты информации в ОС Astra Linux SE // ИБРР-2021. С. 152–154.

УДК 004.056

А.В. СУХАРЕВ
Научный руководитель – А.И. СОЛОПОВ
Национальный исследовательский ядерный университет «МИФИ», Москва

УПРАВЛЕНИЕ БЕСПИЛОТНЫМ ТРАНСПОРТНЫМ СРЕДСТВОМ С УЧЕТОМ ВОЗМОЖНОСТЕЙ И УЯЗВИМОСТЕЙ СОВРЕМЕННЫХ КАНАЛОВ СВЯЗИ

Интегрированное применение различных информационно-коммуникационных технологий, интернета вещей, больших данных и других инновационных решений является актуальным в рамках реализации «умного города» для улучшения качества жизни. Одной из таких технологий является беспилотные транспортные средства (БТС). БТС является объектом критической информационной инфраструктуры (Федеральный закон от 26.07.2017 № 187-ФЗ). В условиях четвертой промышленной революции развитие БТС в России, наряду с другими странами мира неизбежно.

Основная часть

БТС обладают высокой степенью интеграции в систему дорожного движения, для его управления в этой системе используются различные каналы связи:

– Четвертое и пятое поколение мобильной связи (технологии 4G и 5G), имеющие высокую скорость и низкую задержку, что критично для управления в реальном времени.

– Спутниковые каналы связи, имеющие высокую задержку и подверженность помехам, но глобальное покрытие.

– Беспроводные локальные сети на основе стандартов IEEE 802.11 (Wi-Fi), при плотной городской застройке.

– Технологию беспроводной связи V2X (Vehicle-to-Everything) позволяющую обмениваться информацией с другими транспортными средствами, дорожной инфраструктурой, пешеходами и являющейся ключевым элементом обеспечения безопасности развития «умных» дорог и БТС.

Вместе с тем, каналы связи БТС уязвимы к классическим угрозам: глушение сигналов (создание радиопомех для нарушения связи путем создания «белого шума» на частоте целевого сигнала), спуфинг ГНСС (глушение приемника с целью подачи ложного сигнала, атаки «человек

посередине» (перехват и подмена управляющих команд), DDoS-атаки (создание перегрузки каналов и задержки).

Утрата канала связи может вызвать серьёзные последствия: аварии на дорогах, отклонение от маршрутов, создание угрозы жизни/здоровью пассажиров, срывы логистических цепочек и экономическому ущербу, террористические акты.

В связи с этим, современные каналы связи необходимо защищать. Для решения этой задачи и защиты современных каналов связи предлагается использовать следующие методы защиты:

1) Комбинирование нескольких независимых мер (шифрование, аутентификация, резервирование каналов, детектирование аномалий).

2) Частотная перестройка, расширение спектра, спектральный мониторинг.

3) Комплексное применение ГНСС и лидаров с инерционными системами.

4) Криптография и аутентификация. Шифрование каналов: TLS (для IP-трафика), DTLS (для UDP), IPsec для защиты конфиденциальности и целостности с двухсторонней аутентификаций.

5) Мониторинг атак и телеметрии.

Заключение

Развитие БТС обусловливает совершенствование существующих каналов связи и применение передовых комплексных решений.

Список литературы

1. Марков Н.Г., Сонькин Д.М., Фадеев А.С., Шемяков А.О., Газизов Т.Т., «Интеллектуальные навигационно-телекоммуникационные системы управления подвижными объектами с применением технологии облачных вычислений», М.: Горячая линия-Телеком 2014, ISBN 978-5-99120355-5;

2. Жарко Е.Ф., Промыслов В.Г., Исхаков А.Ю., Мещеряков Р.В., Семенков К.В., Абдулова Е.А., Байбулатов А.А., Исхаков С.Ю. «Кибербезопасность беспилотных транспортных средств. Архитектура. Методы проектирования», М.: Радиотехника 2021, ISBN 978-5-93108-205-9;

3. Шемонаев А.Н., Епифанцев К.А., Кессаринский Л.Н., Пончаков М.Ю., «О результатах экспериментального исследования нарушения функционирования компонентов беспилотных транспортных средств от преднамеренных деструктивных электромагнитных воздействий, безопасность информационных технологий» IT Security, том 32, № 1 (2025), DOI: 10.26583/bit.2025.1.12;

УДК 004.056

К.Г. АЙВАЗОВСКИЙ
Научный руководитель – к.т.н, доцент В.С. ГОРБАТОВ
Национальный исследовательский ядерный университет «МИФИ», Москва

ЗАЩИТА ПЕРСОНАЛЬНЫХ ДАННЫХ В ЦИФРОВОЙ ИНФРАСТРУКТУРЕ

В работе рассмотрены современные подходы к обеспечению защиты персональных данных (ПДн) в корпоративных информационных системах. Проанализированы ключевые угрозы безопасности, особенности нормативного регулирования и методы реализации многоуровневой политики доступа на основе отечественных программных решений. Особое внимание уделено техническим механизмам защиты, реализованным в ОС Astra Linux и «Ред ОС». Предложены практические меры по автоматизации управления безопасностью, расширению системы аудита и внедрению доверенной среды обработки данных в корпоративных инфраструктурах.

Введение

В условиях цифровизации бизнеса персональные данные сотрудников, клиентов и партнёров становятся стратегическим активом компании. Их защита в распределённых корпоративных системах и облачных платформах является приоритетной задачей информационной безопасности. Законодательная база, включая Федеральный закон № 152-ФЗ «О персональных данных» [1], а также требования ФСТЭК и ФСБ России, определяют необходимость применения сертифицированных средств защиты информации. Приоритет в корпоративном секторе получают отечественные операционные системы, такие как Astra Linux и «Ред ОС», обеспечивающие реализацию мандатного и дискреционного контроля доступа, криптографической защиты и аудита операций [2].

Основные проблемы и предлагаемые решения

Анализ мер защиты персональных данных показывает, что в корпоративных информационных системах отсутствует единая согласованная архитектура их защиты. Разобщённость политик безопасности и использование изолированных механизмов контроля доступа приводят к несогласованности прав, затрудняют управление и повышают риск утечек. Традиционные дискреционные модели не обеспечивают необходимой гибкости при изменении организационной

структуры, а ручное администрирование без визуальных инструментов порождает ошибки конфигурации. Ситуацию усугубляет слабая детализация аудита и мониторинга действий пользователей, что снижает прозрачность процессов и эффективность реагирования на инциденты [3].

Для устранения указанных проблем необходим системный и комплексный подход к защите ПДн. Оптимальным является построение интегрированной модели управления доступом, объединяющей мандатный, дискреционный и ролевой принципы в единую структуру [4]. Ключевым элементом защиты должны стать технические механизмы отечественных ОС – Astra Linux и «Ред ОС», обеспечивающие мандатный контроль, криптографическую защиту, изоляцию процессов и доверенный контур безопасности. Повышение эффективности возможно через автоматизацию аудита и внедрение SIEM-систем, позволяющих выявлять аномалии в обращениях к ПДн. Завершающим направлением является развитие визуальных инструментов администрирования, которые обеспечивают наглядное управление политиками безопасности и минимизируют влияние человеческого фактора [5].

Заключение

Комплексная защита персональных данных в корпоративной среде требует сочетания нормативных, организационных и технологических мер. Реализация технических механизмов защиты, встроенных в отечественные ОС Astra Linux и «Ред ОС», обеспечивает высокий уровень доверия и соответствие требованиям российского законодательства.

Интеграция этих решений с централизованными системами управления доступом и аудита позволит компаниям построить единый контур безопасности, повысить прозрачность обработки данных и минимизировать риски несанкционированного доступа.

Список литературы

1. Федеральный закон № 152-ФЗ «О персональных данных» от 27.07.2006 г.
2. Акимова Г.П., Даниленко А.Ю., Пашкина Е.В. Мандатный контроль в автоматизированных информационных системах // Прикладные аспекты информатики. 2020. Выпуск 3, С. 3–12.
3. Унижаев Н.В. Особенности моделирования угроз безопасности персональных данных для обеспечения достаточного уровня защищённости // Информационные технологии и вычислительные системы. 2022. №1. С. 95-109.
4. ГОСТ Р 59453.1-2021. Защита информации. Формальная модель управления доступом.

УДК 004.056

А.С. ПОТАПОВА

Национальный исследовательский ядерный университет «МИФИ», Москва

МЕТОДИКА РАННЕГО ОБНАРУЖЕНИЯ DDOS-АТАК НА ОСНОВЕ МОНИТОРИНГА АРТЕФАКТОВ ИНФРАСТРУКТУРЫ ЗЛОУМЫШЛЕННИКОВ

В статье рассматривается инновационный подход к обнаружению DDoS-атак на этапе подготовки злоумышленниками. Предлагаемое решение основано на автоматизированном мониторинге открытых источников с целью выявления конфигурационных файлов, используемых для развертывания инструментов атаки. В отличие от традиционных реактивных методов, методика обеспечивает упреждающее реагирование до начала непосредственного воздействия на целевые ресурсы. Приводится математическая модель для оценки результативности решения, демонстрирующая превосходство над классическими подходами.

DDoS-атаки остаются одной из наиболее распространенных и разрушительных угроз для российской ИТ-инфраструктуры. Исследователи [1] отмечают, что реактивные методы оповещают непосредственно после начала атаки, имеют высокий риск ложных срабатываний, низкую эффективность против сложных и целевых атак. Существующие проактивные способы, работающие на предсказании атаки на основе косвенных признаков, часто носят характер намерения, а не непосредственной подготовки, а также большое время упреждения делает уведомления непригодными для оперативного принятия решений.

Актуальность исследования обусловлена необходимостью сместить внимание на раннее предотвращение инцидента. Цель: разработать методику проактивного обнаружения DDoS-атак с высоко достоверным оповещением и временем, достаточным для принятия мер по нейтрализации угрозы до ее реализации.

Методика основана на гипотезе о том, что подготовка к DDoS-атаке сопровождается появлением в открытом доступе специфических цифровых артефактов, к которым относятся конфигурационные файлы, содержащие целевые IP-адреса и домены, а также списки прокси-серверов. Мониторинг этих данных позволяет зафиксировать факт подготовки атаки в момент ее организации, а не исполнения. Система реализована в виде набора взаимодействующих скриптов, развернутых на изолированной системе. Это обеспечивает скрытность и независимость от нагрузки на основные каналы связи.

На рис. 1 представлена схема методики. Цикл мониторинга повторяется каждые 2 минуты, что обеспечивает высокую оперативность.

Рис. 1. Схема этапов методики обнаружения DDoS-атак

Для количественной оценки предлагаемого решения введем метрику Результативности предотвращения (PE):

$$\mathrm{PE} = \frac{(\mathrm{T\_prepare\_attacker} - \mathrm{T\_response})}{\mathrm{T\_prepare\_attacker}},$$

где T_prepare_attacker – расчетное время подготовки атаки злоумышленником; T_response – время, необходимое защищающейся стороне на принятие мер после получения оповещения.

Для реактивных систем T_prepare_attacker = 0, так как атака уже началась, T_response = 2 мин. $\mathrm{PE} = \frac{(0-2)}{0} \to -\infty$ (неэффективность). Формально PE = 0, так как предотвратить начало атаки невозможно.

Для предлагаемой системы среднее значение T_prepare_attacker = 4,5 мин. Временной интервал в 3–6 минут, доступный для принятия решений, рассчитывается на основе анализа Kill Chain DDoS-атаки, исходя из затрат на развертывание инфраструктуры, запуск атаки и выход на полную мощность, и подтвержден практически $\mathrm{PE} = \frac{(4{,}5-2)}{4{,}5} = 0{,}5$. Методика позволяет нейтрализовать от 50% потенциального ущерба от атаки, когда реактивные системы этот ресурс полностью утрачивают.

В результате создана проактивная методика обнаружения DDoS-атак, принципиальным отличием которой является использование прямых артефактов подготовки атаки в качестве высоко достоверного источника. Система обеспечивает надежное временное окно для принятия упреждающих мер. В дальнейшем планируется расширение источников мониторинга, а также интеграция с системами реагирования.

Список литературы

1. Руднев, Д.В. Исследование методов обнаружения DDOS атак на объект КИИ. VI Международная научная конференция по междисциплинарным исследованиям. Сборник статей конференции, Екатеринбург, 15 июня 2024 г. – Екатеринбург: ООО «Институт цифровой экономики и права», 2024. – С. 506-510.

УДК 004.056

С.В. БОБЫРЕВ
МИРЭА – Российский технологический университет, Москва

СПОСОБЫ ПРОТИВОДЕЙСТВИЯ DDOS-АТАКАМ НА API ИНФОРМАЦИОННЫХ СИСТЕМ

В статье рассматриваются современные методы и технологии защиты API (интерфейсов прикладного программирования) информационных систем от распределённых атак отказа в обслуживании (DDoS). Обосновывается необходимость комплексного подхода, включающего фильтрацию трафика, аутентификацию, использование антибот-систем и масштабирование ресурсов. Представлены рекомендации по реализации эффективной защиты, обеспечивающей непрерывность работы сервисов и сохранность данных.

Введение

Современные информационные системы всё активнее используют API для интеграции и взаимодействия с другими сервисами, автоматизации бизнес-процессов и предоставления удалённого доступа к данным [1]. Однако открытость API создаёт уязвимость перед DDoS-атаками – предприятию грозит потеря доступности сервисов и репутационные риски. В связи с этим защита API от DDoS-атак становится ключевой задачей кибербезопасности.

Методы противодействия DDoS на API

Эффективная защита основана на многоуровневом подходе, включающем [1, 2]:

- Фильтрация и анализ трафика. Использование комплексных систем, включая Web Application Firewall (WAF) и специализированные Anti-DDoS решения, позволяет выявлять и блокировать вредоносный трафик. Особенное внимание уделяется анализу поведения клиентов для определения аномалий, характерных для DDoS. К таким методам относятся ограничение числа запросов (rate limiting), анализ географической и сетевой принадлежности, а также проверка заголовков и параметров запросов.
- Аутентификация и авторизация. Применение многофакторной аутентификации и механизмов контроля доступа ограничивает круг лиц и приложений, имеющих право обращаться к API. Использование OAuth, JWT токенов и сертификатов предотвращает несанкционированное использование интерфейсов и затрудняет проведение массовых атак.

- Антибот-системы и поведенческий анализ. Современные антибот-решения основаны на машинном обучении и анализе паттернов поведения трафика. Они способны выявлять автоматизированные атаки, замаскированные под легитимный трафик, и блокировать их ещё на уровне сетевого периметра.
- Масштабирование и инфраструктурная устойчивость. Наличие достаточной пропускной способности, использование CDN и балансировщиков нагрузки позволяет распределять входящий трафик и обеспечивать бесперебойную работу на время возможных атак.
- Шифрование и минимизация зоны атаки. Перевод трафика на защищённый протокол HTTPS, ограничение доступа к API с помощью белых списков IP и виртуальных частных сетей (VPN) снижает возможности злоумышленников.

Для бизнеса рекомендуется внедрять комплексные решения по защите API, интегрирующие несколько методов одновременно. Следует регулярно проводить тестирование уязвимостей и анализировать логи для своевременного обнаружения новых видов атак. Использование облачных и гибридных Anti-DDoS сервисов обеспечивает адаптивность и масштабируемость защиты.

Заключение

Защита API от DDoS-атак – это обязательный элемент современной кибербезопасности. Только комплексный многоуровневый подход с применением современных технологий и постоянным мониторингом позволяет эффективно противостоять угрозам, обеспечивая стабильность работы информационных систем и безопасность данных.

Список литературы

1. Иванченко Р., Уваров А. Защита API от бот-атак и эксплуатации уязвимостей / Роман Иванченко, Арсений Уваров // ITSec.ru. – 29 апреля 2025. – URL: https://www.itsec.ru/articles/zashchita-api-ot-bot-atak-i-ekspluatacii-uyazvimostej (дата обращения: 15.10.2025).

2. Как защитить API от DDoS-атак // StormWall.pro. – 20 апреля 2025. – URL: https://stormwall.pro/resources/blog/kak-zashchitit-api-ot-ddos-atak (дата обращения: 15.10.2025).

УДК 004.457

З.С. КУТЬИН

Национальный исследовательский ядерный университет «МИФИ», Москва

ПРОГРАММА ДЛЯ АНАЛИЗА СВЕДЕНИЙ АВТОРИЗАЦИИ LINUX

Рассматривается проблема обеспечения информационной безопасности (ИБ) Linux-систем. В качестве решения предлагается программа для анализа сведений авторизации. Инструмент, созданный на Python, предоставляет удобный графический интерфейс для централизованного просмотра и анализа данных из лог-файлов secure/auth, wtmp и btmp. Ключевые функции включают фильтрацию по дате, поиск и сортировку, что значительно упрощает для администратора ИБ задачу выявления подозрительных попыток входа и несанкционированного доступа. Программа представляет собой готовый инструментарий для усиления защиты Linux-инфраструктур.

Введение

Подавляющая часть информационных систем, которые являются в том числе объектами критической информационной инфраструктуры, включая серверные платформы, облачные вычисления и автоматизированные рабочие места рядовых сотрудников, построены на базе ОС Linux [1]. С ростом актуальности этой ОС растет и количество угроз ИБ. Одним из основных способов защиты является анализ сведений об авторизации пользователей. Функциональный инструмент, обрабатывающий эти данные, позволит администратору ИБ быстрее выявлять аномалии, предотвращать утечки информации и своевременно реагировать на потенциальные угрозы.

Разработка программы анализа сведений авторизации Linux

При разработке программы использовалась парадигма объектно-ориентированного программирования, которая удовлетворяет требованиям современной разработки, а также позволяет реализовать функционал для анализа данных: подсвечивание выбранных строк, поиск по ключевым словам, сортировка [2]. В программе реализована авторизация пользователей с повышенными привилегиями (root) для доступа к сведениям авторизации без ограничений. Интерфейс программы обеспечивает вывод информации из различных лог-файлов: secure/auth (авторизация пользователей, включая удачные и неудачные попытки входа в систему, а также задействованные механизмы аутентификации.), wtmp (записи о всех сессиях входа в систему), btmp (записи о неуспешных попытках входа). Кроме того, в программе

реализован вывод всех событий авторизации по указанной дате. Интерфейс программы показан на рис. 1.

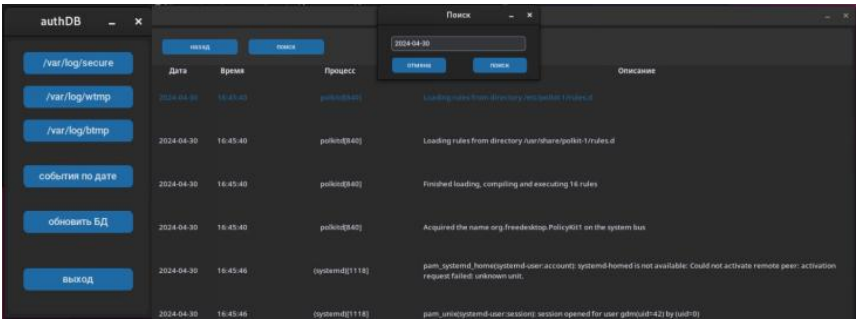

Рис. 1. Интерфейс программы для анализа сведений авторизации Linux

Логика и интерфейс программы разработаны с помощью языка программирования Python и совместимых с ним библиотек. Этот язык является одним из лучших вариантов для разработки качественных и полезных инструментов, в том числе в сфере ИБ [3].

Заключение

Разработанная программа сможет повысить эффективность работы администратора ИБ благодаря широкому функционалу для анализа сведений авторизации и применения современных технологий, а значит снизить процент реализации угроз ИБ в организации. При этом использование актуального языка программирования позволит расширять функционал программы под нужды развивающегося мира технологий.

Список литературы

1. Зайнабидинов Рахматулло Мадаминович Обзор ядра Linux и его роль в современных информационных системах // Universum: технические науки. 2024. №3 (120). URL: https://cyberleninka.ru/article/n/obzor-yadra-linux-i-ego-rol-v-sovremennyh-informatsionnyh-sistemah (дата обращения: 23.10.2025).

2. Гулджанова Д., Мырадов Р., Довранов С. История развития объектно-ориентированного программирования // Вестник науки. 2024. №5 (74). URL: https://cyberleninka.ru/article/n/istoriya-razvitiya-obektno-orientirovannogo-programmirovaniya (дата обращения: 23.10.2025).

3. Дронов В.Ю., Дронова Г.А. Python как средство автоматизации в информационной безопасности. Контроль актуальности защиты от вредоносного кода // ОмГТУ. 2021. № 4. URL: https://cyberleninka.ru/article/n/python-kak-sredstvo-avtomatizatsii-v-informatsionnoy-bezopasnosti-kontrol-aktualnosti-zaschity-ot-vredonosnogo-koda (дата обращения: 23.10.2025).

УДК 004.056

Е.А. ЖАРКОВ
Научный руководитель – доцент Г.П. ГАВДАН
Национальный исследовательский ядерный университет «МИФИ», Москва

ПРОБЛЕМНЫЕ ВОПРОСЫ НОРМАТИВНОГО РЕГУЛИРОВАНИЯ БЕЗОПАСНОСТИ ОБЪЕКТОВ КИИ В СФЕРЕ ЭНЕРГЕТИКИ

В статье анализируются системные противоречия в нормативно-правовом регулировании безопасности объектов критической информационной инфраструктуры (КИИ) в энергетике. Выявлена ключевая проблема: юридический статус значимого объекта КИИ (ЗОКИИ) определяется не его реальными свойствами, а формальным включением в реестр ФСТЭК России, что создает правовые коллизии и ослабляет режим защиты на этапе идентификации.

Введение

Глобальная цифровая трансформация, превратила информационно-коммуникационные технологии (ИКТ) из вспомогательного инструмента в критическую основу функционирования государства и общества. Особое значение это приобретает в энергетике, где отказоустойчивость объектов генерации и распределения энергии [1] напрямую зависит от сложных автоматизированных систем управления (АСУ) и информационно-телекоммуникационных систем (ИТКС), формирующих цифровой каркас отрасли. Однако стремительная киберфизация создает качественно новые риски [1]. Поэтому высокая сложность и взаимозависимость компонентов энергетической инфраструктуры означают, что компрометация даже одного элемента может вызвать сбой с серьезными последствиями для государства. По данным национальных и международных CERT-центров, энергетический сектор продолжает оставаться одним из основных мишеней для целенаправленных кибератак. Следовательно, энергетика, как критически важная отрасль продолжает оставаться одной из наиболее уязвимых сфер для кибератак [2]. Динамичное развитие угроз требует адекватного и непротиворечивого нормативного регулирования. Однако действующая правовая база Российской Федерации (РФ) в области КИИ содержит ряд внутренних противоречий, снижающих ее эффективность.

Ключевые противоречия в государственном правовом регулировании критической информационной инфраструктуры

Анализ Федерального закона № 187-ФЗ и подзаконных актов выявил следующие проблемные аспекты:

1. Формальный подход к определению значимости.

Юридический статус объекта КИИ как «значимого» возникает только после его включения в реестр ФСТЭК России. До этого момента, даже если объект фактически обеспечивает критический процесс в энергетике, требования по его защите (например, по приказам ФСТЭК №№ 235, 239) формально не действуют. Это создает «серую зону», в которой реально значимые объекты не защищены в установленном порядке [3].

2. Коллизия административной и уголовной ответственности.

Субъекта КИИ нельзя привлечь к административной ответственности за невыполнение требований по защите ЗОКИИ, пока объект не внесен в реестр. В то же время, в соответствии со ст. 274.1 УК РФ, возможно привлечение к уголовной ответственности [3] за нарушение правил эксплуатации, повлекшее вред КИИ, даже для объекта, не имеющего формального статуса ЗОКИИ. Данная коллизия создает правовую неопределенность для организаций топливно-энергетического комплекса.

Заключение

Проведенный анализ подтверждает, что актуальной проблемой обеспечения безопасности ЗОКИИ в энергетике является не только техническая защита, но и системные недостатки нормативного регулирования. Ключевой вывод: существующий формально-реестровый принцип присвоения категории значимости вступает в противоречие с необходимостью обеспечения непрерывной и заблаговременной защиты критически важных объектов (КВО). Для повышения устойчивости КИИ энергетики требуется устранение выявленных правовых коллизий и переход к более гибкой модели регулирования, основанной на объективных критериях значимости инфраструктуры.

Список литературы

1. Наталичев, Роман В. et al. Эволюция и парадоксы нормативной базы обеспечения безопасности объектов критической информационной инфраструктуры. Безопасность информационных технологий, [S.l.], v. 28, n. 3, p. 6–27, сен. 2021. ISSN 2074-7136. Доступно на: <https://bit.spels.ru/index.php/bit/article/view/1359/1234>. Дата доступа: 24 окт. 2025. doi: http://dx.doi.org/10.26583/bit.2021.3.01.

2. Голунов, С.В. Энергетическая безопасность постсоветских де-факто государств: вызовы, конфликты и взаимозависимость / С.В. Голунов // Мировая экономика и международные отношения. – 2023. – Т. 67, № 12. – С. 116-126. – DOI 10.20542/0131-2227-2023-67-12-116–126. – EDN SXTIHO.

3. О нормативном регулировании в области использования атомной энергии / Г.А. Ершов, В.Н. Семериков, В.В. Соколов, П.Б. Стрелков // Стандарты и качество. – 2024. – № 6. – С. 40–44. – DOI 10.35400/0038-9692-2024-6-30-24. – EDN DWDBXI.

УДК 004.056

И.А. БАТИЩЕВ, А.А. ЗУЕНКО

МИРЭА – Российский технологический университет, Москва

ОЦЕНКА ПРОТИВОРЕЧИЙ В НОРМАТИВНЫХ АКТАХ НА ОСНОВЕ ПОИСКА ЗАИМСТВОВАНИЙ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

Целью исследования является нахождение противоречий в нормативных актах. Предложен подход по выявлению данных проблем с применением искусственного интеллекта. Для этого используется комбинация строкового сравнения, семантического анализа, моделей NLI и анализа ассоциативных правил, что позволяет автоматизировать аудит документов и наглядно выявлять проблемные зоны.

Введение

В современных организациях со временем накапливается большой объем внутренней нормативной документации. Разработкой таких документов зачастую занимаются разные подразделения.

Постоянная необходимость актуализации закономерно приводят к появлению противоречий – несогласованных, дублирующих или конфликтующих положений. Из-за отсутствия системы актуализации, документы старого и нового образца могут существовать одновременно.

В реалиях крупных организаций ручной аудит слишком трудоёмок. Подобные противоречия создают серьезные риски для информационной безопасности, приводя к правовой неопределенности, ошибкам в настройке систем контроля доступа и созданию уязвимостей [1].

Оценка противоречий в нормативных актах на основе поиска заимствований

Задача заключается в создании автоматизированного метода обнаружения логических конфликтов между положениями разными документа, определения семантических расхождений в трактовках одинаковых понятий, выявления наслоений версий и устаревших формулировок, нахождения дублированных норм с различными интерпретациями [2].

Комплексный метод на основе машинного обучения способен выявлять и классифицировать разногласия во внутренней документации. Метод представлен в трёх этапах. На первом этапе применяются строковые

методы сравнения для эффективного поиска дословных совпадений [3]. Для обнаружения одинаковых перефразированных терминов используется семантический анализ, который переводит текстовые фрагменты в векторные представления и вычисляет семантическое сходство между ними [4].

Вторым этапом является применение моделей Natural Language Inference (NLI), которые классифицируют пары предложений на три отношения: противоречие, нейтральность и корректность [2]. Это позволяет автоматически находить конфликтующие утверждения. Третий этап заключается в выявлении скрытых структурных проблем. Для этого применяются методы ассоциативных правил.

Анализируется частота совместного упоминания терминов и ссылок между документами. При помощи этого строится граф связей и визуализируются группы документов, наиболее подверженные риску противоречий [1].

Заключение

Противоречия во внутренней документации представляют собой очень недооцененную угрозу информационной безопасности. Благодаря современным методам машинного обучения и обработки языка можно автоматизировать выявление подобных противоречий.

Предложенные методы обнаружения разногласий в документации способны систематизировать процесс обнаружения и классификации потенциально конфликтующих положений. Однако данный вопрос требует дополнительных исследований.

Список литературы

1. Раткин Л.С. Разработка методологии выявления противоречий в нормативно-правовых документах // Юриспруденция и государство. 2020. – с. 93-95.

2. Li, Jierui; Raheja, Vipul; Kumar, Dhruv. ContraDoc: Understanding Self-Contradictions in Documents with Large Language Models // arXiv preprint arXiv:2311.09182. – 2023. – URL: https://arxiv.org/abs/2311.09182.

3. Arabi, H. Improving plagiarism detection in text document using structural and semantic hybrid similarity // Neurocomputing. – 2022. – URL: https://www.sciencedirect.com/science/article/abs/pii/S0957417422012489.

4. Álvarez-Carmona, M.A., Franco-Salvador, M., Villatoro-Tello, E., Montes-y-Gómez, M., Rosso, P., Villaseñor-Pineda, L. Semantically-informed distance and similarity measures for paraphrase plagiarism identification // arXiv preprint arXiv:1805.11611. – 2018. – URL: https://arxiv.org/abs/1805.11611.

УДК 004.056

М.М. БАРИНОВА[1,2]
Научный руководитель – доцент Г.П. ГАВДАН[1]
[1]*Национальный исследовательский ядерный университет «МИФИ», Москва*
[2]*Федеральная таможенная служба, Москва*

ЗАЩИТА ИНФОРМАЦИИ В ГИС С ИСПОЛЬЗОВАНИЕМ СЕРВИСОВ ИИ: НОРМАТИВНО-ТЕХНИЧЕСКИЙ АНАЛИЗ И ПЕРСПЕКТИВЫ

В работе проведен анализ проблем и методов защиты информации в государственных информационных системах (ГИС), интегрирующих сервисы искусственного интеллекта (ИИ). Особое внимание уделено противоречию между стратегической важностью ИИ для цифровой трансформации и рисками, связанными с использованием иностранных программных компонентов. Предложены ключевые направления для развития системы защиты информации.

Введение

Цифровая трансформация государственного управления, стимулируемая такими документами, как «Национальная стратегия развития искусственного интеллекта до 2030 года», делает внедрение ИИ в ГИС обязательным (imperative). Однако интеграция интеллектуальных систем, особенно с имеющейся зависимостью от иностранного программного обеспечения (ПО), создает новые векторы атак и риски утечек конфиденциальных данных. Это требует разработки специализированных средств и методологий защиты.

Ключевые проблемы и анализ литературы

Анализ работ [1–3] выявил основные исследовательские фокусы в области безопасности ИИ:

– таксономия ИИ-атак и разработка методов противодействия (adversarial attacks);

– риски использования ПО с открытым исходным кодом в жизненном цикле моделей ИИ;

– пробелы в нормативном регулировании и оценке соответствия ИИ-сервисов.

На примере сервиса автоматического анализа изображений в российских таможенных органах выявлена фундаментальная проблема: недостаточность чисто технических мер защиты без комплексного методологического и нормативного обеспечения на государственном уровне.

Перспективные направления обеспечения безопасности

На основе проведенного анализа предлагается сконцентрироваться на следующих ключевых мерах:

– *защита данных*: строгое шифрование данных, как при хранении, так и при передаче (с использованием TLS/SSL);

– *контроль доступа*: внедрение многофакторной аутентификации и ролевой модели доступа для минимизации рисков несанкционированного доступа;

– *защита моделей ИИ*: применение методов обфускации алгоритмов для противодействия реверс-инжинирингу (например, регулярное тестирование моделей на устойчивость к adversarial-атакам);

– *повышение осведомленности*: обязательное обучение персонала лучшим практикам кибербезопасности;

– *обеспечение целостности*: исследование применения блокчейн-технологий для верификации данных и отслеживания изменений в системе.

Выводы

Несмотря на сформировавшуюся терминологию, практическое применение безопасных ИИ-сервисов в ГИС остается нетривиальной задачей. Технические меры защиты (шифрование, аутентификация) необходимы, но недостаточны. Ключевой проблемой является отсутствие развитой системы отечественных стандартов и методик оценки безопасности и надежности ИИ. Требуется развитие комплексного подхода, объединяющего технические решения, нормативное регулирование и методологическую поддержку государственных органов. Это позволит обеспечить не только информационную безопасность, но и доверие, надежность и устойчивость критически важных ИИ-сервисов.

Список литературы

1. Костогрызов А.И., Нистратов А.А. Анализ угроз злоумышленной модификации модели машинного обучения для систем с искусственным интеллектом // Вопросы кибербезопасности. 2023. № 5 (57). DOI: http://dx.doi.org/10.21681/2311-3456-2023-5-9-24.

2. Марков, А. С. Важная веха в безопасности открытого программного обеспечения / А.С. Марков // Вопросы кибербезопасности. – 2023. – № 1(53). – С. 2–12. – DOI 10.21681/2311-3456-2023-1-2-12. – EDN OHYLTR.

3. Арустамян, Сас С.; Варениса, Виталий В.; Марков, Алексей С. Методические и реализационные аспекты внедрения процессов разработки безопасного программного обеспечения. Безопасность информационных технологий, [S.l.], v. 30, n. 2, p. 23–37, мая 2023. ISSN 2074-7136. Доступно на: <https://bit.spels.ru/index.php/bit/article/view/1499>. (Дата доступа: 24 окт. 2025) DOI: http://dx.doi.org/10.26583/bit.2023.2.01.

УДК 004.056

А.А. ГОЛУБЕВ
Научный руководитель – доцент Г.П. ГАВДАН
Национальный исследовательский ядерный университет «МИФИ», Москва

БЕЗОПАСНОСТЬ ГОСУДАРСТВЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ В РОССИИ

В работе анализируются проблемы обеспечения безопасности государственных информационных систем (ГИС) в условиях роста количества и сложности кибератак. Исследование сосредоточены на современных вызовах, правовом регулировании и ключевых направлениях защиты, включая объекты критической информационной инфраструктуры. Особое внимание в исследовании уделяется проблеме несогласованности источников данных об угрозах.

Введение

Государственные информационные системы (ГИС) являются критическим компонентом системы государственного управления, обеспечивающим его бесперебойное функционирование. В условиях цифровой трансформации они становятся основной мишенью для кибератак, что подтверждается статистикой, где уровень угроз для систем управления остается стабильно высоким.

Ключевые проблемы обеспечения информационной безопасности (ИБ) ГИС включают:

– запаздывание нормативно-правового регулирования, не успевающего за динамикой появления новых угроз;

– субъективность при формировании моделей угроз;

– постоянное обнаружение новых уязвимостей в программно-аппаратных комплексах.

Основные вызовы и правовой контекст

ГИС концентрируют национальные информационные ресурсы (ИР) - обширные фонды данных, требующие особой защиты и др. [1-3]. Это делает информационную инфраструктуру органов власти Российской Федерации (РФ) одним из главных объектов для целенаправленных атак.

Правовую основу защиты составляют законы в сфере информатизации и защиты (объектов) КИИ [1-3]. Однако на практике возникает проблема несогласованности источников данных об угрозах. Например, уязвимости, присутствующие в Национальной базе уязвимостей США (NVD) или др., могут отсутствовать в Базе данных уязвимостей ФСТЭК России (БДУ ФСТЭК России), что создает «слепые зоны» в системе защиты.

Ключевые направления повышения безопасности

Для противодействия актуальным угрозам необходимы:

– Совершенствование нормативной базы РФ: ускорение процессов актуализации требований к безопасности ГИС и КИИ;

– Практическая подготовка кадров: регулярное проведение учений по кибербезопасности для отработки действий при реальных инцидентах и повышения квалификации специалистов;

– Устранение информационных разрывов: необходима гармонизация отечественных и международных реестров угроз для формирования полной и непротиворечивой картины угроз безопасности информации.

Заключение

Проведенный анализ подтверждает критическую важность защиты ГИС и их отнесения к объектам КИИ. Основными проблемами являются не только технические угрозы безопасности, но и системные: запаздывание нормативного регулирования государства и несогласованность источников информации об уязвимостях. Полученные выводы будут использованы для разработки методики оценки угроз безопасности информации в ГИС.

Список литературы

1. Тулупов, А.С. Модель национальной безопасности: общие подходы и вопросы построения / А.С. Тулупов // Вестник Московского университета. Серия 6: Экономика. – 2023. – № 3. – С. 181–192. – DOI 10.55959/MSU0130-0105-6-58-3-9. – EDN WYPTHL.

2. Подсистема предупреждения компьютерных атак на объекты критической информационной инфраструктуры: анализ функционирования и реализации / И.В. Котенко, И.Б. Саенко, Р.И. Захарченко, Д.В. Величко // Вопросы кибербезопасности. – 2023. – № 1(53). – С. 13–27. – DOI 10.21681/2311-3456-2023-1-13-27. – EDN WUUDFK.

3. Наталичев, Роман В. et al. Эволюция и парадоксы нормативной базы обеспечения Безопасности объектов критической информационной инфраструктуры. Безопасность информационных технологий, [s.l.], v. 28, n. 3, p. 6–27, сен. 2021. ISSN 2074-7136. Доступно на: <https://bit.spels.ru/index.php/bit/article/view/1359/1234>. (Дата доступа: 24 окт. 2025) DOI: http://dx.doi.org/10.26583/bit.2021.3.01.

4. Пенерджи, Рустем В.; Гавдан, Григорий П. Информационная безопасность государственных информационных систем. Безопасность информационных технологий, [S.l.], v. 27, n. 3, p. 26-42, сен. 2020. ISSN 2074-7136. Доступно на: <https://bit.spels.ru/index.php/bit/article/view/1290>. (Дата доступа: 24 окт. 2025) DOI: http://dx.doi.org/10.26583/bit.2020.3.03.

5. Гавдан, Григорий П. et al. Обеспечение безопасности государственных информационных систем как объектов критической информационной инфраструктуры. Безопасность информационных технологий, [S.l.], v. 32, n. 3, p. 26-43, июля 2025. ISSN 2074-7136. Доступно на: <https://bit.spels.ru/index.php/bit/article/view/1822/1479>. Дата доступа: 24 окт. 2025. DOI: http://dx.doi.org/10.26583/bit.2025.3.03.

КИБ-2025
КИБЕРНЕТИКА
И ИНФОРМАЦИОННАЯ
БЕЗОПАСНОСТЬ

Направление

Защищенные компьютерные системы и технологии

Руководитель секции – ИВАНОВ М.А., д.т.н.,
заведующий кафедрой №12

УДК 004.056

М.П. ГРИГОРЬЕВ

Национальный исследовательский ядерный университет «МИФИ», Москва
ООО «Гексагон», Москва

МАНИПУЛЯЦИЯ СПЕКУЛЯТИВНЫМ ПОТОКОМ УПРАВЛЕНИЯ

Рассматривается атака, использующая спекулятивное исполнение для временного перехвата потока команд и обхода механизмов защиты памяти. Показано, что при целенаправленной манипуляции предсказателем ветвлений и задержкой проверок возможно спекулятивно перенаправить выполнение на произвольный участок кода до срабатывания защитных средств. Спекулятивно выполненные инструкции получают доступ к конфиденциальным данным и через побочные эффекты утечка информации становится возможной, хотя впоследствии архитектурное состояние программы корректируется.

Уязвимости памяти (переполнения буфера, «use-after-free» и пр.) продолжают систематически обнаруживаться в программных системах. Для защиты применяются как проверенные, так и новые меры: обнаружение и предотвращение переполнений (включая стековые канарейки), программно-аппаратные механизмы обеспечения целостности хода выполнения (CFI), а также переход к языкам с автоматическим управлением памятью, снижающим риск ошибок. В результате классические методы эксплуатации уязвимостей существенно затруднены, общий уровень безопасности современных приложений повысился.

Параллельно развитию защит возник новый класс угроз – микроархитектурные атаки, эксплуатирующие спекулятивное исполнение команд. Спекулятивное исполнение ускоряет программу, но иногда команды выполняются «заранее» и меняют содержимое кеша. Эти изменения не откатываются вместе с программой, поэтому по ним можно косвенно вычитать секреты. Ключевой вопрос: можно ли временно обойти традиционные защиты памяти до завершения проверок? В работе показано, что при тренировке предсказателя и задержке загрузки контрольных данных процессор начинает выполнять ветку, предполагая корректность, и тем самым допускает спекулятивный уход за границы разрешённых переходов.

Атака включает четыре этапа. (1) Подготовка: многократные «чистые» вызовы функции тренируют предсказатель; элементы, участвующие в проверках, вытесняются из кеша. (2) Вызов уязвимости: перезаписывается

адрес возврата/указатель функции, формируя адрес гаджета. (3) Спекулятивное выполнение: до завершения проверки CPU совершает переход к гаджету и выполняет несколько инструкций, в т. ч. обращение к массиву-передатчику по индексу, зависящему от секрета. (4) Откат и извлечение: после обнаружения нарушения архитектурное состояние откатывается, но выбранная линия массива остаётся в кеше.

Допускается использование последовательности гаджетов: первый извлекает секрет, второй формирует индекс выборки, третий выполняет доступ, вследствие чего возникает измеримый побочный эффект. Длина цепочки ограничена окном спекуляции, поэтому инструкции минимальны и предсказуемы. Подбор гаджетов ведётся в коде библиотек и исполняемого файла.

Экспериментально подтверждена реализуемость описанного сценария. В демонстрационной реализации на языке Си с использованием уязвимости переполнения удалось спекулятивно выполнить цепочку из нескольких гаджетов и получить доступ к данным за пределами допустимого диапазона памяти, тем самым обходя встроенные механизмы защиты. Успех атаки зависит от нескольких факторов: достаточной длительности спекулятивного окна, наличия в адресном пространстве пригодных гаджетов, а также устойчивости сайд-канального сигнала. В современных процессорах длина окна спекуляции ограничена, что позволяет выполнять лишь короткие цепочки ROP, но этого достаточно для утечки информации при аккуратной настройке. Показано также, что описанный подход применим не только к переполнению стека: аналогичным образом возможен спекулятивный обход forward-edge защиты и обход проверок границ массивов.

Для защиты от подобных уязвимостей эффективны следующие комбинации: барьеры сразу после критических проверок (LFENCE) для закрытия спекуляции, Retpoline/IBRS/STIBP против отравления BTB, компиляторное маскирование индексов перед ветвлением, очистка микробуферов на границах доверия, шум/рандомизация таймеров и предсказателя. Издержки включают понижение IPC, рост латентности и усложнение сопровождения. Полной защиты без архитектурных изменений пока не существует.

Показан сценарий спекулятивного обхода CFI/стековых канареек и практическая схема утечки через кеш-канал. Результаты подчёркивают необходимость сочетать барьерные инструкции и ограничения предсказателя с прикладными приёмами, а также пересматривать границы доверия при проектировании ПО и платформ.

УДК 004.056

М.П. ГРИГОРЬЕВ

Национальный исследовательский ядерный университет «МИФИ», Москва
ООО «Гексагон», Москва

РАСШИРЕНИЕ БАЗОВОЙ АРХИТЕКТУРЫ MORPHEUS С УСИЛЕННОЙ ЗАЩИТОЙ ОТ ВОЗВРАТНО-ОРИЕНТИРОВАННЫХ (ROP) И ПЕРЕХОДНО-ОРИЕНТИРОВАННЫХ (JOP) АТАК

Предлагается система исполнения кода для многоуровневой защиты от атак на основе повторного использования кода. Архитектура развивает идею «скользящей цели» (Morpheus) и сочетает проактивную рандомизацию, аппаратное дешифрование инструкций «на лету» и реактивное обнаружение атак. Встроенный генератор ПСЧ обеспечивает непредсказуемую смену режимов защиты и адресов. Модуль многоверсионного кода позволяет прозрачно ротировать реализации функций без прерывания работы. Модифицированный execute-only кеш ограничивает видимость и анализ исполняемого кода.

Повторное использование кода остаётся практичным способом обхода механик защиты памяти на современных платформах. Даже при адресной рандомизации и метках указателей злоумышленник может собрать цепочку гаджетов из легитимного кода. Базовая реализация Morpheus снижает вероятность успешной эксплуатации за счёт периодической перерандомизации адресов (churn) и маркировки данных. Тем не менее сохраняется теоретическая уязвимость: кратковременные атаки, запускаемые и завершаемые между циклами перестройки, могут обойти эти механизмы. Требуется архитектура, которая реагирует в момент детектирования и нарушает подготовленную цепочку до её завершения.

Предлагаемая система исполнения кода на базе Morpheus усиливает модель «скользящей цели» за счёт трёх принципов: (1) полиморфный код, при котором каждая функция имеет несколько взаимозаменяемых реализаций, прозрачно переключаемых модулем Code Version Manager, (2) execute-only хранение и дешифрование инструкций в момент выборки, привязанное к контексту перехода, (3) реактивный контур безопасности, запускаемый детектором атак. Ключи и смещения обновляются по непредсказуемому графику, задаваемому встроенным ГПСЧ, что разрушает результаты разведки противника.

Для оперативного обнаружения попыток эксплуатации система включает детектор поведенческих сигнатур, отслеживающий аномалии в

последовательностях RET и непрямых переходов. Признаком ROP-атаки служит аномально большое число инструкций возврата RET, выполняемых подряд без соответствующих вызовов CALL, а также частые промахи механизма возвратного стека, указывающие на нарушение привычного порядка вызовов. Для JOP-атак фиксируется череда непрерывных косвенных переходов с необычно короткими промежутками между ними. В систему заложены пороговые значения: при превышении нормальной частоты подряд исполненных RET или непрямых переходов детектор распознаёт атаку по сигнатуре.

При обнаружении ROP/JOP-сигнатуры архитектура инициирует комплекс немедленных защитных мер. Детектор атак подаёт сигнал в модуль CVM на срочное полиморфное переключение текущей выполняемой функции на альтернативную версию кода. Одновременно контроллер кеша инструкций получает команду на немедленную инвалидацию I-кеша и сброс конвейера, удаляя из процессора все оставшиеся инструкции, связанные с вредоносной активностью. Параллельно запускается внеочередной churn: генерируется новый ключ шифрования, и выполняется экстренная перерандомизация всех чувствительных данных в памяти. Таким образом, даже если злоумышленнику ранее удалось получить часть информации о системе, генерация нового ключа и перерандомизация всех чувствительных данных в памяти делают эти данные бесполезными для дальнейшей атаки.

Эффективный ключ дешифрования инструкции формируется из базового секрета и контекста. Если происходит попытка прыжка на гаджет вне предусмотренной последовательности, вычисляется «неправильный» ключ, и дешифровка выдаёт некорректный код операции. Это обеспечивает целостность управления: инструкции исполняются только при достижении адреса легитимным путём.

Полиморфизм и привязанное шифрование добавляют накладные расходы на выборку инструкций и переключение версий. Тем не менее в обычной работе, благодаря кэшированию ключей и редким срабатываниям защитных механизмов, падение производительности обычно не превышает нескольких процентов. В соотношении «безопасность - производительность» выигрыш по безопасности значительно превосходит потери в быстродействии.

Список литературы

1. Gallagher M. et al. Morpheus: A Vulnerability-Tolerant Secure Architecture Based on Ensembles of Moving Target Defenses with Churn. In: Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS ’19), Providence, RI, USA, April 2019. ACM. DOI: 10.1145/3297858.3304037

УДК 004.056

А.А. ВАГАНОВА

МИРЭА – Российский технологический университет, Москва

КОНТЕЙНЕРЫ И БЕЗОПАСНОСТЬ: АНАЛИЗ УГРОЗ И МЕХАНИЗМОВ ЗАЩИТЫ В СРЕДАХ DOCKER И KUBERNETES

В статье исследована проблема обеспечения безопасности контейнеров в средах Docker и Kubernetes. Рассмотрены виды уязвимостей, а также способы их предотвращения. Отмечена важность проведения грамотной настройки и эксплуатации среды контейнеров. Показана необходимость защиты технологий контейнеризации как важной составляющей любой современной инфраструктуры.

Введение

В последние годы контейнеризация стала важной частью разработки и развертывания приложений, предоставляя разработчикам новые возможности для улучшения гибкости, масштабируемости и скорости доставки программных решений. Далее будут рассмотрены риски, которые могут возникнуть при работе с контейнерами, а также практики, позволяющие обеспечить их безопасность.

Уязвимости контейнеров

К уязвимостям контейнеров относятся:

- **Цепочки поставок.** Вирус, внедренный в публичные контейнеры, распространяется на все использующиеся экземпляры, что приводит к массовым компрометациям систем.
- **Привилегированные контейнеры.** Контейнеры, запущенные с повышенными привилегиями (root), могут предоставить злоумышленнику доступ к хост-системе, чтобы использовать уязвимости на уровне его ядра для выполнения вредоносного кода.
- **Образы контейнеров.** Образы могут содержать уязвимости, возникающие из-за неактуальных пакетов или программного обеспечения. Использование устаревших библиотек может подвергнуть контейнер атакам.
- **Небезопасные настройки конфигурации.** Неправильно сконфигурированные контейнеры могут привести к утечкам данных. Например, контейнеры не должны иметь возможность взаимодействовать с сетевыми компонентами.

Способы устранения уязвимостей Docker:

1. **Ограничение привилегий.** По умолчанию Docker запускает контейнеры с базовым набором привилегий. Параметр «--cap-drop» сбрасывает ненужные разрешения.

2. **Управление сетевыми политиками и изоляция**. Изоляция контейнеров через настройки сети помогает предотвратить несанкционированный доступ и утечку данных.

3. **Обновление и проверка образов.** Для обеспечения безопасности контейнеров Docker важно всегда использовать образы из авторитетных источников. Помимо этого, необходимо проверять целостность загружаемых образов, используя цифровые подписи.

4. **Регулярное сканирование на уязвимости.** Необходимо проверять образы на наличие уязвимостей и небезопасного наполнения. Для этого существуют такие инструменты, как Clair или Trivy, а также встроенные функции безопасности.

Способы устранения уязвимостей Kubernetes:

1. **Тщательная настройка RBAC (Role-Based Access Control).** Ограничение и разделение доступа к ресурсам Kubernetes с использованием принципа наименьших привилегий.

2. **Регулярные обновления и сканирования.** Важно периодически проверять контейнеры, например, с помощью программ Clair и Trivy.

3. **Использование инструментов для управления секретами.** Например, HashiCorp Vault, позволяющий безопасно хранить и иметь доступ к секретам.

4. **Настройка сетевых политик.** Использование TLS для защиты всех каналов связи, контроль трафика, минимизация открытия ненужных портов приложений.

Заключение

Защита контейнеров – важнейший аспект современной разработки программного обеспечения. Следуя рекомендациям по обеспечению безопасности Docker и Kubernetes, возможно избежать утечек данных и иных угроз информационной безопасности.

Список литературы

1. Райс Лиз Безопасность контейнеров. Фундаментальный подход к защите контейнеризированных приложений. – СПб.: Питер, 2021. – 224 с.

2. Обзор способов защиты контейнеров Docker: от простого к сложному [электронный ресурс]. – Режим доступа: https://habr.com/ru/companies/selectel/articles/854850/, свободный (дата обращения: 20.10.2025).

УДК 004.056

Б.О. ЛАВРОВ

Национальный исследовательский ядерный университет «МИФИ», Москва

МОДУЛЬ ОБНАРУЖЕНИЯ АНОМАЛИЙ В КОНТЕЙНЕРИЗОВАННЫХ ПРИЛОЖЕНИЯХ НА ОСНОВЕ АНСАМБЛЕВЫХ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Представлен модуль обнаружения аномалий в Docker-контейнерах с использованием ансамбля алгоритмов машинного обучения. Реализовано объединение алгоритмов Isolation Forest, One-Class SVM, Local Outlier Factor и DBSCAN посредством взвешенного голосования. Разработан компонент генерации текстовых объяснений на русском языке. Для валидации системы создано программное средство симуляции аномальной активности. Архитектура модуля предусматривает интеграцию с системами мониторинга Prometheus и развертывание в среде Kubernetes.

Рост популярности контейнеризации сопровождается увеличением количества целевых атак: согласно отчету Google, значительная доля инцидентов безопасности связана с несанкционированным майнингом криптовалют [1]. Традиционные системы мониторинга, основанные на статических пороговых значениях, демонстрируют высокий уровень ложных срабатываний в динамичной среде контейнеров. Сигнатурные анализаторы, такие как Falco [2], требуют постоянного обновления правил для поддержания актуальности. Перспективным направлением является применение методов машинного обучения, демонстрирующих способность к адаптивному выявлению неизвестных аномалий. Известные решения, такие как KAD (Kubernetes Anomaly Detector), реализуют автоматический выбор модели [3], а в работе Kim et al. достигнута точность 99.7% в задаче детекции криптомайнеров с использованием машинного обучения и технологии eBPF (extended Berkeley Packet Filters) [4]. Целью данной работы является разработка модуля обнаружения аномалий, сочетающего несколько алгоритмов машинного обучения для повышения достоверности и надежности детекции.

Ключевым компонентом разработанного модуля является ансамблевый детектор, агрегирующий решения четырех алгоритмов: Isolation Forest для выявления глобальных выбросов, One-Class SVM (Support Vector Machine) для анализа разреженных областей признакового пространства, LOF (Local Outlier Factor) для оценки локальной плотности и DBSCAN (Density-Based Spatial Clustering of Applications with Noise) для

кластерного анализа и поиска аномальных режимов работы. Сбор метрик производительности (загрузка центрального процессора, потребление оперативной памяти, сетевой трафик) осуществляется с использованием Docker API. Для повышения интерпретируемости результатов разработан модуль объяснений, который формирует текстовые описания на русском языке с указанием затронутых метрик и потенциальных причин инцидента. Доступ к результатам мониторинга предоставляется через REST API (Application Programming Interface).

Для тестирования функциональности модуля было реализовано специализированное программное обеспечение, имитирующее различные сценарии аномалий: резкий рост нагрузки на центральный процессор, утечку памяти, сетевые атаки типа «флуд» и неконтролируемое создание файлов.

Разработанный модуль подтверждает практическую применимость ансамблевых методов машинного обучения для обнаружения аномалий в контейнеризованных средах. К основным преимуществам решения относятся адаптация к изменяющемуся профилю нагрузки, повышение надежности детекции за счет комбинации алгоритмов и обеспечение интерпретируемости решений. Перспективными направлениями дальнейших исследований являются интеграция с технологией eBPF для низкоуровневого анализа системных вызовов и развертывание системы в промышленной эксплуатации.

Список литературы

1. Google. Threat Horizons – Cloud Threat Intelligence. Issue 1. November 2021 // Google Cloud Security. URL: https://bit.ly/41THxbT (дата обращения: 15.10.2025).

2. Falco: Cloud Native Runtime Security // Cloud Native Computing Foundation (CNCF). URL: https://falco.org (дата обращения: 15.10.2025).

3. Kosińska J., Tobiasz M. Detection of Cluster Anomalies with ML Techniques // IEEE Access. 2022. Vol. 10. P. 110742–110753. DOI: 10.1109/ACCESS.2022.3216080.

4. Kim R., Ryu J., Kim S., Lee S., Kim S. Detecting Cryptojacking Containers Using eBPF‑Based Security Runtime and Machine Learning // Electronics. 2025. Vol. 14, No. 6. Art. 1208. DOI: 10.3390/electronics14061208.

УДК 004.056

И.Ю. КОРКИН

Positive Technologies, Москва

О ПОДХОДЕ К ЗАЩИТЕ ПРИЛОЖЕНИЙ КЛАССА «МЕНЕДЖЕР ПАРОЛЕЙ» ОТ ПЕРЕХВАТА КЛАВИАТУРНОГО ВВОДА

В работе рассматривается проблема защиты приложений класса «менеджер паролей» в Windows ОС от атак перехвата клавиатурного ввода. Существующие решения не учитывают возможности нарушителей по противодействию защитным механизмам. Предложен и реализован подход, устраняющий недостатки существующих решений путём интеграции возможностей встроенных механизмов защиты ОС и гипервизора. Проведённые эксперименты подтвердили практическую применимость нового подхода в условиях смоделированных атак.

В настоящее время одной из приоритетных целей нарушителей выступают приложения класса «менеджер паролей», используемые для хранения и управления цифровыми секретами. Атаки направлены на компрометацию файлов хранилищ и перехват клавиатурного ввода [1]. Ситуацию усугубляют атаки «Man-At-The-End» (MATE) – нарушитель с помощью легитимного доступа преодолевает механизмы защиты [2].

Решаемая задача заключается в защите приложений класса «менеджер паролей» от атак типа «перехвата клавиатурного ввода» в Windows ОС.

Модель нарушителя рассматривает пользователя, который обладает привилегиями, достаточными для перехвата нажатий клавиш с помощью штатных интерфейсов (функций WinAPI) [3]. Атака «Bring Your Own Vulnerable Driver» (BYOVD) позволяет нарушителю модифицировать содержимое памяти ядра с целью обхода защитных механизмов ОС [4].

Несмотря на существование ряда техник защиты для обеспечения безопасного ввода, они зачастую не учитывают противодействие нарушителей и не в полной мере отвечают современным вызовам. Так, технология «Secure Desktop», используемая для безопасного ввода мастер‑ключа в «KeePass Password Safe», «1Password» и «VeraCrypt», уязвима к несанкционированному подключению процессов к созданному защищённому рабочему столу [5]. Попытка устранить эту проблему путём проверки открытых дескрипторов для всех запущенных процессов оказалась уязвимой к атаке «Time-Of-Check-to-Time-Of-Use» (TOCTOU), что делает предложенное решение непригодным для систем с большим числом активных процессов [6].

В качестве альтернативы A. Nakajima [3] предложила применить технологию «Event Tracing for Windows» (ETW) для отслеживания вызовов WinAPI-функций, используемых для перехвата ввода. Однако технология ETW также может быть отключена нарушителями [7].

Предлагается новый подход для защиты клавиатурного ввода от перехвата, который устраняет недостатки существующих решений и включает в следующие этапы. 1. Отзыв прав доступа «winsta_enumdesktops» у объекта «Station» для запрета получения списка рабочих столов. 2. Создание нового рабочего стола, используя технологию «Secure Desktop». 3. Отзыв прав доступа «desktop_hookcontrol» у объекта «Desktop» нового стола для запрета установки перехвата ввода. 4. Переключение на созданный стол и запуск приложения «менеджер паролей». 5. Отзыв прав доступа «desktop_switchdesktop» для созданного стола для запрета подключения других потоков. д6. Установка флагов «Process Protected Light» (PPL) и «ProcessMitigationPolicy» для приложения «менеджер паролей» с целью защиты от популярных атак внедрения кода. 7. Использование гипервизора MemoryRanger с открытым исходным кодом для защиты целостности объектов «Station» и «Desktop» с применением технологий Intel VT-x и Extended Page Tables (EPT).

Реализованный прототип на C/C++ подтвердил пригодность подхода для Windows 11 25H2 x64: изолированный рабочий стол с ограниченными правами доступа позволил противостоять смоделированным атакам – перехвату ввода, подключению к столу, внедрению кода и атакам BYOVD.

Список литературы

1. Gallus P., Stanek D., Klaban I. Security Evaluation of Password Managers: A Comparative Analysis and Penetration Testing of Existing Solutions // International Conference on Cyber Warfare and Security (ICCWS 2025). – Williamsburg, Virginia, USA, 2025.

2. Collins S., Poulopoulos A., Muench M., Chothia T. Anti Cheat: Attacks and the Effectiveness of Client Side Defences // Workshop on Research on offensive and defensive techniques in the context of Man at The End (MATE) attacks. – New York, NY, USA, 2024.

3. Nakajima A. Windows Keylogger Detection: Targeting Past and Present Keylogging Techniques // NULLCON Goa 2025 Conference. – Goa, India, 2025.

4. Federal Office for Information Security (BSI). The state of IT Security in Germany in 2024. – URL https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/Securitysituation/IT-Security-Situation-in-Germany-2024.pdf (дата обращения: 10.10.2025).

5. Denkiewicz A. Master Password can be keylogged from “Secure desktop” – URL: https://github.com/adenkiewicz/KeepassKeylogger (дата обращения: 10.10.2025).

6. Oliveira B., Almeida M. Bypassing Secure Desktops Protections // Black Hat Europe Arsenal. – Amsterdam, Netherlands, 2014.

7. Korkin I., Teodorescu C., Golchikov A. Blasting Event-Driven Cornucopia: WMI-based User-Space Attacks Blind SIEMs and EDRs // Black Hat USA Conference. – Las Vegas, USA, 2022.

УДК 004.056

Д.В. ФИЛИППОВ, Е.И. НОВОЖИЛОВ

МИРЭА – Российский технологический университет, Москва

ОБЗОР И АНАЛИЗ ПРОБЛЕМ ИСПОЛЬЗОВАНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ НАПИСАНИЯ ИСХОДНОГО КОДА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

В работе рассматриваются ключевые проблемы, возникающие при применении искусственного интеллекта (ИИ) для автоматизации процессов программирования. Проанализированы основные риски, связанные с генерацией ошибочного или небезопасного кода, вопросами конфиденциальности и авторского права. Отмечено влияние качества входных данных и промптов на итоговый результат. Обоснована необходимость разработки методов контроля и верификации кода, созданного ИИ. Сформулированы рекомендации по безопасному использованию языковых моделей в программной инженерии.

Введение

В последние годы большие языковые модели (LLM) заняли центральное место в программной инженерии, предлагая разработчикам мощные инструменты для автоматизации рутинных задач. С их помощью можно генерировать код, проводить рефакторинг, писать тесты и документацию. Однако, несмотря на огромный потенциал, использование искусственного интеллекта для написания кода сопровождается рядом серьёзных проблем и ограничений, требующих внимательного анализа [1].

Одной из ключевых проблем является феномен так называемых «галлюцинаций» моделей. LLM способный создавать синтаксически корректный, но логически ошибочный код. Поскольку генерация основана на вероятностных закономерностях, а не на истинном понимании логики программы, модель может предложить решение, которое выглядит убедительно, но не работает корректно. Это требует от разработчика постоянного контроля и верификации результатов, что снижает эффективность автоматизации.

Обзор и анализ проблем

Современные языковые модели не обладают реальным пониманием предметной области или бизнес-логики проекта. Они оперируют статистическими закономерностями, а не причинно-следственными связями. Это ограничивает их применение в сложных сценариях, где требуется глубокое понимание контекста и архитектуры системы. Даже

при работе с длинным контекстом LLM могут упустить важные зависимости между частями кода, что ведёт к появлению скрытых ошибок [2].

Существенное значение имеют также этические и правовые вопросы. Использование кода, сгенерированного ИИ, вызывает дискуссии о правах собственности и авторстве. Кто является автором программы – разработчик или модель, обученная на чужих данных? Этот вопрос остаётся открытым и требует нормативного урегулирования. Также существует риск переноса предвзятости из обучающих данных в генерируемые решения, что может сказаться на качестве и справедливости создаваемого программного обеспечения [3].

Помимо этого, стоит отметить проблему зависимости от качества промптов. Даже незначительные изменения формулировки запроса могут кардинально повлиять на результат. Эффективная работа с LLM требует навыков промпт-инжиниринга, что создаёт дополнительный барьер для специалистов без соответствующего опыта. Это рождает новый тип компетенций – «инженер промптов», который становится всё более востребованным на рынке труда [4].

Заключение

Таким образом, искусственный интеллект для написания кода является мощным инструментом, но его применение должно сопровождаться строгими процедурами контроля качества, тестирования и проверки безопасности. В перспективе развитие гибридных систем, сочетающих человеческий надзор и автоматизацию, может стать ключом к эффективному и безопасному использованию ИИ в программной инженерии.

Список литературы

1. Кузнецов И.В., Смирнова А.Н. Применение больших языковых моделей в программной инженерии: риски и преимущества. Информационные технологии и вычислительные системы, 2024, № 3(97), с. 14–26. DOI: http://dx.doi.org/10.21455/itvs.2024.3.2.

2. Милославская Н.Г., Толстой А.И. Управление информационной безопасностью. – М.: НИЯУ МИФИ, 2020. – 536 с.

3. Павлов Д.С. Этические и правовые аспекты использования генеративного ИИ в разработке программного обеспечения. Вопросы кибербезопасности, 2023, № 2(54), с. 34–45. DOI: http://dx.doi.org/10.21681/2311-3456-2023-2-34-45.

4. Швецов Е.А., Чернова К.В. Методы контроля качества кода, сгенерированного искусственным интеллектом. Программные продукты и системы, 2024, т. 37, № 1, с. 51–60. DOI: http://dx.doi.org/10.15827/0236-235X.137.051-060.

УДК 004.056

Е.С. ГРИПАС, К.Р. ОПЛЕУХИН, М.В. РИММЕР

МИРЭА – Российский технологический университет, Москва

СПОСОБЫ ПРОТИВОДЕЙСТВИЯ УГРОЗАМ ОТКАЗА В ОБСЛУЖИВАНИИ ДЛЯ СОВРЕМЕННЫХ ИНТЕРПРЕТИРУЕМЫХ ЯЗЫКОВ ПРОГРАММИРОВАНИЯ

В статье рассматриваются современные методы противодействия атакам типа «отказ в обслуживании», нацеленным на приложения, разработанные с использованием интерпретируемых языков программирования. Подчёркивается важность комплексной стратегии, учитывающей особенности интерпретируемых языков и современные тенденции в сфере кибербезопасности, для повышения устойчивости приложений к возникающим угрозам информационной безопасности.

Современное развитие информационных технологий сопровождается постоянным ростом количества и сложности кибератак, направленных на нарушение доступности цифровых сервисов. По оценкам экспертов в области информационной безопасности, количество атак, направленных на отказ в обслуживании, неуклонно растет. Подобные инциденты приводят к временной недоступности информационных ресурсов, снижению производительности систем и значительным экономическим потерям организаций [1].

С другой стороны, интерпретируемые языки программирования стали основой для построения большинства современных приложений и веб-сервисов. Их популярность объясняется гибкостью, и простотой разработки, что одновременно создает ряд уязвимостей. Главные проблемы связаны с более низкой производительностью, динамической типизацией и автоматическим управлением памятью, что потенциально может быть использовано злоумышленниками для реализации угрозы отказа в обслуживании. Таким образом, рассмотрим методы противодействия угрозам отказа в обслуживании для современных интерпретированных языков программирования.

Современные способы противодействия угрозам отказа в обслуживании для интерпретированных языков программирования включают в себя архитектурные решения, оптимизацию кода и использование специальных средств сетевой защиты.

Архитектурные методы противодействия угрозам отказа в обслуживании представляют собой не просто набор отдельных

инструментов, а целостный подход к проектированию систем, изначально наделяющий их устойчивостью к перегрузкам, который базируется на трёх фундаментальных принципах: избыточности, распределенности и масштабируемости [2].

В дополнение к архитектурным методам, оптимизация программного кода направлена на создание устойчивой и ресурсоэффективной системы, способной противостоять целенаправленным попыткам исчерпания вычислительных ресурсов. Ключевой задачей является устранение узких мест в производительности через эффективное управление памятью, а также оптимизацию алгоритмов обработки данных, что особенно важно для интерпретируемых языков программирования.

Кроме того, критически важным является использование специализированных средств сетевой защиты. К таким средствам относятся решения, которые обладают разнообразным функционалом, включая автоматическое реагирование на атаки, защиту от различных форм атак типа «отказ в обслуживании» и возможность использования сложных алгоритмов для анализа и формирования отчетов для улучшения стратегий обеспечения безопасности

Также методы машинного обучения становятся все более популярными для обнаружения и предотвращения угроз отказа в обслуживании. Они позволяют выявлять аномальные паттерны в использовании ресурсов и поведении пользователей, которые могут указывать на попытки атаки.

Защита приложений на интерпретируемых языках от угроз отказа в обслуживании требует глубокого понимания особенностей их работы. Только комплексный подход, учитывающий особенности интерпретируемых языков и современные тенденции в кибербезопасности, может обеспечить надежную защиту от постоянно развивающихся угроз отказа в обслуживании.

Список литературы

1. Карташова Ю.М. Противодействие угрозам типа «отказ в обслуживании» в условиях цифровой трансформации / Ю.М. Карташова, А.Г. Анацкая // Цифровизация и кибербезопасность: современная теория и практика : сборник материалов IV Международной научно-практической конференции, Омск, 30–31 окт. 2024 г. / Сибирский государственный автомобильно-дорожный университет (СибАДИ). – Омск: СибАДИ, 2025. – С. 413–425.

2. D. Mohammed Sharif, H. Beitollahi, and M. Fazeli, "Detection of Application-Layer DDoS Attacks Produced by Various Freely Accessible Toolkits Using Machine Learning", in IEEE Access, vol. 11, pp. 51810-51819, 2023.

УДК 004.056

Д.В. КРАЮШКИН[1], А.М. ЧЕПОВСКИЙ[1,2]

[1]Национальный исследовательский университет «ВШЭ», Москва
[2]Российский экономический университет им. Г.В. Плеханова, Москва

КОМПЛЕКСНОЕ УПРАВЛЕНИЕ ИНФРАСТРУКТУРОЙ PACS/РИС

Разработан программный комплекс для администрирования информационных систем PACS/РИС. Решение позволяет управлять цифровыми сервисами, информационными моделями DICOM и PACS, контролировать доступ и отслеживать события в едином интерфейсе. Внедрение системы позволяет снизить сложность эксплуатации медицинской ИТ-инфраструктуры.

Введение

Масштабирование медицинских информационных систем (МИС), рост числа исследований и добавление новых компонентов усложняют поддержку ИТ-инфраструктуры [1]. В качестве решения предлагается программное обеспечение (ПО) для централизованного администрирования радиологической информационной системы (РИС) в составе МИС [2]. Основу РИС составляет PACS (Picture Archiving and Communication System) – система хранения, передачи и обработки медицинских изображений.

Компоненты для сопровождения PACS/РИС

Разработанное ПО осуществляет управление информационными моделями компонентов системы, мониторинг сетевого трафика с учетом специфики используемых стандартов, управление цифровыми сервисами, контроль доступа к ресурсам, а также журналирование всех событий с установлением взаимосвязей между бизнес-процессами, цифровыми службами и информационными моделями [2]. Администрирование осуществляется по следующим компонентам и функциям:

Информационные модели PACS и DICOM. Модели PACS и DICOM – это наборы атрибутов различных сущностей из информационных моделей стандартов, которые задействованы при хранении, передаче или обработке данных. Являются информационной абстракцией физических, клинических или сетевых объектов (прибор, пациент, назначение и т.д.), а также различных конфигурационных параметров.

Цифровые службы и их конфигурация. Это программные модули и приложения, которые реализуют бизнес-процессы системы, отвечая за хранение, передачу и обработку информации на основе собственной конфигурации и информационных моделей.

Обмен данными. Стандартизированная процедура обменом данными происходит как внутри PACS/РИС, так и с внешними системами, такими как МИС, ЦАМИ (Центральный Архив Медицинских Исследований) и др. Структура сообщений и порядок передачи определяются наиболее используемыми медицинскими стандартами DICOM и HL7.

Контроль доступа. Доступ к ресурсам обеспечивается вспомогательной службой единой авторизации с протоколами аутентификации OAuth 2.0. Управление пользователями и их правами происходит централизовано.

Журналирование событий. Мониторинг событий системы требуется как на уровне конкретной цифровой службы или информационной сущности, так и на уровне этапа жизненного цикла данных с отслеживанием их изменений. Каждое событие ассоциируется с медицинскими данными, цифровыми службами и информационными моделями PACS и DICOM.

Заключение

Разработанное решение упрощает обслуживание и повышает эффективность использования медицинских информационных систем. Это достигается за счет учета взаимосвязей между информационными моделями, цифровыми службами и интегрируемыми системами, а также поддержки отраслевых стандартов DICOM и HL7.

Для защиты данных и контроля доступа внедрена система централизованной аутентификации и авторизации. Данный подход обеспечивает масштабируемость и позволяет внедрять дополнительные средства защиты, такие как электронная подпись и стеганография медицинских изображений.

Список литературы

1. Гаврилов А.В., Краюшкин Д.В., Куликов И.В., Соломинов М.В., Чеповский А.М. Современные подходы к сопровождению медицинских информационных систем при внедрении новых технологий. Вопросы кибербезопасности. 2025, № 2(66), с. 41–51. DOI: http://dx.doi.org/ 10.21681/2311-3456-2025-2-41-51. – EDN: TFUUII

2. Гаврилов А.В., Куликов И.В., Краюшкин Д.В. Программа графического пользовательского интерфейса для сопровождения информационной системы PACS/RIS. Свидетельство о государственной регистрации программы для ЭВМ. Российская Федерация, № 2025613353, опубл. 6.03.2025. – EDN: AQZAMI

УДК 004.056

Т.С. ОЛЕЙНИКОВА, И.Н. ЦЫГУЛЕВ, И.А. ЧЕПУРКО

Южно-Российский государственный политехнический университет (НПИ) им. М.И. Платова, Новочеркасск

ИЗБЫТОЧНОСТЬ КОДОВ РИДА-СОЛОМОНА В ПРОЦЕДУРАХ СОГЛАСОВАНИЯ QKD

В работе исследуется влияние избыточности кодов Рида-Соломона на безопасность квантового распределения ключа (QKD) на этапе классического согласования. Показано, что передаваемые по публичному каналу синдромы увеличивают утечку информации $Leak_{EC}$, что снижает минимальную энтропию и длину итогового ключа. Рассмотрены асимптотический и конечноразмерный подходы, а также потенциальная угроза, связанная с накоплением синдромов при многократных передачах.

Введение

В протоколах квантового распределения ключа (QKD) этап классического согласования непросеянных ключей критически важен для обеспечения безопасности, поскольку после процедуры согласования базисов в ключе могут сохраняться ошибки до 11% в битах просеянного ключа, которые в дальнейшем предлагается исправлять с помощью кодов Рида–Соломона (RS) [1]. При этом использование кодов Рида-Соломона, несмотря на их высокую корректирующую способность, вносит избыточность, которая приводит к утечке информации $Leak_{EC}$ по открытому каналу и, как следствие, снижает длину безопасного ключа [2].

Постановка задачи

Оценить влияние избыточности RS-кодов на безопасность QKD с точки зрения теории информации, в частности через минимальную и гладкие энтропии. Проанализировать, как утечка информации в процессе согласования сокращает длину итогового ключа, и выявить потенциальные атаки, связанные с многократной передачей синдромов.

Пути решения и результаты

В работе предложен комплексный подход к анализу влияния избыточности кодов Рида-Соломона на безопасность протоколов квантового распределения ключа, сочетающий как асимптотические, так и конечноразмерные методы оценки утечки информации [3]. Показано, что объём утечки $Leak_{EC}$ можно аппроксимировать выражением:

$$Leak_{EC} \approx f_{EC} * h(QBER)$$

где коэффициент эффективности кода $f_{EC} \geq 1$ учитывает неоптимальность процедуры декодирования, а бинарная энтропия $h(QBER)$ отражает уровень шума в канале [4]. Особое внимание уделено сценарию многократных передач, при котором злоумышленник, накапливая синдромы от последовательных блоков коррекции, может проводить статистический анализ и постепенно уточнять свои оценки относительно просеянных ключей, не превышая порогового значения *QBER*. Для адекватной оценки безопасности в реальных системах с конечным числом сигналов (порядка 10^5–10^6) применён аппарат гладких энтропий, позволяющий учесть отклонения от идеализированных асимптотических условий и более точно оценить длину безопасного ключа с учётом желаемого уровня параметра безопасности [5]. Полученные результаты демонстрируют, что избыточность кодов Рида-Соломона напрямую снижает минимальную энтропию ключа и, следовательно, сокращает его итоговую длину, что подтверждает необходимость включения параметров кодирования в общую модель безопасности QKD.

Заключение

Избыточность кодов Рида-Соломона в процедурах согласования *QKD* – не просто технический параметр, а критический фактор безопасности. Её влияние необходимо учитывать при проектировании протоколов, особенно в условиях конечного размера и возможности многократного наблюдения за синдромами со стороны злоумышленника. Требуется баланс между корректирующей способностью и минимизацией утечки информации для обеспечения максимальной длины безопасного ключа.

Список литературы

1. Margarida Pereira. Quantum key distribution with flawed and leaky sources. NPJ quantum information. 2019, DOI: https://www.nature.com/articles/s41534-019-0180-9.
2. Macro Tomamichel, Antothony Leverrier. A largely self-contained and complete security proof for quantum key distribution. Quantum. 2017, с. 1-38. DOI: https://arxiv.org/pdf/1506.08458.
3. Walter O. Krawec. A new security proof for Twin-Field Quantum key distribution (QKD). MDPI. 2023, DOI: https://www.mdpi.com/2076-3417/14/1/187.
4. Francisco Revson Fernandes Pereira. Error probability mitigation in quantum reading using classical codes. MDPI, 2021, DOI: https://pmc.ncbi.nlm.nih.gov/articles/PMC8775009.
5. Marco Baldi. Security of generalized Reed-Solomon code-based cryptosystems. The Institution of Engineering and Technology, 2019, DOI: https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/iet-ifs.2018.5207.

УДК 004.056

П.П. КАРАБУЩЕНКО
Научный руководитель – д.ф-м.н., профессор С.О. СТАРКОВ
Обнинский институт атомной энергетики – филиал НИЯУ «МИФИ»

ИНТЕГРАЦИЯ СИСТЕМЫ RAG И МЕТОДА LORA: ПРИЁМЫ ДООБУЧЕНИЯ И СПОСОБЫ ОБРАБОТКИ ИНФОРМАЦИИ В ЗАЩИЩЁННЫХ СРЕДАХ

Целью работы является изучение и применение методов, направленных на создание защищённых систем на основе локальных LLM, интегрированных с RAG для работы с внутренними данными. Подобная стратегия способствует обеспечению информационной безопасности за счёт полного контроля над данными и вычислительной инфраструктурой. Метод LoRA позволяет повысить точность моделей благодаря тонкой настройки на корпоративных данных.

Современные большие языковые модели (LLaMA, GPT, DeepSeek, YandexGPT, Qwen) показали значительный прогресс в обработке естественного языка и в анализе текстовой информации. Тенденции развития искусственного интеллекта сопряжены с рядом системных проблем: использование ИИ с открытым исходным кодом и открытыми весами; проблемы создания набора научных данных мирового класса; степень безопасности контроля и надёжности ИИ; возможность обеспечения гражданских свобод и защиты конфиденциальности личной жизни; степень защищённости государственной тайны и безопасности; ограничение доступа для научного сообщества и промышленного сектора к архитектурам моделей.

Для обеспечения безопасной работы с LLM предлагается использование локальных языковых моделей с открытым исходным кодом и открытыми весами. На первом этапе проведённого эксперимента была доработана система Retrieval-Augmented Generation (RAG) [1] и протестирована на информационной базе данных, взятых из открытых источников сайта ИАТЭ НИЯУ МИФИ. Архитектура RAG включает языковую модель generator и поисковый модуль retriever. В качестве основы для генератора использовались открытые модели YandexGPT-5-Lite и Mistral-v0.3, адаптированные для локального использования с применением 4-битного квантования. Для векторного поиска (retriever) был применён индекс FAISS [2], который позволяет: быстро находить нужную информацию по плотным эмбеддингам и отслеживать её источники. Безопасность обеспечивается локальным хранением

информации. Механизм взаимодействия между retrieval и generation был построен по принципу retriever-centric pipeline [3, 4], который позволяет системно работать с данными и повышать достоверность ответов.

На втором этапе эксперимента с целью улучшения качества обработки данных с помощью LLM, а также снижения риска возникновения галлюцинаций был использован метод Low-Rank Adaptation (LoRA) [5]. Дообучение проводилось на обработанных выборках из массива информации. Данный подход позволил внести доменные знания без изменения базовых весов модели, что существенно снизило вычислительные затраты и позволило сохранить целостность оригинальной архитектуры.

Результаты, полученные в ходе двух серий экспериментов, убедительно показали, что использование системы RAG существенно повышает точность анализа семантических связей и позволяет извлекать нужные данные из специализированных текстовых коллекций в локальных вычислительных средах.

Процедура тонкой настойки модели (LoRA) способствовала упрощению работы со специфичной информацией и инструкциями, положительно повлияв на общий результат системы. Предложенный в ходе исследования комбинированный подход позволил комплексно проанализировать и протестировать узкоспециализированную информацию, который редко применяется специалистами, однако имеет большой потенциал для практического использования.

Разработанный метод интеграции RAG и LoRA в защищённых системах на основе локальных LLM может применяться в ведомственных информационно-аналитических центрах, на предприятиях оборонного комплекса и в научных организациях, для которых необходима автономная обработка данных и соблюдение требований по защите информации.

Список литературы

1. Lewis P. et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP. NeurIPS, 2020. arXiv:2005.11401.

2. Johnson J., Douze M., Jégou H. Billion-Scale Similarity Search with GPUs (FAISS). arXiv:1702.08734.

3. Gupta R., Ranjan S., Singh M. A Comprehensive Survey of Retrieval-Augmented Generation. arXiv:2403.01085.

4. Fan C. et al. A Survey on RAG Meeting LLMs: Towards Retrieval-Augmented LLMs. arXiv:2401.16925.

5. Hu E. J. et al. LoRA: Low-Rank Adaptation of Large Language Models. arXiv:2106.09685.

УДК 004.239

А.Д. ЕСАКОВ

Национальный исследовательский ядерный университет «МИФИ», Москва

ИССЛЕДОВАНИЕ И РАЗРАБОТКА ПРОГРАММНО-АППАРАТНОГО МЕХАНИЗМА ЗАЩИТЫ ИНФОРМАЦИИ В КОМПЬЮТЕРНЫХ СИСТЕМАХ НА ОСНОВЕ МЕТОДА ПОДВИЖНЫХ ЦЕЛЕЙ

Работа посвящена разработке и исследованию эффективности программно-аппаратного механизма защиты информации, реализующего метод подвижных целей при помощи ГПСЧ или иных алгоритмов стохастического преобразования, в целях повышения эффективности защиты информации в системах от атак с выполнением произвольного кода. В результате работы для ПЛИС Xilinx Kintex-7 разработана интегрированная среда тестирования, включающая в себя прототип микропроцессорной системы и механизм защиты; проведено тестирование механизма посредством попытки внедрения вредоносного кода в систему.

Атаки с выполнением произвольного кода являются распространённым классом атак на компьютерные системы [1]. В подобных атаках нарушитель использует одну из присутствующих в программах уязвимостей памяти, чтобы посредством записи данных определённого вида в память нарушить целостность потока управления программы и тем самым получить возможность совершать несанкционированные операции в системе [1, 2]. Механизмы защиты от данного вида атак отличаются значительным разнообразием и многочисленностью, однако большинство из них способно обеспечить защиту лишь от частных разновидностей атак, нацеленных на отдельные уязвимости [2, 3].

Методы защиты на основе подвижных целей демонстрируют большую эффективность в обеспечении защиты от быстро изменяющихся угроз благодаря способности не только предотвращать уже выявленные атаки, но и в некоторой мере предупреждать ещё не описанные [3. 4]. Суть этих методов состоит в периодическом внесении в систему набора изменений, которые усложняют сбор информации, необходимой нарушителю для атаки [4].

В данной работе представлен программно-аппаратный механизм защиты, реализующий метод подвижных целей путём периодического стохастического преобразования адресного пространства программы и изменения представления программного кода и данных в памяти системы. С этой целью на этапе выборки и декодирования команды механизмом,

интегрированным в процессор, выполняется гаммирование указателей кода и данных в программе при помощи последовательностей, которые вырабатываются ГПСЧ или другим алгоритмом стохастического преобразования, поддерживающим работу в режиме счётчика. На очередном этапе периодического преобразования меняется начальное значение ГПСЧ или ключ алгоритма стохастического преобразования. Кроме того, для слов в памяти, относящихся к категориям данных, кода и указателей, предусмотрены различные гаммирующие последовательности, а их выбор определяется соответствующими тегами, которые должны быть предварительно сформированы в исполняемом файле программы.

Интегрированная среда тестирования реализована на языке описания аппаратуры SystemVerilog для ПЛИС Xilinx Kintex-7 и включает в себя прототип микропроцессора с модифицированным конвейером команд, поддерживающим гаммирование кода, данных и их адресов в памяти; IP-ядро оперативной памяти, а также сценарий тестирования. Тестирование микропроцессорной системы выполнено посредством попытки проведения атаки с переполнением буфера памяти на тестовую программу.

Реализованный механизм позволяет повысить эффективность защиты информации в компьютерных системах от атак с выполнением произвольного кода благодаря свойству превентивности, обеспечиваемому методом подвижных целей. Гаммирование программного кода, данных и соответствующих указателей в конвейере команд процессора, а не в оперативной памяти, позволяет добиться относительно малых накладных расходов на производительность системы и избавляет от необходимости модификации и повторной компиляции исходного кода программ, однако требует присвоения коду и данным тегов.

Список литературы

1. Cowan C. et al. Buffer overflows: Attacks and defenses for the vulnerability of the decade // Proceedings DARPA Information Survivability Conference and Exposition. DISCEX'00. – IEEE, 2000. – Т. 2. – С. 119-129.

2. Abadi M. et al. Control-flow integrity principles, implementations, and applications // ACM Transactions on Information and System Security (TISSEC). – 2009. – Т. 13. – №. 1. – С. 1–40.

3. Gallagher M. et al. Morpheus: A vulnerability-tolerant secure architecture based on ensembles of moving target defenses with churn. Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems. – 2019. – С. 469-484.

4. Jajodia S. et al. (ed.). Moving target defense: creating asymmetric uncertainty for cyber threats. Springer Science & Business Media, 2011. – Т. 54.

УДК 004.056

П.Д. ЗЛАТОВРАТСКИЙ, О.А. КОЗЛОВА

Национальный исследовательский ядерный университет «МИФИ», Москва

ИСПОЛЬЗОВАНИЕ ИНФРАСТРУКТУРЫ ОТКРЫТЫХ КЛЮЧЕЙ В ЗАДАЧАХ АВТОРИЗАЦИИ В СКУД

Системы контроля и управления доступом к физическим объектам традиционно полагаются на централизованные базы данных для управления авторизацией, которое неэффективно без стабильной связи между администратором СКУД и базой данных. В докладе исследуется управление авторизацией без постоянной связи путём использования иерархической инфраструктуры открытых ключей.

Введение

Системы контроля и управления доступом, требования к которым устанавливаются стандартом [1], рассматривают процедуру настройки правил прохода в защищаемую зону доступа как отдельную процедуру, выполняемую непосредственно на контроллере доступа или по линиям связи, к которым применяется постоянный контроль.

Стандарт [1] предполагает в первую очередь использование локальных сетей для настройки контроллера, но не устанавливает явных запретов на использование СКУД в глобальных сетях связи. Настройка посредством глобальных сетей актуальна для СКУД, географически удалённых от центра принятия решений о праве доступа, однако в этом случае нестабильность соединения может являться препятствием для эффективного управления контроллером СКУД.

Предлагаемый подход

В работе предлагается избежать проблем нестабильного соединения, используя инфраструктуру открытых ключей. Для этого линия связи рассматривается как набор сообщений, которые авторизованный администратор направляет на контроллер СКУД, выполняя добавление пользователя с заданным идентификатором в качестве обладателя права на проход через точку доступа при определённых условиях.

Использование криптографического преобразования информации с использованием открытого ключа для обнаружения возможного внесения в них изменений, а также аутентификации на основе инфраструктуры открытых ключей [2] позволяет избежать передачи ответных сообщений от контроллера СКУД администратору. Посредником в передаче

сообщений от администратора к контроллеру СКУД является авторизуемый пользователь, который взаимодействует с администратором в момент получения прав и с контроллером СКУД при аутентификации.

Можно провести параллель между такой цепочкой и хорошо известной инфраструктурой открытых ключей на основе X.509 (PKIX) в трёхзвенной реализации: корневой УЦ; УЦ; пользователь [3]. Аналогами данных элементов будут контроллер, администратор и пользователь СКУД, соответственно. При этом предлагаемая структура используется не для аутентификации, как PKIX, а для авторизации.

Перспективы развития

Непосредственное применение PKIX для авторизации не представляется оправданным: в PKIX удостоверяющий центр определяет лишь личность и связанные с ней параметры, а для СКУД администратор, занимающий схожее положение в инфраструктуре, определяет право прохода в защищаемую зону. Разнородность этих характеристик делает их совмещение в одном сообщении крайне сложным из-за комплексных требований к компетенции администратора СКУД и УЦ одновременно.

Параллель между PKIX и предлагаемым подходом позволяет предположить необходимость создания инструментов, которые в настоящее время используются PKIX, такие как: список аннулированных сертификатов, политика работы администратора [3] и другие.

Заключение

Применение инфраструктуры открытых ключей для авторизации прохода через точку доступа – перспективный способ управления правами в географически удалённых СКУД без непосредственного подключения к каналам связи. Данный подход может быть применён для авторизации прохода в тех СКУД, которые не могут быть подключены к сетям связи общего пользования по требованиям информационной безопасности.

Список литературы

1. ГОСТ Р 51241-2008. Средства и системы контроля и управления доступом. Классификация. Общие технические требования. Методы испытаний., М.: ФГУП «СТАНДАРТИНФОРМ», 2009. – 31 с.

2. Physical Access Control System (PACS) in a Federal Identity, Credentialing and Access Management (FICAM) Framework. Security Industry Association, 2013. URL: https://iqdevices.com/pdfFiles/EPTWG_White%20Paper_v5.pdf (дата обращения 22.10.2025)

3. Горбатов В.С., Полянская О.Ю. Основы технологии PKI. М.: Горячая линия – Телеком, 2004. – 248 с.

УДК 004.05

В.Д. РОЛДУГИН, Г.Г. НОВИКОВ

Национальный исследовательский ядерный университет «МИФИ», Москва

РАЗРАБОТКА МОДУЛЯ ОБРАБОТКИ ГРУППОВЫХ ЗАПРОСОВ MODBUS В СРЕДЕ OWEN LOGIC

Доклад посвящен вопросам повышения эффективности обмена данными в системах промышленной автоматизации. Рассмотрены архитектура и алгоритмы модуля групповой обработки запросов для протокола Modbus, интегрируемого в среду разработки Owen Logic. Актуальность темы обусловлена растущими требованиями к скорости и надежности обмена данными в распределенных системах управления технологическими процессами. Приведены результаты тестирования, подтверждающие повышение производительности и надежности взаимодействия с устройствами

Введение

Протокол Modbus остается одним из наиболее распространенных стандартов в промышленной автоматизации. Однако его базовая реализация, особенно в средах конфигурирования программируемых логических контроллеров (ПЛК), зачастую не использует потенциал групповых операций чтения и записи, что приводит к избыточному сетевому трафику и увеличению времени отклика системы.

Среда программирования Owen Logic [1], предназначенная для контроллеров компании «ОВЕН», предоставляет средства для создания алгоритмов управления, но обладает ограниченными возможностями по оптимизации Modbus-обмена. Разработка модуля групповой обработки запросов направлена на устранение этого недостатка.

Учет ограничений протокола Modbus

Важным аспектом при проектировании модуля являлся учет аппаратных и программных ограничений протоколов Modbus RTU [2] и ASCII. Для устройств платформы КС1 размер буфера передачи данных фиксирован и составляет 256 байт. Модуль динамически рассчитывает максимально допустимое количество регистров в одном запросе, опираясь на выбранный протокол и функцию Modbus.

Модуль рассчитывает максимальный размер группы переменных, выбирая меньшее значение между пользовательской настройкой и ограничением протокола. Если количество регистров в группе превышает

этот лимит, модуль автоматически разбивает группу на несколько отдельных запросов.

Условия и алгоритм группировки переменных

При разработке определены условия для автоматического объединения одиночных запросов к переменным Modbus [2] в групповые на основе анализа их атрибутов.

Алгоритм модуля последовательно проверяет условия для каждого набора переменных, что обеспечивает формирование только семантически корректных групповых запросов, исключающих ошибки при взаимодействии с устройствами.

Основная логика построена на последовательном анализе отсортированного списка переменных и формировании групп с проверкой условий совместимости и ограничений протокола с учетом: превышения лимита регистров, наличия разрыва в адресах, изменения функции протокола, совместимости типов данных.

Модуль тесно интегрирован с существующей архитектурой Owen Logic через специализированный сервис и взаимодействует с глобальным словарем переменных.

Тестирование и результаты

Для верификации корректности работы модуля разработан комплекс тестов, включающий: юнит-тесты для алгоритмов группировки и валидации, интеграционные тесты, проверяющие взаимодействие модуля с ядром Owen Logic, функциональные тесты на реальном оборудовании.

Результаты тестирования показали: сокращение количества Modbus-запросов на 60–80% для типовых конфигураций, корректную обработку граничных случаев, соблюдение ограничений протоколов [3], стабильную работу при длительно эксплуатации.

Список литературы

1. Иванов О.П. Основы программирования с OWEN Logic. – М.: Наука, 2020. – 280 с.
2. Петров В.К. Промышленные сети и протоколы связи в АСУ ТП. – М.: Техносфера, 2019. – 320 с.
3. Официальная документация MAP – Modbus Application Protocol. URL: https://modbus.org/docs/Modbus_Application_Protocol_V1_1b3.pdf

УДК 004.8

Ю.В. МАЛЫШЕВ

Национальный исследовательский ядерный университет «МИФИ», Москва

ОБ ИНТЕЛЛЕКТУАЛИЗАЦИИ ТРИГГЕРА ДЛЯ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ БАЗЫ ДАННЫХ

Для повышения уровня информационной безопасности в базах данных известно использование аудита на основе триггеров. Они позволяют применять сложные правила обеспечения безопасности. Это может потребовать решения задач, которые традиционно считаются интеллектуальным, и использовать подход технологии продукционных правил. Задача может решаться созданием триггера, который может собирать статистику, самообучаться, делать выводы, разрешать конфликты и изменять поведение на основе своего опыта.

Для повышения в базах данных уровня информационной безопасности среди различных мер защиты информации применяется аудит доступа к данным. Одним из основных подходов является организация аудита на основе триггеров [1]. Триггеры позволяют автоматически отвечать на определенное событие проведением контроля изменений в таблицах базы данных, проверки соблюдения сложных бизнес-правил, которую невозможно выполнить средствами ограничений целостности, и сложных правил обеспечения безопасности [2]. Указанная сложность может потребовать для обеспечения информационной безопасности решать задачи на основе представления знаний и использования самообучения.

Для решения таких задач добавляется функция, которая создает более умное программное обеспечение [3], обладающее интеллектуальными свойствами [4]. Для этого можно использовать подход технологии систем с продукционными правилами [3, 4]. Для работы триггера данные настройки могут храниться в отдельных таблицах, которые содержат экспертные знания, данные вывода и собранной статистики.

Триггерная функция при интеллектуализации триггера может быть адаптивной, изменяя свое поведение на основе полученного собственного опыта, может учиться и делать выводы. Для получения дополнительных знаний используется возможность применения триггеров для сбора статистики [2] о данных или действиях пользователей. Статистика может использоваться в посылках для разрешения конфликта.

Для применения набора правил можно использовать анонимные блоки, работающие с данными настройки в отдельных таблицах, вызовы хранимых процедур и хранимых функций для модульного

программирования обработки данных. Известно, что доступ к данным клиентских приложений через хранимые процедуры и функции при запрете прямого доступа к таблицам позволяет увеличить безопасность базы данных [2]. Для уменьшения ограничений существующих недостатков систем основанных на правилах [4] и триггеров [1, 2] известны подходы построения компактной системы [1, 4].

Возможное постоянное развитие внешней проблемной среды требует, чтобы в поддерживающих их средствах знания проблемной области могли отображаться адекватно и быть легко изменяемыми. Для этого решаются плохо формализуемые задачи, требуются адаптивность и способность к самообучению используемых программных средств. Интеллектуализация программных средств позволяет придать им адаптивные свойства. Постоянно изменяемая модель проблемной области может поддерживаться в базе знаний [4].

Экспертными данными первоначально настраиваются в правилах значения параметров, которые впоследствии могут изменяться. При малом количестве опытов в первичной статистической совокупности применяются экспертные данные. Используются сведения для формирования правил из привязанных ко времени операций и соответствующих данных и методы математической статистики. Это позволяет выполнять подсчет веса, ранжирование, профилирование (лучших) образцов, выявление исключений, устранение конфликтов [4].

В простейшем случае определяется общий вид правила из базы знаний, впоследствии изменяются входящие в него параметры. Обучение может происходить по стратегии «без учителя» или по стратегии «с учителем» при учете дальнейшей оценки успеха каждой операции. Обучение как приобретение знаний операций с базой данных можно рассматривать как подход машинного обучения.

Применение возможностей рассмотренного умного триггера может расширить возможности обеспечения информационной безопасности.

Список литературы

1. Полищук Ю.В., Боровский А.С. Базы данных и их безопасность: учебное пособие. – М.: ИНФРА-М, 2020. – 210. – (Высшее образование: Специалитет)

2. Ткачев О.А. PostgresSQL: SQL+PL/pgSQL для тех, кто хочет стать профессионалом – СПб.: Издательство Наука и Техника, 2024. – 480 с.

3. Джонс М.Т. Программирование искусственного интеллекта в приложениях. М.: ДМК Пресс, 2018. – 312 с.

4. Андрейчиков А.В., Андрейчиков О.Н. Интеллектуальные информационные системы: Учебник. – М.: Финансы и статистика, 2004. – 424 с.

УДК 004.056

Т.А. НИКИТИН[1], Т.И. КОМАРОВ[1], И.Ю. ЖУКОВ[2]

[1]Национальный исследовательский ядерный университет «МИФИ», Москва
[2]ООО «Группа компаний «Инфотактика», Москва

ВОПРОСЫ БЕЗОПАСНОСТИ CXL

CXL (Compute Express Link) – это эволюционное продолжение стандарта PCIe (Peripheral Component Interconnect Express). На физическом уровне он использует тот же интерфейс, что и PCIe, но благодаря более совершенным протоколам на верхнем уровне, CXL способен устранить некоторые существенные недостатки, присущие PCIe. Однако с новыми возможностями появляются и новые вызовы в области обеспечения безопасности и конфиденциальности данных.

Стандарт CXL был разработан для расширения возможностей PCIe. Его основная особенность заключается в том, что эта технология обеспечивает когерентность основной памяти, которая подключается непосредственно к процессору, и памяти CXL-устройств [1].

CXL состоит из трёх высокоуровневых протоколов: CXL.io (инициализация, управление и обмен данными между хостом и устройством); CXL.cache (кэширование устройством информации из памяти хоста); CXL.mem (возможность для хоста обращаться к памяти устройства как к кэшируемой).

Первые версии CXL (CXL 1.0 и CXL 1.1) предназначались для подключения устройств только к одному хосту. Более поздние версии (CXL 2.0 и CXL 3.0) позволяют создавать сложные сети из множества хостов и устройств, объединять их ресурсы, предоставлять совместный доступ и многое другое. Указанные возможности позволяют использовать технологию CXL в дата-центрах для организации внутренних сетей высокопроизводительных вычислительных систем самого различного назначения (машинное обучение и нейронные сети, в частности; моделирование; потоковая обработка данных и т.д.).

Однако возможности, добавленные в новые стандарты технологии CXL, помимо существенного расширения её функциональных возможностей, делают её подверженной тем атакам, которые хорошо известны в контексте сетевых технологий и могут быть высокоуровнево разделены на три категории: неавторизованный доступ к ресурсам и информации; неавторизованное изменение доступной в сети информации; атаки типа «отказ в обслуживании» [2].

Стандарт предлагает несколько механизмов защиты от указанных выше угроз, одним из которых является возможность сквозного шифрования трафика. Однако данные механизмы являются рекомендованными, а не обязательными, и есть риск, что производители оборудования и разработчики программного обеспечения не будут их использовать по умолчанию. Кроме того, нельзя исключать возможность появления новых аппаратных и программных ошибок, которые могут привести к появлению дополнительных уязвимостей.

Уже сейчас существует ряд исследований [3, 4, 5], направленных на создание дополнительных средств защиты сетей на основе технологии CXL, что свидетельствует о недостаточности мер по обеспечению безопасности сетевого взаимодействия, предусмотренных стандартом.

Таким образом, технология CXL является весьма востребованной для построения высокопроизводительных вычислительных систем, но на текущий момент её стандарт не требует применения изложенных в нём механизмов обеспечения безопасности. Более того, их применение не может обеспечить достаточный уровень безопасности для многих применений, т.к. исследований безопасности применения технологии CXL существует не так много, а в имеющихся отмечается недостаточность мер, предусмотренных стандартом, и предлагаются новые.

Переход на технологию CXL необходим для повышения эффективности организации вычислений, в т.ч. на объектах КИИ, но требует более внимательного анализа сопряженных с этим проблем, в т.ч. из области информационной безопасности.

Список литературы

1. Das Sharma D., Blankenship R., Berger D. An Introduction to the Compute Express Link (CXL) Interconnect // ACM Comput. Surv. 2024. т. 56. № 11. С. 1–37. DOI: http://dx.doi.org/10.1145/3669900.

2. Dastres R., Soori M. A Review in Recent Development of Network Threats and Security Measures // International Journal of Computer and Information Sciences. 2021. HAL Id: hal-03128076.

3. Stark S. W., Markettos A. T., Moore S. W. How Flexible Is CXL's Memory Protection? // Commun. ACM. 2023. т. 66. № 12. С. 46–51. DOI: http://dx.doi.org/10.1145/3617580.

4. Choi K. и др. ShieldCXL: A Practical Obliviousness Support with Sealed CXL Memory // ACM Trans. Archit. Code Optim. 2025. т. 22. № 1. с. 1–25. DOI: http://dx.doi.org/10.1145/3703354

5. Li C., Zhao J., Xu Y. Efficient Security Support for CXL Memory through Adaptive Incremental Offloaded (Re-)Encryption. // Proceedings of the 58th IEEE/ACM International Symposium on Microarchitecture (MICRO '25). ACM. 2025. с. 1102–1116. DOI: https://doi.org/10.1145/3725843.3756119.

УДК 004.056

А.А. СМИРНОВ

Национальный исследовательский ядерный университет «МИФИ», Москва

РАНЦЕВЫЕ КРИПТОСИСТЕМЫ: АНАЛИЗ СТОЙКОСТИ И ПОТЕНЦИАЛЬНЫХ УЯЗВИМОСТЕЙ, ВОЗМОЖНОСТИ ОРГАНИЗАЦИИ СКРЫТЫХ КАНАЛОВ

В работе представлен обзор ранцевых (рюкзачных) криптосистем – класса асимметричных схем шифрования. Рассмотрены различные виды ранцевых схем, варианты известных атак на них. Дополнительно обсуждаются скрытые уязвимости и возможности организации скрытых криптографических каналов (бэкдоров). Приводится программная реализация скрытого канала в ранцевой криптосистеме. Описана эволюция ранцевых схем, выявлены сильные и слабые стороны, а также сформулированы перспективные направления повышения криптостойкости и применения в постквантовой криптографии.

Ранцевые криптосистемы (КС) представляют собой класс асимметричных схем шифрования, основанных на NP-полной задаче о рюкзаке – нахождении подмножества чисел с заданной суммой [1, с. 1–7]. Они являются одними из первых практических реализаций асимметричной криптографии. Несмотря на историческую значимость, многие ранние схемы оказались уязвимы к криптоанализу [1, с. 8–17] [2]. Целью настоящей работы является анализ различных видов ранцевых КС, оценка их криптографической стойкости, выявление возможных уязвимостей и исследование возможности встраивания бэкдоров.

Для достижения цели был проведен обзор различных ранцевых схем шифрования – начиная с классической КС Меркла-Хеллмана [3] и заканчивая современными вариантами ранцевых КС [4–6], в том числе постквантовая КС Шора-Ривеста [4], основанная на арифметике в конечных полях (полях Галуа). Рассмотрены варианты классических атак. Проведен сравнительный анализ схем. Выполнен анализ скрытых уязвимостей, в том числе основанных на структуре ключей, которые могут быть использованы для встраивания бэкдоров без значительного снижения стойкости схем. Программно реализован скрытый канал в ранцевой КС [7]. Разработано приложение, демонстрирующее работу "зараженной" программы генерации ключей (рис. 1) и проведение атаки (рис. 2), результатом которой является восстановление секретного ключа пользователя.

Обзор ранцевых КС позволяет проследить эволюцию их конструкций, выявить сильные и слабые стороны рюкзачных схем шифрования, а также определить их направления развития.

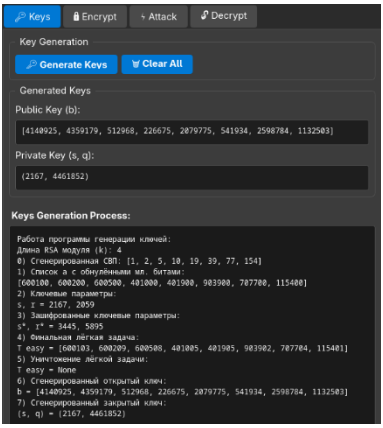

Рис. 1. Работа «зараженной» программы генерации ключей

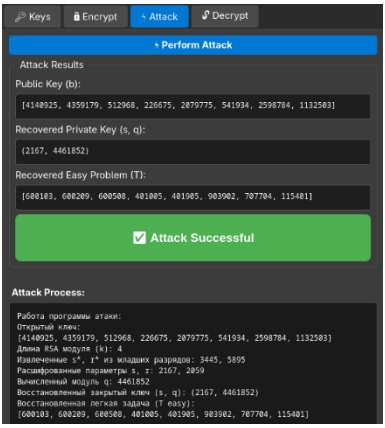

Рис. 2. Проведение атаки

Ранцевые КС более простые с точки зрения математики, в сравнение с другими КС с открытым ключом (например, КС RSA). Также современные модификации рюкзаков демонстрируют потенциал для повышения криптостойкости и применения в постквантовой криптографии [4]. При этом скрытых каналов в ранцевых КС практически неизвестно, тем не менее, существует большое количество потенциальных возможностей для создания бэкдоров [7].

Список литературы

1. Odlyzko A.M. The rise and fall of knapsack cryptosystems, 1998. DOI: 10.1090/psapm/042/1095552.

2. Schnorr C.-P. Attacking the Chor-Rivest cryptosystem by improved lattice reduction. Springer, 1995. p. 1–12.

3. Merkle R., Hellman M. Hiding information and signatures in trapdoor knapsacks // IEEE Transactions on Information Theory. 1978.

4. Chor B. and Rivest R.L. A knapsack-type public key cryptosystem based on arithmetic in finite fields. IEEE Transactions on Information Theory, v. 34, no. 5, p. 901–909, 1988. DOI: 10.1109/18.21214.

5. Kasahara M. Construction of New Classes of Knapsack Type Public Key Cryptosystem Using Uniform Secret Sequence, K (II) PKC, Constructed Based on Maximum Length Code // IACR Cryptology ePrint Archive. 2012.

6. Волков МарияСабина А., Гордеев Э.Н. Применение неинъективных векторов в ранцевых криптосистемах. Безопасность информационных технологий, [S.l.], т. 32, № 1, с. 122–131, 2025. ISSN 2074-7136. DOI: 10.26583/bit.2025.1.08.

7. Иванов М.А. Идея бэкдора в ранцевой криптосистеме. Безопасность информационных технологий, [S.l.], т. 32, № 3, с. 112–120, 2025. ISSN 2074-7136. DOI: 10.26583/bit.2025.3.09.

УДК 004.056

К.В. АГИЕВЕЦ[2], М.А. ИВАНОВ[1,2]

[1]Государственный университет управления, Москва
[2]Национальный исследовательский ядерный университет «МИФИ», Москва

ГЕНЕРАТОР ПСЕВДОСЛУЧАЙНЫХ ЧИСЕЛ, ПРЕДНАЗНАЧЕННЫЙ ДЛЯ СОЗДАНИЯ ВИРТУАЛЬНОГО КАНАЛА СВЯЗИ В СХЕМЕ СТОХАСТИЧЕСКОГО КОДИРОВАНИЯ

Рассматривается устройство, ориентированное на реализацию прямого и обратного стохастического преобразования в стохастическом (n, k, Q) кодеке. Блок прямого стохастического преобразования, реальный дискретный канал связи и блок обратного стохастического преобразования образуют Q-ичный симметричный канал, в котором в случае ошибки в канале связи вероятности различных векторов ошибок Q-ичных символов равны.

При передаче данных по каналам связи приходится решать три задачи защиты информации: обнаружение и исправление ошибок, вызванных действием помех в каналах связи (обеспечение помехозащищенности); обеспечение секретности информации; защиту от навязывания ложных данных (имитозащиту).

Известен способ обеспечения комплексной защиты информации, пересылаемой по каналу связи, основанный на использовании стохастического кодирования [1, 2]. В докладе рассматривается генератор псевдослучайных чисел (ГПСЧ), ориентированный на реализацию прямого и обратного стохастического преобразования в стохастическом (n, k, Q) кодеке, где n – число Q-ичных символов в кодовом слове, k – число информационных Q-ичных символов, (n – k) – число проверочных Q-ичных символов.

На рис. 1 показана схема ГПСЧ, где DI – информационные входы, Mode – вход режима, Clk – тактовый вход, K – ключ (параметр стохастического преобразования); Rg – регистры, Mux – мультиплексоры (два в один), блок сложения в $GF(2^{L/2})$, блок умножения на матрицу $(T_G)^{L/2}$ (T_G – матрица размером $L/2 \times L/2$, определяющая один такт работы генератора Галуа, вид которой зависит от выбранного характеристического двоичного примитивного многочлена $L/2$-й степени), блок замены S, блок сложения по модулю $2^{L/2}$; DO – информационные

выходы устройства в режиме стохастического преобразования, L – разрядность Q-ичных символов.

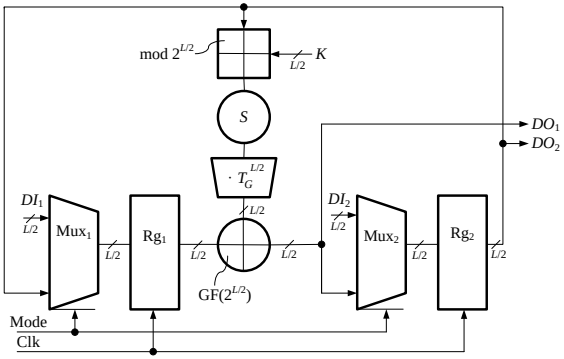

Рис. Схема ГПСЧ

В устройстве реализована сеть Фейстеля, обладающая свойством обратимости выполняемых криптографических преобразований, что дает возможность по одной и той же схеме реализовать прямое (на стороне источника информации) и обратное стохастическое преобразование (на стороне приемника информации), что является необходимым условием при реализации стохастического кодирования информации. При этом раундовая функция петли Фейстеля может быть реализована на основе 2D или 3D (в зависимости от величины L) S-блока и блока умножения на матрицу.

При $L = 32$ выбирается 2D S-блок размером 4×4; при $L = 128$ выбирается либо 2D S-блок размером 8×8, либо 3D S-блок размером $4 \times 4 \times 4$.

Список литературы

1. Алгоритмическое мышление в задаче надежной передачи данных по каналу связи. Агиевец К.В., Иванов М.А., Кондахчан М.А., Стариковский А.В. − Вторая Всероссийская научно-техническая конференция «Кибернетика и информационная безопасность» (КИБ-2024). Сборник научных трудов. – М.: НИЯУ МИФИ, 2024. – С. 92–93.

2. Ватрухин Е.М. Комплексная защита информации в каналах «земля-борт» // Вестник Концерна ВКО «Алмаз-Антей». 2020, № 4. – С. 6–14. https://doi.org/10.38013/2542-0542-2020-4-6-14

УДК 004.056

А.Р. ЛАХИН

Национальный исследовательский ядерный университет «МИФИ», Москва

СКРЫТЫЙ КРИПТОГРАФИЧЕСКИЙ КАНАЛ В ПРОГРАММЕ ГЕНЕРАЦИИ КЛЮЧЕЙ РАНЦЕВОЙ КРИПТОСИСТЕМЫ

В докладе анализируется скрытый механизм компрометации криптосистемы с открытым ключом, основанной на задаче об укладке рюкзака. Целью работы является выяснение возможностей злоумышленников по внедрению бэкдоров. Предлагается метод, позволяющий злоумышленнику, имеющему пару ключей RSA, восстановить секретные параметры жертвы. Метод основан на преобразовании супервозрастающей последовательности с добавлением «шума». Показано, что, обладая открытым ключом жертвы, злоумышленник может вычислить ее закрытый ключ и решить легкую задачу об укладке рюкзака. Результаты подчеркивают риски, связанные со скрытыми уязвимостями в криптосистемах.

Введение

В криптографии ранцевые криптосистемы известны своей сложностью и устойчивостью к классическим атакам. Однако возможность внедрения скрытых каналов с помощью бэкдоров остаётся серьёзной угрозой.

Предположим, что мы в роли «разработчика-злоумышленника», который внедряет бэкдор в ранцевую криптосистему [1, 2]. Возьмем для создания бэкдора криптоалгоритм RSA:

$$PK_W = (e, N), SK_W = d,$$

где PK_W и SK_W – соответственно открытый и закрытый ключи злоумышленника, а N имеет m цифр.

Пусть в атакуемой криптосистеме:

1) $\{a_i\}, i = \overline{0, n}$, где $n \geq 2m - 1$ – в конце каждого члена этой последовательности, определяющей легкую задачу об укладке рюкзака, присутствует член из супервозрастающей последовательности $\{2^i * 10 - 9\}$ с необходимым количеством нулей перед каждым из них, чтобы было однозначное решение задачи об укладке рюкзака. А перед этой супервозрастающей последовательностью с нулями присутствует шум:

$$a_i = * \dots * 0 \dots 0(2^i * 10 - 9),$$

где $*$ – это случайная цифра от 0 до 9;

2) R и S – взаимно простые числа (Для простоты: R и S – простые);

3) $R < N$ и $S < N$;

4) $RS \equiv 1 \bmod T$ (Для простоты возьмём: $T = RS - 1$);

5) $\sum_{i=1}^{n} a_i + 9(n + 1) < T$.

Первое, третье и пятое условия – для возможности реализации самого бэкдора, второе и четвертое – как разрешенный вариант реализации рюкзачной криптосистемы.

Идея бэкдора (с учётом заданных условий)

Со стороны криптосистемы с бэкдором:

$R^e \bmod N = R^* = r_m^* r_{m-1}^* \dots r_1^*, S^e \bmod N = S^* = s_m^* s_{m-1}^* \dots s_1^*$.

Преобразуем $\{a_0, \dots, a_n\}$ в другую последовательность $\{a_0^*, \dots, a_n^*\}$ с помощью алгоритма, схема которого представленного на рис. 1, где $\{t_i\}$ – измененная супервозрастающая последовательность, чтобы можно было получить числа R^* и S^* из открытого ключа. После этого получим открытый ключ:

$$\{a_0^* * R \bmod T, a_1^* * R \bmod T, \dots, a_n^* * R \bmod T\} = \{b_0, b_1, \dots, b_n\},$$

В итоге: (S, T) – секретный ключ жертвы, а $\{b_0, b_1, \dots, b_n\}$ – открытый.

Со стороны разработчика-злоумышленника:

Берем $\{b_0, b_1, \dots, b_n\}$ из списка открытых ключей и собираем из последних цифр первых $2m$ членов этой последовательности числа R^* и S^*.

$$(R^*)^d \bmod N = R^{ed} \bmod N = R, (S^*)^d \bmod N = S^{ed} \bmod N = S;$$

$$T = R * S - 1;$$

$$\{b_0 * S \bmod T, b_1 * S \bmod T, \dots, b_n * S \bmod T\} = \{a_0^*, a_1^*, \dots, a_n^*\}$$

В итоге, получили и секретный ключ жертвы, и оригинальную последовательность, с помощью которой можно легко решить задачу о рюкзаке.

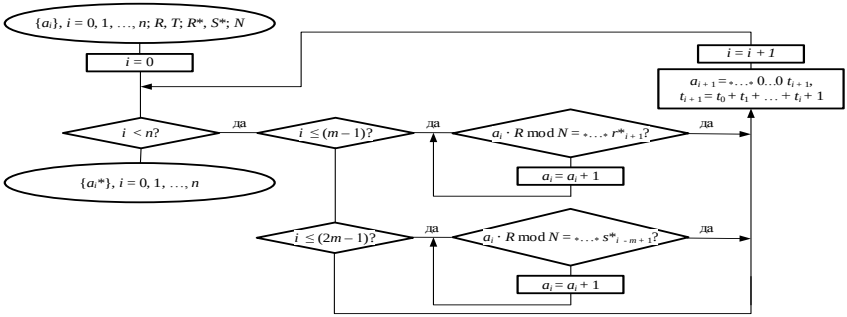

Рис. 1. Схема алгоритма преобразования последовательности $\{a_i\}$.

Список литературы

1. Arto Salomaa. Public-Key Cryptography. URL: http://www.bnti.ru/dbtexts/ipks/old/analmat/1_2002/open_key_crypt.pdf (дата обращения: 15.10.2025).

2. Иванов М.А. Идея бэкдора в ранцевой криптосистеме // Безопасность информационных технологий, 2025, v. 32, № 3. − С. 112–120.

УДК 004.056

Т.С. КАЛИНИНА, А.Р. ЛАХИН

Национальный исследовательский ядерный университет «МИФИ», Москва

РЮКЗАЧНАЯ КРИПТОСИСТЕМА, ИСПОЛЬЗУЮЩАЯ МАТЕМАТИКУ КОНЕЧНЫХ ПОЛЕЙ

Предлагается модификация рюкзачной криптосистемы на основе полей Галуа $GF(p\hat{}n)$. Цель – создание стойкой схемы, устойчивой к классическим атакам. Разработан метод построения последовательностей с уникальными суммами через полиномиальное представление элементов поля. Ключевой особенностью является механизм шифрования с добавлением шума и умножением на примитивный элемент поля. Представлены корректные процедуры шифрования и расшифрования. Результаты подтверждают перспективность предлагаемого решения.

Введение

В отличии от ранцевой криптосистемы Шора-Ривеста, где используется дискретное логарифмирование, предлагается модификация, работающая непосредственно в конечном поле.

В рюкзачных криптосистемах осуществляется нахождение такого набора чисел $\{a_{i_1}, a_{i_2}, \dots, a_{i_m}\}$ из последовательности $\{a_1, \dots, a_n\}$, сумма которых даёт k. Перенесём эту идею в расширенные поля Галуа $\mathrm{GF}(p^n)$.

Элементами поля $\mathrm{GF}(p^n)$ являются полиномы степени, меньшей n, с коэффициентами из конечного поля $\mathrm{GF}(p)$: $a_{k1}a_{k2}\dots a_{kn}$, где $a_{ki} \in \mathrm{GF}(p) = \{0, 1, \dots, p-1\}, i = \overline{1,n}$.

Сумма любых элементов из поля $GF(p^n)$ даёт элемент из того же поля. Но, у нас есть условие для рюкзачной криптосистемы, что любой набор чисел из последовательности должен давать уникальную сумму, иначе сообщение будет невозможно расшифровать. Не все возможные последовательности из элементов поля $\mathrm{GF}(p^n)$ будут удовлетворять такому условию, также максимальное количество элементов, входящих в удовлетворяющую условию последовательность, равно n. Чтобы определить, удовлетворяет ли последовательность длиною n такому условию, найдём её определитель по модулю p:

$$\{a_1, \dots, a_n\}, a_i = a_{i1}a_{i2}\dots a_{in}, i = \overline{1,n}$$

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \bmod p = \begin{cases} 0 \Rightarrow \text{не удовлетворяет} \\ \{1, \dots, p-1\} \Rightarrow \text{удовлетворяет.} \end{cases}$$

При этом если брать подпоследовательности из этих последовательностей, то они тоже будут удовлетворять условию уникальности.

Идея

Имеем поле $\mathrm{GF}(p^n)$, неприводимый многочлен $\varphi(x)$ степени n, примитивный элемент ω, число R $(0 < R < p^n - 1)$, последовательность $\{a_1, \dots, a_k\}$ $(a_i \in \mathrm{GF}(p^n),\ i = \overline{1,k}, k \le n)$.

Сформируем открытый ключ:

1) Добавим шум: $b_i = {*}{*}{*} \dots {*}{*}{*}\ a_i\ {*}{*}{*} \dots {*}{*}{*} = b_{i1} b_{i2} \dots b_{il}$, при этом $l \bmod n = 0, i = \overline{1,k}$.

2) Делим b_i на блоки длиной n и умножаем их на ω^R и берем по $mod\ \varphi(x)$

$$b_i = b_{i_1} \dots b_{i_a} \dots b_{i_s}, \text{где } b_{i_j} \text{ длиною } n, b_{i_a} = a_i, i = \overline{1,k}$$

$$c_i = \left(b_{i_1} * \omega^R \bmod \varphi(x)\right) \dots \left(b_{i_a} * \omega^R \bmod \varphi(x)\right) \dots \left(b_{i_s} * \omega^R \bmod \varphi(x)\right) = $$

$$= c_{i_1} \dots c_{i_a} \dots c_{i_s}, i = \overline{1,k}$$

В итоге: $PK = (p, \{c_i\}), SK = (n, \varphi(x), \omega, R)$.

Шифрование

Зашифрование

Имеется сообщение $m = m_1 m_2 \dots m_k$, где $m_i \in \{0, 1, \dots, p-1\}, i = \overline{1,k}$.

1) Находим сумму $e = m_1 c_1 + m_2 c_2 + \dots + m_k c_k$, причём

$$c_i + c_j = c_{i1} \dots c_{il} + c_{j1} \dots c_{jl} = \left((c_{i1} + c_{j1}) \bmod p\right) \dots \left((c_{il} + c_{jl}) \bmod p\right)$$

$$k * c_i = k * c_{i1} \dots c_{il} = (k * c_{i1} \bmod p) \dots (k * c_{il} \bmod p)$$

2) Отправляем эту сумму получателю

Расшифрование

1) Делим сумму e на блоки длиной n и умножаем их на $\omega^{p^n - R - 1}$ и берем по $\bmod \varphi(x)$:

$$e = e_1 \dots e_s$$

$$d = \left(e_1 * \omega^{p^n - R - 1} \bmod \varphi(x)\right) \dots \left(e_s * \omega^{p^n - R - 1} \bmod \varphi(x)\right) = d_1 \dots d_s$$

2) Делим c_i на блоки длиной n и умножаем их на $\omega^{p^n - R - 1}$ и берем по $\bmod \varphi(x)$:

$$c_i = c_{i_1} \dots c_{i_s}$$

$$b_i = \left(c_{i_1} * \omega^{p^n - R - 1} \bmod \varphi(x)\right) \dots \left(c_{i_s} * \omega^{p^n - R - 1} \bmod \varphi(x)\right) = b_{i_1} \dots b_{i_s},$$

$$i = \overline{1,k}$$

$$b_i = {*}{*}{*} \dots {*}{*}{*}\ a_i\ {*}{*}{*} \dots {*}{*}{*}, i = \overline{1,k}$$

3) Решаем задачу о рюкзаке, зная d_{i_a} и $\{a_i\}, i = \overline{1,k}$, и получим m.

УДК 004.056:372.8

А.К. ЕГОЯН[1], М.А. ИВАНОВ[2,3]
[1]*Государственный университет просвещения, Москва*
[2]*Государственный университет управления, Москва*
[3]*Национальный исследовательский ядерный университет «МИФИ», Москва*

ОПЫТ ПРЕПОДАВАНИЯ АЛГОРИТМИЧЕСКОГО МЫШЛЕНИЯ В УЧРЕЖДЕНИЯХ ОСНОВНОГО ОБЩЕГО И ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

Для поиска решения сложных задач применяется алгоритмическое мышление, при этом процесс решения начинается с формулировки идеального конечного результата (ИКР). Затем начинается процесс движения к ИКР. Сократить длительность этого инкубационного периода позволяют эвристические приемы разрешения технических противоречий и метод контрольных вопросов.

Рассматриваются примеры использования алгоритмического мышления (АМ), эвристических приемов разрешения технических противоречий и метода контрольных вопросов при решении различных задач защиты информации (ЗИ), как новых, так и уже решенных ранее.

Приводятся результаты применения АМ при решении новых задач ЗИ. Например, на рис. 1 приведен ряд новых схем генераторов псевдослучайных чисел (ГПСЧ), созданных с использованием АМ и ориентированных на реализацию стохастических методов ЗИ. Приводятся результаты применения АМ при совершенствовании известных перспективных технических решений. Например, в [2] продемонстрировано применение АМ при построении стохастического кодека, обеспечивающего комплексную защиту информации, пересылаемой по каналу связи. Решение «уже решенных» задач в педагогике необходимо для развития творческого мышления учащихся, отработки ими навыков АМ. Повторение процесса решения известных задач на основе использования нового опыта и новых знаний часто позволяет найти иные, более эффективные способы их решения.

В заключении показано, что алгоритмическое мышление – это универсальный подход к решению задач, который может использоваться не только в технических университетах, но и в учреждениях основного общего образования, например, для повышения эффективности процесса обучения иностранному языку [3]. Сформулируем ИКР: надо научить мыслить на другом языке. Как этого добиться? Представляется, что только так: осваивая эвристические приемы разрешения технических

противоречий и преодолевая инерцию мышления на другом языке, осваивая метод контрольных вопросов и изучая ошибки (ловушки) мышления на другом языке, решая на уроках иностранного языка задачи на развитие латерального мышления [4].

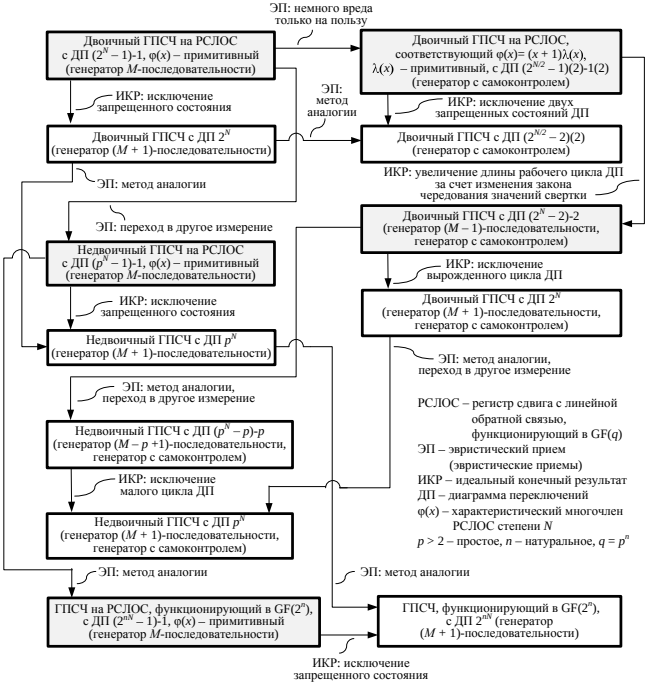

Рис. 1. Последовательность совершенствования ГПСЧ на РСЛОС. Серым цветом выделены известные технические решения.

Список литературы

1. Иванов М.А. Алгоритмическое мышление в задачах защиты информации. − Вторая Всероссийская научно-техническая конференция «Кибернетика и информационная безопасность» (КИБ-2024). Сборник научных трудов. – М.: НИЯУ МИФИ, 2024. – С. 94–95.

2. Алгоритмическое мышление в задаче надежной передачи данных по каналу связи. Агиевец К.В., Иванов М.А., Кондахчан М.А., Стариковский А.В. − Вторая Всероссийская научно-техническая конференция «Кибернетика и информационная безопасность» (КИБ-2024). Сборник научных трудов. – М.: НИЯУ МИФИ, 2024. – С. 92–93.

3. Егоян А.К. Методический потенциал использования спортивной лексики при обучении иностранному языку в младшей школе (рукопись). – М., 2025. – 72 с.

4. Moore Gareth. Lateral logic: puzzle your way to smart thinking. – London: Michael O'Mara Books Limited, 2016. – 192 p.

УДК 004.056.53

Х.Н. КАНАЕВ, Л.К. КАЧАЕВА

МИРЭА – Российский технологический университет, Москва

СОВРЕМЕННЫЕ ПУТИ РЕШЕНИЯ ПРОБЛЕМ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ КОНТЕЙНЕРНОЙ ВИРТУАЛИЗАЦИИ НА ПРИМЕРЕ DOCKER

В статье рассматриваются современные проблемы безопасности контейнерной виртуализации на примере Docker, включая уязвимости образов и зависимостей, риски выхода из контейнера, избыточные привилегии, пробелы в управлении секретами и атаки на цепочку поставок ПО; предлагается комплекс мер: «shift-left» практика в CI/CD, автоматическое сканирование, генерация и проверка SBOM, аттестации BuildKit в соответствии с SLSA, минимизация привилегий и строгие сетевые политики, что снижает вероятность компрометации инфраструктуры при сохранении преимуществ контейнерной модели [1, 4].

Введение

Технология Docker не новая, но оказалась достаточной уязвимой с позиции информационной безопасности в данной статье приводятся основные проблемы, которые необходимо решить, чтобы сделать ее более надежной, потому что ее использование в настоящий момент не является безопасным и повсеместно ее использовать нельзя [1]. Во-первых, специфика контейнеров заключается в разделении одного ядра между множеством изолированных сред, что делает критичной правильную конфигурацию механизмов изоляции процессов, файловых систем и сети; любая ошибка в настройках или использование образов с неизвестным происхождением повышает вероятность выхода за пределы контейнера и несанкционированного доступа к ресурсам хоста [2]. Во-вторых, распространённая практика повторного использования публичных базовых образов и многоуровневых зависимостей приводит к переносу известных уязвимостей в продуктивную среду, а отсутствие прозрачности состава образов затрудняет оперативное реагирование на новые CVE и устранение технического долга безопасности [1, 3]. В-третьих, уязвимости контейнерного рантайма и инструментов сборки подтверждают, что при небезопасной конфигурации возможны атаки с выходом из контейнера и воздействием на хостовую систему, что исключает возможность безусловного доверия к контейнерной изоляции как таковой [2].

Кроме технологических факторов существенную роль играет цепочка поставок программного обеспечения: компрометация зависимостей, подмена артефактов при сборке и публикации образов, а также отсутствие криптографической верификации и аттестаций приводят к тому, что даже формально «чистые» образы могут содержать скрытые закладки или уязвимые

компоненты [3, 4]. Для снижения этих рисков современный подход предусматривает смещение контроля безопасности влево, то есть интеграцию проверок на этапах разработки и сборки: автоматическое сканирование Dockerfile и образов, генерацию перечня компонентов программного обеспечения (SBOM) и закрепление происхождения сборок с помощью аттестаций, что обеспечивает прослеживаемость и возможность формировать политики допуска только проверенных артефактов [4]. Наконец, операционная дисциплина и человеческий фактор оказывают прямое влияние на итоговую безопасность: отказ от привилегированного режима, минимизация набора способностей ядра, запуск процессов от непривилегированных пользователей, строгие сетевые политики и централизованный мониторинг событий контейнеров являются обязательными условиями для перехода к контролируемой и управляемой модели эксплуатации Docker [1].

Заключение

Суммируя выявленные проблемы, представляется целесообразным внедрить единый «минимально доверенный конвейер» для Docker, объединяющий три взаимосвязанных компонента: минимизацию привилегий исполнения, проверяемую цепочку поставок артефактов и автоматизированный контроль допуска на этапе развёртывания [3, 4]. На практике это означает, что контейнеры запускаются от непривилегированного пользователя с запретом выдачи новых привилегий и только строго необходимыми Linux-capabilities; сборка образов выполняется многоэтапно с обязательной генерацией перечня компонентов (SBOM) и аттестацией происхождения; публикация сопровождается криптографической подписью, а система доставки допускает к деплою лишь образы, подпись и аттестации которых успешно проверены политиками [4]. Дополнительно для каждого сервиса применяются базовые сетевые политики с минимальными правилами входа/выхода и профили изоляции ядра (seccomp/AppArmor) из набора по умолчанию [1]. Такой связный подход одновременно уменьшает поверхность атаки за счёт отказа от избыточных прав, предотвращает попадание в эксплуатацию недоверенных сборок и ограничивает последствия потенциального компрометационного инцидента, обеспечивая контролируемую и воспроизводимую безопасность контейнерной инфраструктуры без отказа от её производственных преимуществ [3, 4].

Список литературы

1. Habr Selectel. Обзор способов защиты контейнеров Docker [Электронный ресурс]. 2024. URL: habr.com/ru/companies/selectel/articles/854850/habr.

2. opennet.ru. Уязвимость в runc, позволяющая выбраться из контейнера (CVE-2024-21626) [Электронный ресурс]. 2024. URL: www.opennet.ru/605453.

3. H3llo Cloud. Что вам надо знать в 2025 году про контейнеры, чтобы… [Электронный ресурс]. 2025. URL: habr.com/ru/companies/h3llo_cloud/articles/890512/habr.

4. Docker Docs. SBOM attestations – документация BuildKit [Электронный ресурс]. 2025. URL: docs.docker.com/build/metadata/attestations/sbom/.

КИБ-2025
КИБЕРНЕТИКА
И ИНФОРМАЦИОННАЯ
БЕЗОПАСНОСТЬ

Направление

Интеллектуальное управление сетевой безопасностью

Руководитель секции – МИЛОСЛАВСКАЯ Н.Г.,
д.т.н., профессор

УДК 004.056

Н.Г. МИЛОСЛАВСКАЯ

Национальный исследовательский ядерный университет «МИФИ», Москва

СОВРЕМЕННЫЕ СРЕДСТВА ЗАЩИТЫ ИНФОРМАЦИИ В ЦЕНТРАХ УПРАВЛЕНИЯ СЕТЕВОЙ БЕЗОПАСНОСТЬЮ

С целью определения необходимого и достаточного набора современных средств защиты информации (СЗИ), которые позволяют центрам управления сетевой безопасностью (ЦУСБ) эффективно решать функциональные задачи по обеспечению информационной безопасности (ИБ) активов информационно-телекоммуникационных сетей (ИТКС), проведено аналитическое исследование и изучены лучшие в данной области мировые практики функционирования ЦУСБ.

Введение

ЦУСБ является специализированным структурным элементом современных ИТКС различных организаций [1]. Он призван осуществлять упреждающее управление сетевой безопасностью (СБ) ИТКС за счет прогнозирования развития событий ИБ на основе интеллектуальных подходов к обработке больших относящихся к ИБ ИТКС данных и реагирования на инциденты ИБ за счет применения передовых средств защиты информации (СЗИ). В [1] предложена зональная архитектура ЦУСБ, в каждой зоне (подзоне) которой должны размещаться соответствующие СЗИ.

Постановка задачи

Целью исследования является определение необходимого и достаточного набора современных СЗИ, которые позволяют ЦУСБ эффективно решать функциональные задачи по обеспечению ИБ активов ИТКС.

Результаты исследования

Еще в 2021 г. компания *Gartner* выделила передовые СЗИ, используемых в ЦУСБ. Это системы, предназначенные для обнаружения и реагирования на угрозы для конечных точек (*EDR*) и в сети (*NDR*), управления информацией и событиями безопасности (*SIEM*), анализа поведения пользователей и их идентичности (*UEBA*), аналитики угроз (*TI*), защиты веб-приложений (*SWG*), а также облачные службы безопасного доступа (*SASE*) и брокер безопасного доступа к облаку (*CASB*) [2]. Так, в ЦУСБ для обнаружения вредоносного ПО используются

системы *EDR, CASB* и *SWG/SASE*, обнаружения нарушений политик ИБ – *SIEM* и *CASB*, обнаружения фишинга – *UEBA* и *SWG/SASE*, взлома учетных записей пользователей и злоупотребления ими – *UEBA* и *CASB*, обнаружения неприемлемого использования ПО – *SWG/SASE*, выполнения контекстного поиска – *CASB* и *TI*, проверки оповещений безопасности – *SIEM* и т.д. В 2024 г. институт *SANS* выявил наибольшую удовлетворенность от применения современных СЗИ самих работников ЦУСБ от использования на уровне хостов *EDR*, средств защиты от вредоносного ПО (*MPS*) и средств непрерывного мониторинга и оценки, а на уровне сети – виртуальных частных сетей (*VPN*), средств защиты электронной почты (*SEG*), межсетевых экранов нового поколения (*NGFW*), сетевых систем обнаружения/предотвращения вторжений (*NIDS/NIPS*), средств мониторинга трафика и средств защиты от атак «отказ в обслуживании». В качестве аналитических средств лучше всего зарекомендовали себя *SIEM*-системы, средства управления «поверхностью» атак и платформы *TI*. Также важны средства мониторинга и логирования событий операционных систем и управления журналами регистрации событий, имеющих отношение к ИБ ИТКС [3]. Аналитическое исследование *SANS* 2025 г. передовых технологий обеспечения СБ в ЦУСБ в США, Европе, Латинской и Южной Америках и Канаде показало, что различные по области деятельности, размеру и бюджету организации отдают предпочтение собственным центрам или используют «облачные» сервисы ЦУСБ [4].

Заключение

На основе проведенного исследования сформировано подтверждаемое практикой представление о составе СЗИ, необходимых для работы ЦУСБ.

Список литературы

1. Милославская Н.Г. Научные основы построения центров управления сетевой безопасностью в информационно-телекоммуникационных сетях. М.: Горячая Линия-Телеком, 2021. – 431 с.

2. Ahlm E. (Gartner). How to Build and Operate a Modern Security Operations Center. 7 June 2021. [Электронный ресурс]. – Режим доступа: https://emt.gartnerweb.com/ngw/eventassets/en/ conferences/2022/sec23i/documents/how-to-build-and-operate-a-modern-security-operations-center.pdf (дата обращения: 01.10.2025).

3. SANS 2024 SOC Survey: Facing Top Challenges in Security Operations [Электронный ресурс]. – Режим доступа: https://www.sans.org/white-papers/sans-2024-soc-survey-facing-top-challenges-security-operations (дата обращения: 01.10.2025).

4. SANS 2025 SOC Survey [Электронный ресурс]. – Режим доступа: https://www.sans.org/white-papers/sans-2025-soc-survey (дата обращения: 01.10.2025).

УДК 004.000

Д.В. ШУЛИНИН[1], Н.Г. МИЛОСЛАВСКАЯ[2]

[1]*ООО «Юзергейт», Москва*

[2]*Национальный исследовательский ядерный университет «МИФИ», Москва*

ПРОЦЕСС РАЗРАБОТКИ ПРАВИЛ ДЛЯ СРЕДСТВ ЗАЩИТЫ ИНФОРМАЦИИ

Цель работы – создание процесса разработки контента (правил выявления атак, сигнатур сетевых приложений) для средств защиты информации (СЗИ) на примере UserGate NGFW [1]. Основной результат – формирование систематизированного подхода к разработке контента СЗИ, начиная со сбора данных из открытых и закрытых источников, их оценку и приоритезацию, и заканчивая формированием пакета контента и загрузки его на СЗИ.

Введение

Защитный контент (правила выявления сетевых атак, сигнатуры сетевых приложений, идентификаторы компрометации (списки URL, IP, хэшей файлов, содержащих вредоносный код)) является неотъемлемой частью любого сетевого СЗИ. Благодаря наличию этого контента и функционала, способного его обрабатывать, сетевое устройство может называться СЗИ. Также необходимо помнить, что технологии проведения атак эволюционируют, что говорит о необходимости постоянного развития защитного контента и поддержания его в актуальном состоянии.

Постановка задачи

Цель работы – создание унифицированного процесса по разработке защитного контента для сетевых СЗИ, который бы охватывал полный цикл создания нового правила или сигнатуры, включая поиск информации о новых видах атак, оценку критичности и приоритезацию, моделирование атаки, разработку правила и тестирование его на продуктивном трафике, фильтрацию ложноположительных срабатываний и формирование пакета правил для загрузки на СЗИ.

Пути решения задачи

Разработан процесс, который трансформирует данные из открытых и закрытых источников в пакеты защитного контента, загружаемого на СЗИ, состоящий из следующих этапов:

1) формирование запроса на разработку экземпляра защитного контента. В ходе формирования запроса происходит сбор ключевой информации из

источников данных об атаках, методах защиты, уязвимостях программного обеспечения (ПО) включая, но не ограничиваясь, публикациями производителей ПО и компаний, занимающихся информационной безопасностью (ИБ), общедоступными репозитории, и информационными рассылками регуляторов в сфере ИБ [2];

2) оценка приоритета, для чего используется методика, опирающаяся на ряд критериев, в том числе вектор атаки; сложность реализации; необходимые привилегии; уровень влияния на конфиденциальность, целостность и доступность;

3) стендирование с задачей воспроизведения атаки в лабораторных условиях. Результат данного этапа – полностью сформированная лабораторная среда, которая позволила воспроизвести атаку;

4) сбор и анализ ключевых идентификаторов атаки, в рамках чего осуществляется запись трафика, журналов событий и сбор сопутствующих идентификаторов (например, файлы) в момент проведения атаки;

5) разработка и тестирование с написанием правила выявления атаки, основываясь на собранных на предыдущем этапе, образцах. Для тестирования разработанного правила трафик, имитирующий действия атакующего, пропускается через СЗИ, на котором включено разработанное правило;

6) фильтрация ложноположительных срабатываний правила, для чего используется генератор трафика с образцами трафика, имитирующими работу сети компании. Успешным завершением этапа является отсутствие ложноположительных сработок на тестовом образце трафика;

7) сборка пакета контента. Правило добавляется в пакет контента для централизованной публикации на СЗИ.

Заключение

Созданный процесс разработки контента для СЗИ охватывает полный цикл создания новых правил выявления атак, делая его прозрачным, и измеримым за счет отслеживания ключевых показателей, таких как кол-во производимого контента и время, затраченное на разные этапы разработки.

Список литературы

1. Российские межсетевые экраны нового поколения [Электронный документ]. – Режим доступа: https://www.anti-malware.ru/analytics/Technology_Analysis/Russian-NGFW-selection-criteria (дата обращения: 23.10.2025).

2. Банк данных угроз безопасности информации [Электронный документ]. – Режим доступа: https://bdu.fstec.ru/vul (дата обращения: 23.10.2025).

УДК 004.056

В.Е. МОРОЗОВ

ООО «Либрасофт», Минск, Республика Беларусь

ПРЕИМУЩЕСТВА СОВМЕСТНОГО ИСПОЛЬЗОВАНИЯ DCAP- И DLP-СИСТЕМ

Рассматривается назначение DCAP-систем, указываются риски информационной безопасности, которые с их помощью можно минимизировать. Обсуждаются достоинства и недостатки автономного использования DCAP-решения. Показываются преимущества интеграции функционала DCAP и DLP-систем. Рассматривается пример интеграции DCAP-системы «FileAuditor СёрчИнформ» и DLP-системы «КИБ СёрчИнформ» в рамках единого программного комплекса.

Решения класса DCAP (data-centric audit and protection) предназначены для обнаружения, категоризации и защиты т.н. «данных в покое». На практике за этим скрываются информация на компьютерах сотрудников, а также та, что разбросана по сетевым папкам, облачным хранилищам, базам данных и т.д. В реальной ситуации критичные для бизнеса данные часто не содержатся в единой базе, различным образом видоизменяются, сохраняются пользователями на персональных устройствах, в облаках и в общедоступных папках. В результате персональные, платежные, учетные данные, файлы с коммерческой тайной, чертежи и прочие технические документы могут бесконтрольно множиться, что создает как бизнес-риски, так и риски информационной безопасности. При этом доступ к данным получают те, кто не должен; информацию сложнее защитить от слива, так как непонятно, где она находится; привилегированных пользователей становится слишком много, их действия невозможно проконтролировать; сотрудники работают с устаревшими версиями документов.

DCAP-система как самостоятельный продукт предназначена для аудита данных внутри корпоративного периметра. Она находит и классифицирует данные из разных источников, обеспечивает мониторинг прав доступа, ведет историю событий, следит за соблюдением политик безопасности. Ее возможности гораздо шире, чем возможности просто ИТ-инструмента, который поможет упорядочить хранение данных. Тем не менее, активное использование DCAP-систем службами информационной безопасности выявило недостатки автономной работы и, прежде всего, ограниченность получаемой информации. DCAP-решение дает точную статичную картинку, его работа на продолжительном временном отрезке не может

ответить на важные для ИБ-специалистов вопросы, часто возникающие при расследовании инцидентов. Например, откуда взялся этот файл в корпоративной сети? Каким образом он попал к другому сотруднику? Мог ли его получить кто-то извне? DCAP не контролирует каналы передачи информации и дает только ограниченное видение, а чтобы его расширить, необходимо подключать другие средства корпоративной безопасности.

Интеграция DLP и DCAP, напротив, позволяет охватить все имеющиеся данные, выявлять события, происходящие внутри сети, собирать больше сведений о пользователях, их правах доступа, читать черновики почты, определять классы и атрибуты данных и многое другое.

Кроме того, слияние DLP и DCAP позволяет использовать единые политики и бесшовную защиту; общую консоль и подходы к управлению данными (при использовании продуктов от одного вендора); взаимное обогащение данных для получения полной картины и точного прогнозирования; единое хранилище данных и событий, которое позволяет сократить расходы на его внедрение и обслуживание.

В качестве примера рассмотрим интеграцию DCAP-системы «FileAuditor СёрчИнформ» и DLP-системы «КИБ СёрчИнформ» в рамках единого программного комплекса [1]. FileAuditor позволяет отслеживать наличие критичной информации на компьютерах пользователей и права доступа к файлам. Сканирование ресурсов согласно настроенным правилам, предоставление дерева папок/файлов, а также прав доступа к ним может осуществляться либо с помощью агента, либо с помощью службы анализа данных на сервере. В свою очередь, метки, установленные на файлы в результате сканирования, могут использоваться КИБ для блокировки доступа или отправки файлов через различные каналы передачи данных.

В заключение следует отметить, что если говорить о выстраивании комплексной системы информационной безопасности [2], то никакого смысла отделять DCAP от DLP нет – информация, которую собирают оба продукта, взаимно дополняет друг друга. Кроме того, использование DLP и DCAP от одного вендора позволяет сократить потребление системных ресурсов, упростить и удешевить эксплуатацию системы.

Список литературы

1. СёрчИнформ FileAuditor. – Текст: электронный // SearchInform: [сайт]. – 2025. – URL: https://searchinform.ru/products/fileauditor/ (дата обращения: 03.09.25)

2. Морозов В.Е., Милославская Н.Г. Комплексные решения для минимизации внутренних угроз информационной безопасности // Вопросы кибербезопасности. 2024. № 5(63). С. 67–78. DOI: 10.21681/2311-3456-2024-5-67-78.

УДК 004.056

М.А. КАЗАКОВ, Н.Г. МИЛОСЛАВСКАЯ

Национальный исследовательский ядерный университет «МИФИ», Москва

РЕКОМЕНДАЦИИ ПО ВНЕДРЕНИЮ КОНЦЕПЦИИ НУЛЕВОГО ДОВЕРИЯ ДЛЯ ПОВЫШЕНИЯ БЕЗОПАСНОСТИ АСУ ТП

Исследование направлено на разработку практических рекомендаций по интеграции концепции доступа с нулевым доверием (*ZTA*) в системы автоматизированного управления технологическими процессами (АСУ ТП). На основе анализа модели *Purdue* и ключевых принципов *ZTA* предложен поэтапный подход к внедрению, выделены защищаемые объекты и определены приоритетные меры для снижения рисков ИБ в промышленных сетях.

Введение

АСУ ТП становятся целью для сложных кибератак, а традиционных периметровых средств защиты уже недостаточно. Концепция доступа с нулевым доверием (*Zero Trust Access*), предполагающая постоянную проверку всех субъектов доступа, представляет собой перспективный подход.

Постановка задачи

Цель работы – разработать структурированные рекомендации по адаптации концепции *ZTA* для АСУ ТП на основе модели *Purdue* [1]. Данная модель представляет собой иерархию уровней от датчиков и контроллеров (*PLC*) до корпоративных систем, обеспечивающую сегментацию сети, что является основой для построения защищенных доменов доверия в *ZTA*. Задача исследования заключается в устранении противоречия между требованием максимальной защищенности АСУ ТП от современных киберугроз [2] и необходимостью сохранения бесперебойного функционирования технологических процессов. Критичными активами, определяющими эту задачу, являются *PLC* (программируемые контроллеры, напрямую управляющие оборудованием), *SCADA* (системы диспетчеризации) и *MES* (системы управления производственными операциями), защита которых требует особого подхода. Необходимо разработать адаптированный для АСУ ТП план внедрения концепции *ZTA*, который бы учитывал иерархическую структуру сети на основе модели *Purdue*, критичность активов и

минимизировал риски нарушения производственного цикла на этапе трансформации системы обеспечения безопасности [3].

Результаты

В результате исследования разработан пошаговый план внедрения концепции *ZTA* в АСУ ТП. Основная идея – начинать с небольших некритических систем, постепенно расширяя область защиты на всё более критичные компоненты. Рекомендуется чётко разделить сеть на изолированные зоны в соответствии с моделью *Purdue*, чтобы ограничить перемещение злоумышленников. Для постоянного контроля безопасности предлагается настроить мониторинг ключевых активов (*PLC, SCADA*-серверов) и строго ограничить права доступа пользователей только необходимыми для их задач функциями и критическим системам.

Заключение

Разработанные рекомендации позволяют системно подойти к внедрению концепции *ZTA* в АСУ ТП. Использование модели *Purdue* в качестве каркаса для *ZTA* обеспечивает логическую сегментацию сети и помогает идентифицировать ключевые активы для защиты. Предложенный поэтапный подход минимизирует риски нарушения технологических процессов при переходе на новую модель безопасности, что является критически важным для промышленных предприятий.

Список литературы

1. Zero Trust Architecture [Электронный ресурс]: NIST Special Publication 800-207 / National Institute of Standards and Technology. 2020. URL: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf (дата обращения: 24.10.2025).

2. ISA/IEC 62443 Series. Industrial communication networks – Network and system security [Электронный ресурс] / International Society of Automation. URL: https://www.isa.org/standards-and-publications/isa-standards/isa-iec-62443-series-of-standards (дата обращения: 24.10.2025).

3. Милославская, Н. Г. Управление информационной безопасностью: Конспект лекций: учебное пособие / Н.Г. Милославская, А.И. Толстой. – М.: НИЯУ МИФИ, 2020. – 536 с.

УДК 004.056.5

В.А. АРТАМОНОВ[1], В.Ю. ЖИВЦОВ[2]

[1]Самарский государственный экономический университет
[2]Самарский государственный медицинский университет

ОБЕСПЕЧЕНИЕ ДОВЕРИЯ ПРИ ДЕЦЕНТРАЛИЗОВАННОМ ВЗАИМОДЕЙСТВИИ АВТОНОМНЫХ АГЕНТОВ

В работе рассматриваются механизмы формирования доверия между автономными агентами в распределённых системах на основе технологий блокчейна и децентрализованных идентификаторов. Предложен подход к интеграции различных инструментов, включающих смарт-контракты и токенизацию для формирования неизменяемой репутационной истории. Рассматривается комплексная защита архитектуры доверия от ключевых векторов атак.

В распределённых системах взаимодействие автономных агентов с искусственным интеллектом требует надёжных механизмов подтверждения идентичности и оценки добросовестности. Обеспечение доверия базируется на совокупности криптографических методов и организационных мер. В целях удостоверения личности, учёта репутации, и защиты от манипулятивных стратегий предлагается использование технологии блокчейна.

Ключевым элементом архитектуры доверия выступают верифицируемые и независимые от посредников децентрализованные идентификаторы (DID), проверка подлинности которых осуществляется через криптографические механизмы консенсуса сети, отраслевые стандарты W3C, в свою очередь, гарантируют совместимость между платформами [1, 2].

Интеграция экономических функций в архитектуру автономных агентов достигается путём привязки цифровых активов к блокчейн-адресам, что позволяет агентам самостоятельно осуществлять финансовые операции и выступать полноправными участниками договорных отношений. Для обеспечения автоматизированного исполнения соглашений и гарантии выполнения обязательств служат смарт-контракты, исполняемые в детерминированной среде распределённого реестра без возможности одностороннего изменения условий. Токенизация цифровых и материальных объектов через невзаимозаменяемые токены (NFT) создаёт надёжный механизм подтверждения и передачи прав собственности. Децентрализованные финансовые протоколы (DeFi) предоставляют агентам доступ к

автоматизированным финансовым инструментам для формирования основы автономной цифровой экономики с минимальным человеческим посредничеством.

В качестве мер обеспечения информационной безопасности необходимо реализовать комплексную защиту от широкого спектра векторов атак, таких как создание множественных фиктивных идентификаторов с целью манипуляции репутационными метриками (Sybil), перехват или подмена данных (MITM), эксплуатация уязвимостей смарт-контрактов для несанкционированного изменения идентификационных данных, а также DoS/DDoS-атаки на узлы сети с целью нарушения доступности верификационных сервисов [3].

Для противодействия подделкам децентрализованных идентификаторов требуется комплекс решений, включающих многофакторную верификацию на основе криптографических доказательств с нулевым разглашением (ZK-proofs), механизмы подтверждения контроля над ключами через временные метки блокчейна, а также кроссплатформенная валидация данных. Дополнительную защиту обеспечивают непередаваемые soulbound-токены (SBT), которые привязываются к конкретному идентификатору и формируют неотчуждаемую репутационную историю агента.

Интеграция с протоколом Agent-to-Agent (A2A) и платёжными стандартами x402/AP2 создаёт полноценную экосистему для автономной машинной экономики с криптографически верифицируемыми взаимодействиями [4].

В качестве средства обеспечения отказоустойчивости системы хранение метаданных идентификаторов и верификационных документов целесообразно реализовать за счёт распределённого контент-адресуемого хранения InterPlanetary File System (IPFS).

Список литературы

1. Олин Р.А. Методы обеспечения безопасности данных в распределённых системах искусственного интеллекта / Р.А. Олин, В.В. Бондаренко // Материалы Всероссийской научно-технической конференции. – Самара, 2025. – С. 198–200. – EDN ITHYTA

2. Живцов В.Ю. Токенизация репутации /В.Ю. Живцов, Р.А. Олин, К.В. Портнов // Экономика и предпринимательство. – 2025. – № 11(184). – С. 715–719. – EDN PWBNAV.

3. Ranjbar O., Rastagari M. A. Blockchain Applications in Network Engineering: From Secure Routing to Decentralized Identity Management [Электронный ресурс] // JACKA. – 2025. – Vol. 2, № 4. – URL: https://ojs.unimal.ac.id/jacka/article/view/23670 (дата обращения: 22.09.2025).

4. Muttoni A., Zhao J. Agent TCP/IP: An Agent-to-Agent Transaction System [Электронный ресурс] // arXiv preprint. – 2025. – arXiv:2501.06243. – URL: https://arxiv.org/pdf/2501.06243.pdf (дата обращения: 30.09.2025).

УДК 004.056

И.Г. РАЗЕНКОВ, Н.Г. МИЛОСЛАВСКАЯ
Национальный исследовательский ядерный университет «МИФИ», Москва

ПОСТРОЕНИЕ ГРАФА КОРПОРАТИВНОЙ СЕТИ НА ОСНОВЕ ПАССИВНОГО АНАЛИЗА СЕТЕВОГО ТРАФИКА

Рассматриваются вопросы обеспечения информационной безопасности (ИБ) корпоративных сетей (КС) и мониторинга трафика в них. Цель исследования – разработка методики построения графа КС на L2/L3-уровнях модели OSI/ISO. Создан программный комплекс, позволяющий строить граф КС на основе пассивного анализа сетевого трафика.

Введение

Задача построения топологии сети на основе пассивного анализа трафика является актуальной. Современные КС [1] представляют собой сложные многослойные структуры, включающие тысячи различных серверов, рабочих станций, маршрутизаторов и других устройств. Без визуализации интенсивного сетевого трафика в КС трудно обнаруживать малозаметные угрозы ИБ. Графическое отображение сетевой инфраструктуры КС предоставляет специалистам по ИБ возможность визуально отслеживать динамику трафика, аномалии в поведении пользователей, а также быстро выявлять устройства, которые становятся целями компьютерных атак. Граф КС позволяет получить более наглядную визуализацию топологии КС и анализируемого трафика в ней, нежели любой другой метод. Это упрощает поиск уязвимостей в активах КС и предсказание потенциальных компьютерных атак на защищаемую сетевую инфраструктуру.

Постановка задачи

Известно много программных средств, позволяющих осуществлять построение топологии КС на основе активного сканирования ее активов [2]. Однако анализ описаний этих средств показал, что не существует готового решения, удовлетворяющего требованию к программному сканеру анализировать данные трафика различных наиболее часто используемых сетевых протоколов в пассивном («бесшумном») режиме и на этой основе строить граф КС. Поэтому цель исследования – разработка пассивного сканера сетевого трафика, устраняющего выявленные недостатки активных, например, невозможность обнаружить в сети новые устройства, или устройства, намеренно остающиеся скрытыми. Также в некоторых ситуациях активное сканирование может быть недопустимо из-за слишком

большой нагрузки на КС, взаимодействия с чувствительным оборудованием или каких-либо других причин, требующих скрытого анализа КС [3].

Результаты

В ходе исследования создана методика построения графа КС на L2/L3-уровнях модели OSI/ISO, которая заключается в получении, накоплении и фильтрации информации о топологии сети из трафика, построении на ее основе графа и обогащении графа дополнительными данными. Перечень задач, при решении которых может быть применена данная методика, это мониторинг сетевой активности в КС, что обязательно для эффективного обеспечения ИБ, выявление активных сетевых устройств и информационных потоков между ними, контроль целостности сетевой инфраструктуры, проведение аудита ИБ КС и тестирований на проникновение, а также анализ неизвестной сетевой инфраструктуры на основе записанного трафика.

Разработана программная реализация предлагаемой методики на языке C++ [4]. Программный комплекс позволяет строить граф КС в автоматическом режиме (без взаимодействия с пользователем) на основе пассивного анализа трафика наиболее часто встречающихся в КС сетевых протоколов, таких как ARP, DHCP, CDP, LLDP, DNS, MDNS, SNMP, OSPF, BGP, HTTP, FTP и т.д. Комплекс обладает высокой производительностью и позволяет производить обработку сотен гигабайт трафика – так, построение графа на основе предобработанной записи трафика (предобработка заключается только в выделении нужных протоколов) размером ~110 Гб занимает порядка 30 с.

Заключение

Далее на основе полученных результатов планируется тестирование разработанного программного комплекса на реальных задачах, оптимизация его работы и расширение функционала.

Список литературы

1. Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы, технологии, протоколы: Юбилейное издание. – СПб.: Питер, 2021. – 1008 с.

2. Андреев А.А. Математические модели, методы и комплекс программ для описания структуры локальной вычислительной сети при неполных исходных данных: дис. канд. техн. наук: 05.13.18 / Андреев Антон Александрович. – Петрозаводск, 2021. – 144 с.

3. Hosmer C. Python passive network mapping: P2NMAP (1sh Edition). Syngress, 2015. 162 p.

4. Straustrup, B. The C++ Programming Language (4th Edition). Addison-Wesley Professional, 2013. 1376 p.

УДК 004.056

С.С. ВЕЛИГОДСКИЙ, Н.Г. МИЛОСЛАВСКАЯ

Национальный исследовательский ядерный университет «МИФИ», Москва

ВИЗУАЛИЗАЦИЯ УРОВНЯ ЗРЕЛОСТИ ЦЕНТРОВ УПРАВЛЕНИЯ СЕТЕВОЙ БЕЗОПАСНОСТЬЮ

Для разработанной ранее модели оценки уровня зрелости центров управления сетевой безопасностью (ЦУСБ) информационно-телекоммуникационных сетей (ИТКС) субъектов критической информационной инфраструктуры (КИИ) Российской Федерации (РФ) предлагается способ визуализации итогового уровня зрелости ЦУСБ по пяти направлениями оценки (НО).

Введение

Модель зрелости ЦУСБ ИТКС субъекта КИИ РФ – это структурированный набор элементов, объединяющих информационную потребность установления уровня зрелости ЦУСБ (знания, необходимые для управления целями, задачами, рисками и проблемами ЦУСБ [1]) с их атрибутами (свойствами или характеристиками ЦУСБ, определяемыми количественно или качественно вручную или автоматизированно). Выявленные недостатки известных моделей, отличающихся отсутствием детального описания модели и методики ее применения, узкой направленностью на отдельные процессы, неполным охватом области оценки [2], позволили предложить собственную унифицированную модель зрелости (УМЗ) ЦУСБ (рис. 1) [3].

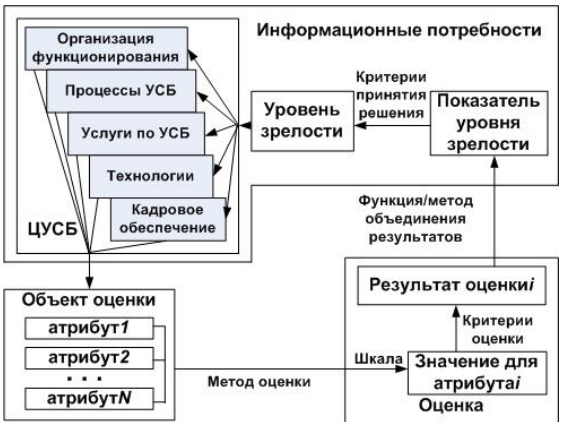

Рис. 1. Оценивание уровня зрелости ЦУСБ по пяти НО

Постановка задачи

Цель исследования – предложить способ визуализации итогового уровня зрелости ЦУСБ, наглядно демонстрирующий возможные направления его дальнейшего совершенствования.

Результаты исследования

Выбран способ визуализации итогового уровня зрелости ЦУСБ в виде круговых диаграмм (рис. 2): первая и третья диаграммы отображают обобщенные результаты оценки значений показателей уровней зрелости организационного и кадрового обеспечения соответственно, а на второй одновременно отображаются результаты оценки уровней зрелости процессов управления сетевой безопасностью (УСБ), услуг УСБ и используемых для этого технологии.

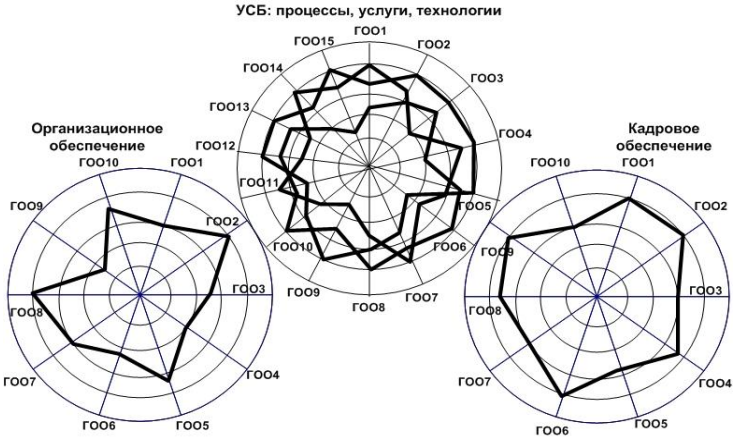

Рис. 2. Круговые диаграммы отображения показателей уровней зрелости

Список литературы

1. Милославская Н.Г. Научные основы построения центров управления сетевой безопасностью в информационно-телекоммуникационных сетях. М.: Горячая Линия-Телеком, 2021. – 431 с.

2. Велигодский С.С., Милославская Н.Г. Анализ подходов к оценке уровня зрелости центров управления сетевой безопасностью. – Сборник научных трудов Всероссийской научно-технической конференции «Кибернетика и информационная безопасность «КИБ-2023». – МИФИ, 2023. – С. 70–71.

3. Велигодский С.С., Милославская Н.Г. Унифицированная модель зрелости центров управления сетевой безопасностью информационно-телекоммуникационных сетей. – Известия ЮФУ. Технические науки, 2023. – №. 3(233). – С. .157–172. – Режим доступа: https://izv-tn.tti.sfedu.ru/index.php/izv_tn/article/view/811 (дата обращения: 02.07.2025).

УДК 004.056:004.8

Н.Г. МИЛОСЛАВСКАЯ

Национальный исследовательский ядерный университет «МИФИ», Москва

СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СРЕДСТВАХ ОБЕСПЕЧЕНИЯ СЕТЕВОЙ БЕЗОПАСНОСТИ

Опираясь на основные положения проекта ГОСТ Р «Искусственный интеллект в критической информационной инфраструктуре..», аналитические исследования и лучшие мировые практики, определены средства обеспечения сетевой безопасности (СОСБ) информационно-телекоммуникационных сетей (ИТКС), в которых могут использоваться системы искусственного интеллекта (ИИ), и их типичные уязвимости, снижающие защищенность ИТКС организаций.

Введение

В июле 2025 г. опубликован проект ГОСТ Р «Искусственный интеллект в критической информационной инфраструктуре. Общие положения» [1], регулирующий применение ИИ в критической информационной инфраструктуре (КИИ). Актуальность стандарта связана с тем, что внедрение систем ИИ, способных для заданной цели генерировать выводы и моделировать человеческое поведение на основе анализа внешних данных и опыта без подробного программирования логики решения конкретных задач, в объекты КИИ (например, ИТКС) требует особого внимания к вопросам обеспечения безопасности таких систем.

Постановка задачи

Технологии ИИ могут как повысить эффективность и автоматизацию процессов в КИИ, так и одновременно создать новые уязвимости. С одной стороны, встроенный в СОСБ ИИ позволяет оптимизировать работу администраторов, например, автоматизирует обнаружение угроз, анализируя в реальном времени огромный объем данных в лог-файлах и поведение пользователей и выявляя отклонения и закономерности. С другой стороны, сам ИИ требует защиты и контроля, поскольку доверять ПО, содержащему его, можно не всегда. В тоже время ИИ используется и злоумышленниками, например, для создания самообучающегося вредоносного ПО или написания стилистически грамотных фишинговых писем. Необходимо определить СОСБ, в которых могут использоваться системы ИИ, и их типичные уязвимости, снижающие защищенность ИТКС субъектов КИИ.

Результаты исследования

Проведя аналитические исследования и опираясь на лучшие мировые практики, в качестве СОСБ, в которых могут использоваться системы ИИ, выделены средства управления доступом, обнаружения и реагирования на угрозы, симуляции атак, контроля безопасности конфигураций, распознавания фишинга, спама и другого нежелательного контента, тестирования программного кода приложений, поиска информации по открытым источникам, а также создания центров управления сетевой безопасностью [2] и киберполигонов.

Типовыми уязвимостями применения ИИ при ОИБ активов ИТКС [3] являются, например, возможность заражения данных, инверсии и отравления модели, кражи модели, утечки конфиденциальной информации, обхода защиты, вывода данных, социальной инженерии с использованием ИИ, атак через «тайный вход» и *API*, атак с переносом обучения, *DoS*- атак и т.п., а также уязвимости аппаратного обеспечения.

Исследован интернет-ресурс *AI Vulnerability Database,* представляющий собой открытую расширяемую базу знаний в области уязвимостей моделей ИИ. Определены и другие источники по уязвимостям систем ИИ, в совокупности формирующие многогранное представление для более полного понимания рисков, связанных с ИИ, например, *MIT AI Risk Repository, MITRE ATLAS, NIST AI Risk Management Framework и т.п.*

Заключение

Далее на основе полученных результатов планируется сформулировать рекомендации для повышения защищенности организаций, применяющих системы ИИ в КИИ.

Список литературы

1. Проект ГОСТ Р «Искусственный интеллект в критической информационной инфраструктуре. Общие положения» [Электронный ресурс]. – Режим доступа: https://fstec.ru/tk-362/deyatelnost-tk362/rassmotrenie-dokumentov-smezhnymi-tk/tk-164-iskusstvennyj-intellekt (дата доступа: 01.10.2025).

2. Милославская Н.Г. Научные основы построения центров управления сетевой безопасностью в информационно-телекоммуникационных сетях. М., Горячая Линия-Телеком, 2021. – 431 с.

3. Top 14 AI Security Risks in 2025 [Электронный ресурс]. – Режим доступа: https://www.sentinelone.com/cybersecurity-101/data-and-ai/ai-security-risks (дата доступа: 01.10.2025).

УДК 004.056

М.Ю. ТОЛСТЫХ, Е.Д. ФИЛИППОВА

Московский государственный лингвистический университет

ОЦЕНКА УГРОЗ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ, НАПРАВЛЕННЫХ НА ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Статья посвящена вопросам анализа угроз безопасности информации, связанных с интеграцией технологий искусственного интеллекта. Обосновывается необходимость подхода, связанного с использованием методологии, разработанной регулятором и модифицированной в алгоритмическом решении с учетом специфики угроз, характерных для указанных технологий.

Наряду с неоспоримым преимуществом своего применения технологии искусственного интеллекта (ТИИ) приводят к возникновению новых угроз в сфере информационной безопасности (ИБ). Проблема защиты данных в ИИ актуальна не только для крупных и широко применяемых моделей (GPT-5, YandexGPT2), взлом которых может привести к нежелательным последствиям для государства, но и для локальных, используемых различными компаниями для ведения бизнеса.

Из-за отсутствия регламентированного процесса классификации и оценки угроз ИБ, направленных непосредственно на ТИИ, для реализации данной процедуры предлагается взять за основу алгоритм, представленный в методике, утвержденной ФСТЭК России [1].

При осуществлении анализа угроз выделим несколько этапов: характеристика негативных последствий; идентификация возможных объектов воздействия; определение источников угроз безопасности информации; оценка способов реализации угроз; оценка их актуальности.

Целесообразно выделить три основные группы негативных последствий от реализации угроз, направленных на ТИИ: для физических лиц (например, потеря контроля над устройством, угроза здоровью человека при взломе оборудования), для юридических лиц (простой и финансовые потери при отказе сервисов, ущерб репутации при утечках данных) и для государства (атаки на критическую информационную инфраструктуру); при этом корректная идентификация негативных последствий позволит адекватно оценить риски, распределить ресурсы на защиту наиболее уязвимых объектов, а также сформировать стратегию реагирования на инциденты.

Далее необходимо идентифицировать компоненты системы, согласующиеся с этапами жизненного цикла ТИИ, которые могут быть целью атаки: данные (обучающие датасеты); модель (код, API); инфраструктура (сервер, облачные хранилища); потоки данных (логи, промежуточные результаты).

Следующий этап заключается в определении видов нарушителей. Предлагается следующая классификация: первая группа (злоумышленник в ходе атаки может напрямую взаимодействовать с вычислительной моделью ИИ, имеет доступ к обучающим данным модели и алгоритмам ее функционирования); вторая группа (злоумышленник не имеет доступа к внутренним параметрам и архитектуре модели, но имеет возможность анализировать входные и выходные данные модели, с помощью которых можно раскрыть ее характеристики); третья группа сочетает возможности злоумышленников первой группы, но с ограничениями [2].

Далее следует произвести анализ атак, которые могут быть совершены нарушителями. В базе данных угроз регулятора [3] представлено несколько, направленных непосредственно на ТИИ: угроза раскрытия информации о модели машинного обучения; угроза хищения обучающих данных; угроза нарушения функционирования («обхода») средств, реализующих технологии ИИ и др. В марте 2025 г. Сбер разработал «Модель угроз для кибербезопасности AI», содержащую 70 угроз, которые могут быть актуальны на различных этапах жизненного цикла ИИ: при сборе и подготовке данных; разработке модели и обучении; эксплуатации модели и интеграции с приложениями.

Таким образом, несмотря на отсутствие единого алгоритма оценки угроз ИБ, направленных непосредственно на ТИИ, возможно реализовать данную процедуру, используя методику регулятора. Идентифицировав актуальные угрозы, необходимо обеспечить комплексную защиту, сочетающую технические и организационные меры, чтобы обеспечить надежность информационной инфраструктуры и доверие пользователей.

Список литературы

1. Методический документ. Методика оценки угроз безопасности информации: утв. ФСТЭК России 05.02.2021. [Электронный ресурс] // URL: https://www.consultant.ru/document/cons_doc_LAW_378330/ (дата обращения: 09.09.2025).

2. Ветров И.А. Особенности использования марковских процессов принятия решений при моделировании атак на системы искусственного интеллекта / И.А. Ветров, В.В. Подтопельный // Вестник Самарского университета. Естественнонаучная серия. 2024. С. 147–160.

3. Банк данных угроз безопасности информации ФСТЭК России. [Электронный ресурс] // URL: https://bdu.fstec.ru (дата обращения: 09.09.2025).

УДК 004.056

А.А. СЕРОВА

МИРЭА – Российский технологический университет, Москва

СОСТЯЗАТЕЛЬНЫЕ АТАКИ НА НЕЙРОННЫЕ СЕТИ: ВИДЫ АТАК И МЕТОДЫ ЗАЩИТЫ

В работе представлена классификация атак по типу доступа к модели «белый» и «черный ящик», детально рассмотрены три ключевых метода компрометации: атака одного пикселя, метод знака градиента и атака на основе карты значимости Якобиана. Особое внимание уделено методам предотвращения компрометации сообщений.

Состязательные атаки на нейронные сети – это атаки, которые основываются на особенностях машинного обучения и эксплуатируют уязвимости глубинных нейронных сетей, что позволяет компрометировать распознавание сообщения и кардинально менять результат работы модели. Существуют два типа состязательных атак. Если злоумышленник обладает полными знаниями о модели, говорят об атаке белого ящика. Если же злоумышленник может только подавать на вход изображения и получать результат работы модели, то такую атаку называют атакой чёрного ящика. Можно выделить несколько наиболее часто используемых методов компрометации изображений [1].

Атака одного пикселя – атака, работающая в условиях «чёрного ящика» и целью которой является изменение значения всего одного наиболее значимого пикселя в исходном изображении. Выделяют два вида атаки: направленная и ненаправленная. Направленная атака применима в тех случаях, когда злоумышленник хочет заставить модель предсказать конкретный, заранее выбранный целевой класс. Задача ненаправленной атаки – сделать так, чтобы изменённое изображение нейронная сеть отнесла к любому другому целевому классу, главное не к желаемому. С помощью алгоритмов оптимизации находится позиция одного пикселя и его новое значение цвета, которые максимизируют вероятность ошибки классификации [2]. Атака, основанная на знаке градиента – атака, являющаяся разновидностью метода атаки «белого ящика», в основе которой лежит многократное обновление через серию итераций, в каждой из которых значения пикселей изменяются в направлении увеличения потерь целевой функции. Различают два вида атаки: быстрый метод знака градиента и итеративный метод. Математически атака описывается выражением: $x'=x+\varepsilon\cdot\mathrm{sign}(\nabla x J(\theta.x,y))$, где x – исходные данные, ε –

минимальная константа шума, J – функция потерь, θ – параметры модели, y – истинная метка. Быстрый метод знака градиента помогает выявлять и защищать уязвимые участки в архитектурах искусственного интеллекта [2].

Атака карты значимости на основе Якобиана. Алгоритм начинается с выбора целевой модели для атаки. Затем вычисляется вектор градиентов, позволяющий определить те части, которые влияют на классификацию картинки как подделку. Затем выбранные признаки начинают итеративно изменяться, увеличивается их и уменьшается влияние других частей изображения, что приводит к результату, когда нейронная сеть классифицирует подделку.

Для защиты нейронных систем от состязательных атак существуют различные методы. Маскирование градиентов - техника искажения или скрытия информации о градиентах модели, что усложняет для злоумышленника вычисление направлений для создания эффективных атакующих возмущений.

В настоящее время существует множество способов атак, для простых и сложных сетей. Основными способами защиты нейронных сетей от состязательных атак являются состязательное обучение, оборонительная дистилляция, нейронная очистка, маскирование градиентов, реконструкция входных данных с помощью автоэнкодеров, детекция атакующих примеров и применение ансамблей и других продвинутых методов повышения устойчивости моделей [3]. Каждая из этих техник имеет свои особенности, преимущества и ограничения. Эффективная защита от состязательных атак требует сочетания нескольких подходов, а также постоянного обновления защиты в ответ на развитие новых методов атак. Кроме того, важна балансировка между устойчивостью к атакам и сохранением точности модели на нормальных данных.

Список литературы

1. Котенко И.В., Саенко И.Б., Лаута О.С., Васильев Н.А., Садовников В.Е. Атаки и методы защиты в системах машинного обучения: анализ современных исследований. Вопросы кибербезопасности 2024 №1(59). DOI: 10.21681/2311-2024-1-24-37.

2. Чехонина Е.А., Костюмов В.В. Обзор состязательных атак и методов защиты для детекторов объектов. International Journal of Open Information Technologies. 2023. №7. URL: https://cyberleninka.ru/article/n/obzor-sostyazatelnyh-atak-i-metodov-zaschity-dlya-detektorov-obektov.

3. Sber. 3 метода состязательных атак на глубокие нейронные сети: как обмануть ИИ. Habr. URL: https://habr.com/p/918702.

УДК 004.056

А.В. КУЗНЕЦОВ

ООО «РТК ИБ», Москва

Финансовый университет при Правительстве Российской Федерации, Москва

КАТАЛОГ ЛОГИЧЕСКИХ ДЕЙСТВИЙ ПО ЛОКАЛИЗАЦИИ КОМПЬЮТЕРНЫХ ИНЦИДЕНТОВ

Исследование направлено на повышение качества и своевременности локализации компьютерных инцидентов в рамках реализации мер реагирования. По результатам исследования предложен каталог логических действий по локализации компьютерных инцидентов, который, в отличие от известных, систематизирует (структурирует) действия на основе жизненного цикла объектов в компьютерной сети, инвариантен к объектам воздействия и может быть использован для автоматизации мер реагирования (как словарь синонимов).

Введение

Реагирование на компьютерные инциденты является неотъемлемой частью противодействия нарушениям и компьютерным атакам в компьютерных сетях, которых с каждым годом становится всё больше. Локализация компьютерных инцидентов является первичным мероприятием, активно противодействующим нарушителю [1]. На практике превалируют экспертные формулировки для действий по локализации компьютерных инцидентов, напрямую зависящие от квалификации членов групп реагирования (ГР), в т.ч. используются англицизмы, сленг и т.п. И, к сожалению, отсутствуют формализованные систематизированные каталоги действий по локализации компьютерных инцидентов (словари синонимов), которые можно было бы использовать членам ГР как справочники.

Результаты частотного анализа рекомендаций по реагированию

Автором проведён частотный анализ 1 185 карточек компьютерных инцидентов (разделов с текстами рекомендаций по реагированию, включая локализацию компьютерных инцидентов), зарегистрированных за одну неделю для распределённой информационной инфраструктуры, находящейся под компьютерной атакой (от разведки до нанесения ущерба). Анализ проводился в рамках процесса информационно-аналитического обеспечения деятельности ГР с учётом работы [2], включая токенизацию, удаление стоп-слов, лемматизацию и разметку частей речи.

По итогам анализа определены наиболее часто предлагаемые ГР логические действия (допускается вхождение нескольких логических действий в одну карточку, от одного до четырёх действий включительно):

– действия по блокировке учётных записей: 69%;
– действия по сетевой изоляции хостов компьютерной сети: 65%;
– действия по удалению объектов с хостов компьютерной сети: 25%;
– действия по восстановлению исходных параметров конфигураций хостов компьютерной сети: 9%.

Каталог логических действий

По итогам анализа предложен словарь синонимов (базовых форм существительных), базирующихся на жизненном цикле (ЖЦ) любого объекта в компьютерной сети (хоста, приложения и т.п.), а именно:

– возникновение, создание, установка, инсталляция, восстановление, (пере)запись, начало, старт, ввод;
– активация, включение, подключение, (пере)запуск, разблокировка;
– изменение, смена, перемещение, перенаправление, изолирование, (пере/до)настройка, (пере)конфигурирование, доработка, модификация, трансформация, коррекция, исключение;
– деактивация, выключение, отключение, завершение, прекращение, остановка, блокировка, сброс, разрыв, прерывание, откат, вывод;
– исчезновение, удаление, уничтожение, очистка, зануление, обезвреживание, устранение.

Предложенный каталог может быть перенесён в средства реагирования (EDR, NDR и др.) и управления ими (IRP, SOAR) для автоматизации [3].

Заключение

Предложенный каталог действий систематизирует действия на основе ЖЦ объектов компьютерной сети, может быть использован в средствах автоматизации как справочник, который повысит унификацию наполнения и парсинга карточек, и, как следствие, своевременность локализации.

Список литературы

1. Милославская Н.Г., Толстой А.И. Управление инцидентами информационной безопасности. М.: Горячая Линия - Телеком, 2024. – с. 105–109, 169–193.
2. Крашенинников М.С., Лавреш И.И., Устюгов В.А. Разработка и организация бизнес-процессов подготовки частотных словарей в целях автоматизации обработки естественного языка при выполнении задач по лингвистическому анализу текстов. Вестник Сыктывкарского университета. Серия 1: Математика. Механика. Информатика. 2023, № 3 (48), с. 72–89. DOI: https://doi.org/10.34130/1992-2752_2023_3_72. – EDN: LCFWZK.
3. Кузнецов А.В. Методология реагирования на инциденты информационной безопасности в распределенных автоматизированных информационных системах. Вопросы кибербезопасности. 2025, № 4 (68), с. 65–72. DOI: 10.21681/2311-3456-2025-4-65-72. – EDN: ULFTTD.

УДК 004.056; 004.8

Е.А. ВОРОБЬЕВА, М.Ю. ТОЛСТЫХ

Московский государственный лингвистический университет

ИНТЕЛЛЕКТУАЛЬНАЯ КООРДИНАЦИЯ ДЛЯ ОБНАРУЖЕНИЯ И РЕАГИРОВАНИЯ НА СЕТЕВЫЕ ИНЦИДЕНТЫ В РАСПРЕДЕЛЕННЫХ СИСТЕМАХ

Предлагается концепция интеллектуальной координации на основе искусственного интеллекта для повышения эффективности в крупномасштабных средах. Подход интегрирует машинное обучение для выявления аномалий, глубокое обучение для классификации угроз и обучение с подкреплением для адаптивного реагирования.

Современная ИТ-инфраструктура активно развивается в сторону распределенных моделей (облака, периферийные вычисления, IoT). Для решения актуальных проблем их безопасности, в том числе реализации законодательных требований [1, 2], требуется переход к национальной и более интеллектуальной и адаптивной киберзащите. Искусственный интеллект (ИИ), включая машинное обучение (ML) и глубокое обучение (DL), предоставляет мощные инструменты для анализа больших данных и автоматизации принятия решений. Однако его применение в распределенных системах требует особого внимания к работе с данными, обеспечению из защиты и координации множества автономных агентов.

Актуальна разработка концепции интеллектуальной координации для повышения эффективности обнаружения и реагирования на сетевые инциденты в распределенных системах. Основной упор ориентирован на создание синергии различных ИИ-технологий, где совместная работа обеспечит надежный интеллектуальный уровень защиты.

Предлагаемая архитектура основана на принципах многоуровневой интеллектуальной координации, где каждый уровень выполняет свою специфическую роль в общей стратегии защиты.

• Уровень обнаружения на узлах (Node-Level Detection, NLD): каждый узел оснащается автономными модулями обнаружения, которые используют легковесные ML-алгоритмы для локального анализа данных, генерируемых самим узлом; основная задача – выявление локальных аномалий; благодаря ML модули способны обнаруживать не только известные угрозы, но и не встречавшиеся атаки, основываясь на поведенческом анализе.

• Уровень анализа и классификации угроз (Threat Analysis and Classification, TAC): применяются более мощные модели DL, способные коррелировать разрозненные аномалии, собранные с различных узлов, для построения более полной картины инцидента, классифицировать выявленные угрозы и оценивать степень риска; результат работы – семантически обогащенная информация об угрозах.

• Уровень интеллектуальной координации и принятия решений (Intelligent Coordination and Decision-Making, ICDM): ядро системы, отвечающее за принятие стратегических решений; объединяет возможности RL и многоагентных систем для выработки оптимальной и адаптивной рефлексии; генерирует динамические политики реагирования, подстраивающиеся к изменяющимся условиям и характеру угроз.

• Уровень выполнения реагирования (Response Execution, RE): отвечает за исполнение предписанных действий на соответствующих узлах (изоляция скомпрометированного узла, блокировка вредоносного трафика); обеспечивает трансляцию решений ICDM в управляющие команды, которые приводят в действие защитные механизмы системы.

Предложенная архитектура предоставляет ряд преимуществ для обнаружения и реагирования на сетевые инциденты в распределенных системах: повышенное качество обнаружения (за счет синергии локального анализа аномалий на NLD и глобальной классификации угроз на TAC); оперативное и адаптивное реагирование (благодаря применению RL на ICDM); масштабируемость системы (ввиду децентрализованного характера сбора и первичного анализа данных).

Таким образом, предложенная архитектура интегрирует различные ИИ-технологии: от обнаружения локальных аномалий на периферии до принятия стратегических решений на основе ML и DL. Ключевым элементом является синергия компонентов, позволяющая анализировать разрозненные события и выстраивать целостную картину угроз, а также вырабатывать адаптивные, своевременные меры реагирования.

Список литературы

1. Указ Президента РФ «О дополнительных мерах по обеспечению информационной безопасности РФ» от 1.05.2022 № 250 [Электронный ресурс] // URL: https://base.garant.ru (дата обращения: 10.09.2025).

2. Постановление Правительства РФ «О порядке перехода субъектов критической информационной инфраструктуры РФ на преимущественное применение доверенных программно-аппаратных комплексов на принадлежащих им значимых объектах критической информационной инфраструктуры РФ» от 14.11.2023 № 1912 [Электронный ресурс] // https://www.garant.ru (дата обращения: 09.09.2025).

УДК 004.056

П.А. ИВАНОВА

МИРЭА – Российский технологический университет, Москва

СОВРЕМЕННЫЕ ТЕНДЕНЦИИ ПО ОБМЕНУ ИНДИКАТОРАМИ КОМПРОМЕТАЦИИ ПО ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Статья рассматривает современные тенденции обмена индикаторами компрометации (IoC) в информационной безопасности, включая использование автоматизированных платформ SIEM и TIP для оперативного выявления угроз. В России развивается практика создания отраслевых центров для централизованного сбора и анализа IoC. Современный тренд – автоматизация, стандартизация и совместный анализ для повышения эффективности защиты.

Введение

Современные тенденции обмена индикаторами компрометации (IoC) в области информационной безопасности включают активное использование автоматизированных платформ и систем, таких как SIEM (Security Information and Event Management) и TIP (Threat Intelligence Platform). Эти системы позволяют эффективно собирать, сопоставлять и анализировать данные об угрозах для оперативного обнаружения инцидентов и реагирования на них.

Индикатор компрометации (Indicator of Compromise, IoC) – это фрагмент технических данных или цифровой артефакт, который с высокой вероятностью указывает на несанкционированный доступ или вредоносную активность в информационной системе [1].

Современные тенденции по обмену индикаторами компрометации

Использование платформ SIEM и TIP для автоматического сопоставления телеметрии инфраструктуры с коллекциями IoC позволяет оперативно выявлять угрозы и снижать нагрузку на аналитиков. TIP обеспечивают более гибкую и специализированную обработку IoC, включая поддержку соответствующих схем данных и настроек фильтрации.

Рост применения технологий машинного обучения и искусственного интеллекта для более точного анализа и фильтрации индикаторов приводит к уменьшению количества ложных срабатываний и повышению

релевантности данных, усиливая аналитические возможности команд безопасности [2].

Активное использование открытых и коммерческих Threat Intelligence фидов (TI feeds) с распространёнными стандартами обмена данных, такими как STIX/TAXII, OpenIOC, CybOX, которые содержат списки хеш-сумм вредоносных файлов, IP-адресов, доменных имен и других атрибутов, обеспечивает интеграцию с системами защиты.

В России развивается идея создания отраслевых информационно-аналитических центров под эгидой регуляторов (например, ФСТЭК), которые будут централизованно агрегировать и анализировать IoC для критической инфраструктуры, обеспечивая координацию и упреждающее обнаружение атак [3].

Происходит популяризация решений для автоматизации верификации и очистки IoC от мусора (например, GOSINT), что позволяет повысить качество и надежность информации для SOC и центров реагирования.

Также, имеют место увеличение роли объединённых усилий по обмену разведданными через открытые и закрытые сообщества, стандартизация форматов и расширение возможностей для совместного анализа угроз [4].

Заключение

Таким образом, современные тенденции направлены на максимальную автоматизацию, использование искусственного интеллекта, стандартизацию форматов обмена и развитие отраслевых платформ для эффективного обмена и анализа индикаторов компрометации.

Список литературы

1. Индикатор компрометации (Indicator of Compromise, IoC) // Энциклопедия Kaspersky. URL: https://encyclopedia.kaspersky.ru/glossary/indicator-of-compromise-ioc/ (дата обращения: 13.10.2025).

2. Москвин А. Кибербразведка для бизнеса и ИБ: руководство / Андр. Москвин. – Электрон. дан. – URL: https://gdspace.ru/articles/2530 (дата обращения: 13.10.2025).

3. Scattered Lapsus$ и «Троица хаоса»: новая волна киберугроз // CISO Club. URL: https://cisoclub.ru/scattered-lapsus-i-troica-haosa-novaja-volna-kiberugroz/ (дата обращения: 13.10.2025).

4. Кибербразведка по-русски: как развивается отечественный Threat Intelligence // Habr, 2025. URL: https://habr.com/ru/articles/933642/ (дата обращения: 13.10.2025).

УДК 004.056

В.Е. ВЕБЕР, И.О. ВИНОГРАДОВА, М.М. АФАНАСЬЕВА

МИРЭА – Российский технологический университет, Москва

СОВРЕМЕННЫЕ МЕТОДЫ ПОВЕДЕНЧЕСКОГО АНАЛИЗА РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ НА ОСНОВЕ СЕРВИС-ОРИЕНТИРОВАННОЙ АРХИТЕКТУРЫ

В статье проведено исследование современных методов поведенческого анализа взаимодействия сервисов распределённых информационных систем (РИС), построенных на основе сервис-ориентированной архитектуры (SOA, service-oriented architecture). Рассмотрены и проанализированы ключевые подходы к повышению безопасности, киберустойчивости и надёжности РИС посредством мониторинга взаимодействия их сервисов. Наиболее перспективным является использование методов причинно-следственного логического анализа взаимодействия сервисов.

Введение

Современные распределенные информационные системы, основанные на сервис-ориентированной архитектуре (SOA), сталкиваются с серьезными проблемами в области информационной безопасности. Традиционные методы защиты оказываются недостаточно эффективными против атак, направленных на нарушение логики взаимодействия между сервисами [1]. Основная проблема заключается в том, что существующие подходы к мониторингу не способны обнаруживать аномалии, связанные с отклонениями в последовательности вызовов сервисов и несоответствием бизнес-логике системы [2].

Методы поведенческого анализа распределенных информационных систем и их реализация

Слабая связанность сервисов в SOA-системах создает новые уязвимости, которые невозможно обнаружить традиционными средствами безопасности. Для решения этой проблемы предлагается комплекс методов поведенческого анализа, интегрирующий различные подходы к мониторингу и диагностике распределенных систем.

Анализ логов и трассировок направлен на диагностику распределенных сбоев путем восстановления полного пути выполнения запросов. Использование инструментов трассировки позволяет выявлять

каскадные сбои и точки отказа в цепочках взаимодействия сервисов, обеспечивая видимость сквозных транзакций.

Контроль соответствия бизнес-процессам обеспечивает обнаружение атак на бизнес-логику через сравнение реальных последовательностей вызовов сервисов с эталонными моделями бизнес-процессов. Это позволяет выявлять попытки обхода критических шагов выполнения операций.

Проактивный мониторинг аномалий решает проблему обнаружения неизвестных угроз через анализ паттернов взаимодействия и выявление статистически значимых отклонений от нормального поведения системы, обеспечивая превентивное реагирование на потенциальные инциденты безопасности.

Для практической реализации предложенных методов используются техники моделирования UML и SoaML, обеспечивающие формальное описание взаимодействий сервисов, а также инструменты автоматизированного мониторинга, системы сбора и корреляции метрик, средства формальной верификации взаимодействий. Автоматизация процессов мониторинга позволяет снизить сложность анализа поведения распределенной системы и обеспечить своевременное обнаружение аномалий [3]. Интеграция перечисленных инструментов и методик создает единую систему поведенческого анализа, способную эффективно противостоять современным угрозам безопасности в SOA-системах.

Заключение

Предложенный комплекс методов поведенческого анализа позволяет эффективно решать проблемы информационной безопасности SOA-систем, связанные с нарушением логики взаимодействия сервисов. Анализ трассировок, контроль бизнес-процессов и мониторинг обеспечивают комплексную защиту от современных угроз, направленных на нарушение целостности распределенных информационных систем.

Список литературы

1. Blal R., Leshob A., Mili H., Benzarti I., Hadaya P., Raqeebir R. SOA Services Identification and Design Methods from Business Models: A Systematic Literature Review / Redouane Blal, Abderrahmane Leshob, Hafedh Mili, Imen Benzarti, Pierre Hadaya, Raqeebir Rab // IEEE Access. – 2025. – Vol. 13. – P. 9879–9892. – DOI: 10.1109/ACCESS.2025.3528297.

2. Dragoni N., Lanese I., Larsen S.T., Mazzara M., Mustafin R., Safina L. Microservice Architecture: A Comprehensive Survey // Journal of Systems and Software. – 2020. – Vol. 15, № 1. – P. 1–20. – DOI: 10.1016/j.jss.2020.110535.

3. 10 Ways Microservices Create New Security Challenges // Kong Blog. – 2025. – URL: https://konghq.com/blog/engineering/10-ways-microservices-create-new-security-challenges (дата обращения: 13.10.2025).

УДК 004.056

В.А. ТИХОМИРОВ, Н.Г. МИЛОСЛАВСКАЯ

Национальный исследовательский ядерный университет «МИФИ», Москва

ПОСТРОЕНИЕ ПРОЦЕССА УПРАВЛЕНИЯ УЯЗВИМОСТЯМИ АКТИВОВ ТИПОВОЙ ИТКС НА ОСНОВЕ ГРАФОВЫХ МОДЕЛЕЙ КОМПЬЮТЕРНЫХ АТАК

Предложены подходы к приоритизации уязвимостей активов информационно-телекоммуникационных сетей (ИТКС) на основе анализа разработанного гиперграфа компьютерных атак (КА) и рассмотрению связей между этими уязвимостями при попытке злоумышленника реализовать КА. Исследованы основные характеристики получившегося гиперграфа КА.

Введение

Анализ показал, что существующие методы построения графов КА не подходят для задач ранжирования уязвимостей активов ИТКС, поскольку они не позволяют управлять уязвимостями, а строят пути КА с использованием всех возможных техник и тактик. Эти модели не могут в реальном времени и автоматизированном режиме обрабатывать информацию об уязвимостях. Также не решены проблемы вычислительной сложности, обработки циклических зависимостей и интеграции вероятностных методов оценки. Актуальность исследования обусловлена необходимостью разработать процесс управления уязвимостями активов ИТКС, который бы учитывал не только критичность уязвимости для конкретного актива, но и то, как уязвимость сочетается с другими уязвимостями и некорректными конфигурациями.

Постановка задачи

Цель исследования – рассмотрев основы моделирования графов КА, провести систематический анализ современных подходов к построению и анализу графов КА, включая их классификацию, сравнение и математический аппарат оценки критичности маршрутов КА. Для достижения поставленной цели решались следующие задачи: анализ существующих подходов к описанию графов КА, формулирование методики создания графа КА для задач управления уязвимостями активов ИТКС, разработка алгоритма построения гиперграфа КА [1] и процесса управления уязвимостями [2] с применением графов КА.

Результаты исследования

В рамках исследования проанализированы актуальные подходы к формированию графов КА [3]. В результате выявлено, что существующие подходы к построению графов КА нерелевантны для выбранной задачи приоритизации уязвимостей активов ИТКС. Предложен новый подход к построению графов КА, с учетом специфики задачи приоритизации уязвимостей в рамках процесса управления уязвимостями активов ИТКС. Основа графа – это хостовой граф, который описывает возможные сетевые взаимодействия серверов между собой. В то же время, каждую вершину графа можно рассмотреть, как граф зависимостей эксплуатации, основанный на уязвимостях. В итоге получен гиперграф – пара $H = (X, E)$, где X − множество вершин, а E − семейство подмножеств X, называемых гиперребрами [1]. С учетом особенностей двухуровневого гиперграфа (уровень хоста и уровень сети) разработан алгоритм приоритизации уязвимостей на основе степени посредничества уязвимости в графе.

Выявление многофакторных уязвимостей активов ИТКС, как и выявление уязвимостей архитектуры ИТКС, в автоматизированном режиме возможны лишь при учете связей между разными активами и их уязвимостями, которые злоумышленник может проэксплуатировать при проведении КА. Выявить такие связи можно при помощи построенной графовой модели перемещения злоумышленника в ИТКС организации, включая использование базы данных *CVE*.

Заключение

Созданы методика формирования графа КА, ее формализованное представление и модель гиперграфа КА. На этой основе и NIST SP 1800-5 разработан процесс управления уязвимостями активов ИТКС. Далее предлагаемая модель и алгоритм оценки и приоритизации уязвимостей будут реализованы на практике.

Список литературы

1. Hypergraphs: Combinatorics of Finite Sets [Hypergraphes: Combinatoire des ensembles finis] / Claude Berge. Amsterdam: North-Holland, 1989. 255 p. (North-Holland Mathematical Library; vol. 45). – ISBN 0- 444-87489-5

2. О безопасности критической информационной инфраструктуры Российской Федерации [федер. закон от 26 июля 2017 г. № 187-ФЗ: принят Гос. Думой 12 июля 2017 г.; одобрен Советом Федерации 19 июля 2017 г.]. – 2017.

3. Zenitani K. Attack graph analysis: an explanatory guide. Computers & Security. 2023. Vol. 126. DOI: 10.1016/j.cose.2022.103081.

УДК 004.056

Д.В. ЖОРОВ, Н.Г. МИЛОСЛАВСКАЯ

Национальный исследовательский ядерный университет «МИФИ», Москва

ПРОЦЕСС УПРАВЛЕНИЯ УЯЗВИМОСТЯМИ АКТИВОВ ИТКС С ИСПОЛЬЗОВАНИЕМ СКАНЕРА OPENVAS(GVM)

Цель работы – разработка регламентированного процесса управления уязвимостями активов информационно-телекоммуникационных сетей (ИТКС) с использованием сканера OpenVAS(GVM), что обеспечивало бы соответствие требованиям ФСТЭК России. Основной результат – замкнутый цикл «обнаружение, оценка, устранение и верификация», интегрированного в методологию регулятора.

Введение

Эффективное управление уязвимостями активов ИТКС является критически важным компонентом обеспечения информационной безопасности (ОИБ) организаций, особенно в свет ужесточения регуляторных требований. Сканер OpenVAS(GVM) представляет собой мощный инструмент с открытым исходным кодом для обнаружения уязвимостей [1]. Однако его эффективность напрямую зависит от интеграции в регламентированный процесс, соответствующий современным рекомендациям ФСТЭК России.

Постановка задачи

Цель работы – создание унифицированного процесса управления уязвимостями активов ИТКС с использованием OpenVAS(GVM), который бы охватывал полный жизненный цикл уязвимости – от обнаружения до верификации ее устранения, строго соответствовал методическому документу ФСТЭК России от 17.05.2023 «Руководство по организации процесса управления уязвимостями в органе (организации)» [2], эффективно использовал ключевые функции OpenVAS(GVM), интегрируя их результаты в этапы ручного анализа и принятия решений, а также обеспечивал получение измеримых метрик для оценки эффективности и непрерывного улучшения процесса управления уязвимостями.

Пути решения задачи

Разработан циклический процесс, который превращает разрозненные операции сканирования активов ИТКС в целостную и измеримую систему управления рисками ИБ. Он состоит из последовательных этапов [2], где OpenVAS(GVM) играет ключевую роль на этапах мониторинга и контроля:

1) мониторинг и оценка применимости. Процесс начинается с планирования и запуска сканирования OpenVAS(GVM). Ключевая новая идея – этап предобработки и фильтрации, на котором из «сырого» отчёта исключаются уязвимости, не относящиеся к текущей инфраструктуре. Это значительно снижает нагрузку на аналитиков;

2) оценка критичности уязвимостей активов ИТКС, когда данные OpenVAS(GVM) систематизируются и используются аналитиками для ручной приоритизации. Результат – упорядоченный список уязвимостей с присвоенными уровнями риска ИБ и рекомендованными сроками устранения (SLA);

3) устранение уязвимостей, где OpenVAS(GVM) не задействован напрямую, но предоставленные им данные являются основанием для выбора методов исправления (установки обновлений безопасности, изменения конфигураций);

4) контроль устранения. Наиболее значимый результат – верификация исправлений при повторном сканировании OpenVAS(GVM) с последующим сравнением отчётов «до/после», что обеспечивает объективное подтверждение устранения уязвимостей активов ИТКС,

На основе результатов контроля предложены ключевые метрики эффективности (процент закрытых уязвимостей, среднее время от обнаружения до устранения, количество возвратов), которые легли в основу процедуры непрерывного улучшения процесса, позволяя корректировать политики безопасности, периодичность сканирования и приоритеты устранения уязвимостей.

Заключение

Разработанный процесс управления уязвимостями активов ИТКС с использованием OpenVAS – это комплексное решение, стандартизирующее действия по ОИБ. Его преимущество – создание замкнутого, измеримого цикла, строго соответствующего требованиям ФСТЭК России [2].

Список литературы

1. Сканеры уязвимостей – обзор мирового и российского рынков [Электронный документ]. – Режим доступа: https://www.anti-malware.ru/analytics/Market_Analysis/Vulnerability-scanners-global-and-Russian-markets#part72 (дата обращения: 20.10.2025).

2. Руководство по организации процесса управления уязвимостями в органе (организации). Методический документ (утв. ФСТЭК России 17.05.2023). 2023. 33 с. [Электронный документ]. – Режим доступа: https://fstec.ru/files/1096/---17--2023-/2011/---17--2023-.pdf (дата обращения: 20.10.2025).

УДК 004.056

А.М. РУСАКОВ, Е.А. МАКСИМОВА

МИРЭА – Российский технологический университет, Москва

МЕТОДИКА ВЫЯВЛЕНИЯ УГРОЗ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ИНФРАСТРУКТУРНОГО ГЕНЕЗА

В работе приведена методика для оценки эффектов инфраструктурного деструктивизма. Под инфраструктурным деструктивизмом понимается феномен, возникающий в распределенной информационной системе, когда в результате деструктивных воздействий инфраструктурного генеза происходят системные изменения, ведущие к нарушению устойчивости, целостности, доступности, функциональности и управляемости информационной системы [1, 2]. Даны рекомендации по использованию методики.

Введение

В настоящее время отмечается увеличение количества кибератак на распределенные информационные системы (РИС) связанных с реализацией угроз информационной безопасности отказа в обслуживании. При этом классические методы обеспечения их киберустойчивости не всегда эффективны из-за усложнения и постоянной эволюции инфраструктуры, а также возникающих в системе угроз инфраструктурного генеза (т.е. происхождения), реализация которых может привести к инфраструктурному деструктивизму – неконтролируемому саморазрушению инфраструктуры [2].

Описание методики выявления угроз информационной безопасности инфраструктурного генеза РИС

Основные процедуры частной методики выявления угроз информационной безопасности инфраструктурного генеза в РИС [2] характеризуются следующим.

Шаг 1. Построение сетевой инфраструктуры для заданного количества серверов. Обеспечивается стабильная работа всех сетевых клиентов с серверами.

Шаг 2. Имитация работы РИС. Производится запуск РИС в штатном режиме с имитацией деятельности пользователей по получению информации с серверов. Своевременность доставки информации напрямую зависит от количества серверов и их взаимодействия в аспекте обслуживания пользовательских запросов.

Шаг 3. Сбор данных о работе сетевой инфраструктуры. На данном шаге происходит сбор различной информации (например, журналов событий работы

серверов) при обслуживании пользовательских запросов в штатном режиме работы РИС. Шаг выполняется параллельно с Шагом 2.

Шаг 4. Обработка данных о работе сетевой инфраструктуры. Процедура выполняется после завершения имитации работы РИС и предназначена для обработки информации, собранной в процессе с серверов.

Шаг 5. Выявление эффектов инфраструктурного деструктивизма. Применяется алгоритмы выявления ИД к информации, собранной на Шаге 3 и обработанной на Шаге 4. Например, если на некотором сервере за счет кэширования последующие запросы были выполнены более оперативно, то это будет считаться положительным эффектом; и, наоборот.

Работа данного шага основана на следующем принципе. Поведенческие особенности для запроса описаны в виде наборов правил антропоморфических типов взаимодействия на основе продукционной модели представлений знаний. Для каждого из запросов (которые выгружаются из журналов событий) выполняется пространственно-временная локация на основе последовательных шаблонов. В начале работы алгоритма загружается информация о запросах, для каждого из которых формируется множество анализируемых взаимодействующих запросов и оценивается причинно-следственная связь. Если зависимость детектируется, то формируется множество для описания количественных характеристик антропоморфического взаимодействия запросов, что позволяет определить соответствующий тип эффекта. Количественные характеристики рассчитываются на основе фактического времени выполнения запросов по каждому из антропоморфических типов взаимодействия по отношению к общему времени выполнения исследуемых запросов (измеряется в процентах)

Заключение

Данная методика позволяет оперативно выявлять эффекты ИД в РИС на ранних этапах его возникновения за счет автоматизации процедур обнаружения [3], что является важным как в научном, так и в практическом плане.

Список литературы

1. Максимова Е.А., Русаков А.М. Проактивная оценка динамики рисков инфраструктурного деструктивизма для распределенной системы распознавания лиц // Защита информации. Инсайд. 2025. № 4(124). С. 66–71.

2. Максимова Е.А. Аксиоматика инфраструктурного деструктивизма субъекта критической информационной инфраструктуры // Информатизация и связь. – 2022. – № 1. – С. 68–74. – DOI 10.34219/2078-8320-2022-13-1-68–74. – EDN ZMOPQB.

3. Максимова Е.А., Буйневич М.В. Метод оценки инфраструктурной устойчивости субъектов критической информационной инфраструктуры // Вестник УрФО. Безопасность в информационной сфере. – 2022. – № 1(43). – С. 50–63. – DOI 10.14529/secur220107.

УДК 004.056

И.С. КРЫЛОВ
Научный руководитель – д.т.н., доцент МИЛОСЛАВСКАЯ Н.Г.
Национальный исследовательский ядерный университет «МИФИ», Москва

МОДЕЛЬ ОЦЕНКИ ЗРЕЛОСТИ ПРОЦЕССОВ ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ КОМПАНИИ

Цель исследования – разработка автоматизированного решения для оценки уровня зрелости процессов обеспечения информационной безопасности (ПОИБ) компании, включающего модель, методику и сервис. Основные результаты – создание модели оценки зрелости ПОИБ и методики для сопоставления данных опросов с результатами автоматического сканирования активов компании.

Постановка задачи

Традиционные методы оценки зрелости ПОИБ, основанные на анкетировании и периодических аудитах, демонстрируют системные недостатки: высокую субъективность, запаздывание информации и отсутствие механизмов верификации заявленных практик ОИБ реальному состоянию инфраструктуры. Это формирует «иллюзию безопасности», приводя к принятию необоснованных управленческих решений, и создает значительные риски, в частности, в сфере страхования при определении размеров страховых премий. Особую актуальность проблема приобретает в контексте возросших атак на цепочки поставок, когда уязвимость одного звена может привести к каскадным нарушениям ИБ по всей экосистеме партнеров. Возникает потребность в создании автоматизированного решения, обеспечивающего непрерывный, объективный и автоматически верифицируемый анализ уровня зрелости ПОИБ организации.

Пути решения

Для решения задачи предлагается модель оценки, сочетающая три компонента: 1) *опросный модуль*, основанный на адаптированных вопросниках стандартов NIST CSF, ISO 27001 [1–3] для оценки формальных процессов и политик ИБ; 2) *модуль автоматического непрерывного сканирования (OSINT)*, выполняющий мониторинг инфраструктуры компании, доступной для внешнего злоумышленника, для выявления уязвимостей, неучтенных активов и ошибок конфигурации, и 3) *модуль сопоставления и оценки,* анализирующий данные из первых двух источников и выявляющий рассогласования и рассчитывает итоговую оценку зрелости ПОИБ.

Новизна подхода – в разработке алгоритма сопоставления данных из этих разнородных источников. Модель автоматически сопоставляет ответы респондентов (например, «процесс управления уязвимостями внедрен и функционирует») с объективными техническими данными (например, «наличие неустраненных критических уязвимостей»). Применение решения позволяет перейти от субъективных оценок к динамической и верифицируемой системе анализа уровня зрелости ПОИБ компаний – поставщиков.

Разработанная модель позволяет преодолеть ключевые ограничения традиционных подходов к оценке зрелости ПОИБ. Интеграция опросных и технических данных позволяет убрать субъективность оценки за счет автоматической верификации данных, обеспечить непрерывность мониторинга и динамический пересчет уровня зрелости ПОИБ компаний и повысить обоснованность принятия решений при управлении риском атак на цепочки поставок. Таким образом, модель создает основу для интеллектуального управления безопасностью, ориентированного на данные, и реальные, а не декларируемые показатели. Модель является особенно актуальной для страховых компаний при проведении первоначального анализа уровня страховых рисков, поскольку позволяет получать объективные и проверяемые данные об уровне зрелости ПОИБ компании – страхователя. Применение автоматизированного решения позволит компаниям – покупателям быстро и наглядно определять степень зрелости ПОИБ своих поставщиков, что позволит им более взвешено подходить к принятию решений, связанных с риском атак на цепочки поставок.

В ходе дальнейшей работы по уточнению модели планируется дополнить как сами критерии оценки, так и их весовые коэффициенты модели. В частности, необходимо учитывать размеры компании, вид деятельности. Также создается автоматизированное решение, позволяющее в реальном времени отслеживать изменение оценки зрелости.

Список литературы

1. Информационные технологии. Информационная безопасность во взаимоотношениях с поставщиками. Стандарт ГОСТ ИСО/МЭК 27036 (ISO/IEC 27036). Стандартинформ. Москва. 2021. –117 с.

2. NIST CSWP 29, The NIST Cybersecurity Framework. Gaithersburg: National Institute of Standards and Technology. 2024, – 32 с. DOI: https://doi.org/10.6028/NIST.CSWP.29.

3. NIST SP 800-161r1, Cybersecurity Supply Chain Risk Management Practices for Systems and Organizations. Gaithersburg: National Institute of Standards and Technology, 2022. – 326 с. DOI: https://doi.org/10.6028/NIST.SP.800-161r1

УДК 004.056

Д.А. МАРГАРЯН, Д.С. ЮДИНА

Национальный исследовательский ядерный университет «МИФИ», Москва

ВЫБОР ОПТИМАЛЬНОЙ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ ДЛЯ СИСТЕМЫ ОБНАРУЖЕНИЯ DDOS-АТАК

Распределённые атаки типа «отказ в обслуживании» (DDoS) остаются одной из ключевых угроз сетевой безопасности. В работе проведено сравнение двух алгоритмов машинного обучения для обнаружения аномалий – Isolation Forest и One-Class SVM. На основе набора данных CIC-DDoS2019 установлено, что Isolation Forest обеспечивает более высокие показатели точности и полноты, демонстрируя устойчивость к несбалансированным данным. Полученные результаты позволяют рекомендовать данный алгоритм как базовый инструмент для построения систем раннего обнаружения DDoS-атак.

Современная цифровая инфраструктура критически зависит от бесперебойной работы сетевых сервисов [1]. Распределённые атаки типа DDoS направлены на исчерпание ресурсов сервера или сети. Традиционные методы, основанные на сигнатурах, неэффективны против новых гибридных атак, поэтому применяются алгоритмы машинного обучения, способные выявлять аномалии в сетевом трафике [2].

В исследовании проведено сравнение моделей Isolation Forest и One-Class SVM. Isolation Forest изолирует редкие наблюдения посредством случайных деревьев, тогда как One-Class SVM строит границу, отделяющую нормальные данные от начала координат [3, 4]. Метрики оценки включали точность, полноту, F1-меру и AUC-ROC.

Эксперименты на наборе CIC-DDoS2019 показали, что Isolation Forest обеспечивает точность 0.98, полноту 0.96 и AUC-ROC 0.99, тогда как One-Class SVM – 0.94, 0.91 и 0.95 соответственно. Таким образом, Isolation Forest демонстрирует лучшие результаты и устойчивость к несбалансированным данным.

Дополнительно было проведено сравнение стабильности работы моделей при изменении объёма обучающей выборки и степени шумности данных. Результаты показали, что алгоритм Isolation Forest демонстрирует меньшую чувствительность к выбросам и способен сохранять высокие показатели точности даже при снижении доли обучающих данных на 30%. В то время как One-Class SVM в подобных условиях терял до 7% точности и демонстрировал рост ложноположительных срабатываний.

Анализ вычислительной эффективности показал, что Isolation Forest требует меньше ресурсов по сравнению с One-Class SVM, особенно при обработке больших объёмов потоковых данных. Среднее время предсказания для одной записи оказалось на 25–30% меньше, что делает возможным его применение в системах обнаружения атак в режиме реального времени.

Алгоритм Isolation Forest можно рекомендовать для реализации систем раннего обнаружения DDoS-атак. Перспективы дальнейших исследований включают использование ансамблей и нейронных сетей для анализа временных рядов сетевого трафика с целью повышения точности и адаптивности обнаружения. Традиционные сигнатурные подходы, применяемые в системах обнаружения вторжений, не способны распознавать новые или модифицированные типы атак, поскольку они основаны на заранее известных шаблонах вредоносного поведения. В отличие от них, методы машинного обучения позволяют выявлять ранее неизвестные аномалии за счёт анализа статистических закономерностей в сетевом трафике, что делает их особенно актуальными для современных систем защиты. Датасет CIC-DDoS2019 содержит более 80 типов сетевых атак, включая UDP Flood, SYN Flood, HTTP Flood и другие, что делает его эталонным для тестирования моделей обнаружения DDoS.

Список литературы

1. Назаров А.Ш., Ли И.Т. DDoS-атаки и средства защиты от них // Политехнический вестник. Серия: Интеллект. Инновации. Инвестиции. – 2023. – № 1(61). – С. 42–45.
2. Терновой О.С. Раннее обнаружение DDOS атак на основе статистического анализа // Перспективы развития информационных технологий. – 2011. – № 6. – С. 212–215.
3. Шелухин О.И., Полковников М.В. Исследование алгоритма Isolation Forest при бинарной классификации сетевых аномалий // Безопасные информационные технологии: Сборник трудов Десятой международной научно-технической конференции, Москва, 03–04 декабря 2019 года. – Москва: Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет), 2019. – С. 387–393
4. Ляпин А.Д. Сравнительный анализ алгоритмов SVM (Support Vector Machine) и RF (Random Forest) // Вестник науки. – 2025. – Т. 3, № 6(87). – С. 1816–1824.

УДК 004.056; 004.8

А.А. ВОРОБЬЕВА, М.Ю. ТОЛСТЫХ

Московский государственный лингвистический университет

ПРИМЕНЕНИЕ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ АНОМАЛЬНОГО СЕТЕВОГО ТРАФИКА, ГЕНЕРИРУЕМОГО БОТНЕТАМИ НОВОГО ПОКОЛЕНИЯ

Проводится исследование применимости методов глубокого обучения для выявления аномалий, ассоциированных с активностью ботнетов нового поколения. Рассматриваются ключевые архитектуры нейронных сетей, их потенциал в моделировании сложных паттернов сетевого поведения. Статья направлена на осмысление перспектив применения этих передовых технологий в контексте интеллектуального управления сетевой безопасностью.

Современные ботнет-угрозы эволюционировали до уровня, при котором их активность зачастую остается незамеченной для традиционных систем безопасности. Имитация обычного поведения, использование зашифрованных протоколов и сокрытие под стандартными сетевыми средствами делают их практически невидимыми для классических подходов, опирающихся на поиск известных сигнатур или оценку простых статистических показателей [1]. Также наблюдается тенденция к распределенной и децентрализованной архитектуре, которые повышают отказоустойчивость системы и усложняют процесс полного контроля над ботнетом [2].

Возникает потребность в разработке и внедрении более совершенных, самообучающихся методов анализа сетевого трафика. Здесь методы машинного обучения, и в особенности глубокое обучение (DL), демонстрируют свой значительный потенциал. Их способность выявлять сложные, нелинейные закономерности и автоматизированно извлекать релевантные признаки из больших объемов данных открывает новые горизонты для обнаружения аномалий в сфере сетевой безопасности.

Важнейшим аспектом является способность DL-моделей обучаться на «нормальном» сетевом поведении, а затем, на основе усвоенного представления, идентифицировать аномалии. Для анализа последовательного характера сетевого трафика, ключевую роль играют такие архитектуры, как рекуррентные нейронные сети (RNN), включая LSTM и GRU, которые умеют работать с последовательностями и долгосрочными зависимостями, а также сверточные нейронные сети (CNN), способные выявлять локальные паттерны. Автокодировщики

также демонстрируют потенциал, обучаясь реконструировать нормальный трафик и выявляя аномалии по ошибке реконструкции [2].

Для успешного применения этих моделей критически важна правильная подготовка данных: агрегация статистических показателей, преобразование байтовых последовательностей пакетов в числовые векторы и анализ трафика в рамках временных окон или сессий. Следовательно, сетевой трафик необходимо подготовить. Необходимо собрать достаточное количество данных, представляющих как нормальное поведение сети, так и, по возможности, примеры вредоносной активности. Затем информацию следует преобразовать: извлечь важные статистические показатели (частоту и размер пакетов, используемые порты, время между пакетами) и представить их в числовом виде.

Обучение сложных DL-моделей является ресурсоемким процессом, зачастую требующим использования специализированного оборудования, такого как графические процессоры (GPU). Это может создавать ограничения для практического применения в условиях недостаточного ресурсного обеспечения. Необходимым условием является также скорость обнаружения и реагирования. Система безопасности должна не только выявлять аномалии, но и делать это в режиме, близком к реальному времени, для своевременного предотвращения потенциального ущерба.

Актуальной проблемой остается минимизация ложноположительных срабатываний. Важно достичь высокой точности обнаружения при одновременном снижении числа ложных тревог, вызываемых легитимной, но нестандартной сетевой активностью.

Несмотря на перечисленные трудности, перспективы развития глубокого обучения в сфере кибербезопасности весьма позитивны. DL предлагает существенный потенциал для повышения эффективности систем сетевой безопасности благодаря своей способности выявлять сложные, нелинейные паттерны в сетевом трафике. Дальнейшее развитие и внедрение технологий обещает значительное повышение надежности обнаружения новейших ботнет-атак. Таким образом, глубокое обучение становится неотъемлемым элементом современной кибербезопасности, направленным на противодействие скрытым и адаптивным угрозам.

Список литературы

1. Зуев Н.В. Обнаружение аномалий сетевого трафика методом глубокого обучения. Программные продукты и системы / Software & Systems, 2021.

2. Abdulrahman Alzahrani. An Optimized Approach to Deep Learning for Botnet Detection and Classification for Cybersecurity in Internet of Things Environment. Computers, Materials and Continua. Volume 80, Issue 2, 15 August 2024, p. 2331–2349.

УДК 004.056

А.Е. ПАВЛОВ, Г.С. БРЕЖНЕВ, А.А. СЕРОВА

МИРЭА – Российский технологический университет, Москва

СОВРЕМЕННЫЕ СПОСОБЫ ЗАЩИТЫ ПРОГРАММНЫХ ИНТЕРФЕЙСОВ ОТ КИБЕРАТАК НАПРАВЛЕННЫХ НА РЕАЛИЗАЦИЮ УГРОЗ «ОТКАЗ В ОБСЛУЖИВАНИИ»

В работе рассматриваются современные методы борьбы с угрозами и методы защиты от угроз вида «отказ в обслуживании» для программных интерфейсов сложных распределённых информационных систем. Предлагается использование комбинированного метода на основе искусственного интеллекта для эффективной защиты, а также показан практический опыт его использования.

Программные интерфейсы – это специальный инструментарий для взаимодействия распределённых систем между собой. В настоящий момент существует огромное количество возможных атак для подрыва работы программных интерфейсов [1].

Масштабируемая распределённая инфраструктура. Этот способ пригоден для облачных решений за счёт использования дополнительных ресурсов и заключается в том, чтобы создать такую систему, которая динамически сможет увеличивать свои вычислительные ресурсы в ответ на растущую нагрузку [2]. Этот способ реализуется через 2 основных этапа масштабирования: при горизонтальном масштабировании система мониторинга засекает рост нагрузки и автоматически запускает новые экземпляры сервиса, равномерно распределяя трафик между большим количеством целей. При вертикальном масштабировании облачный провайдер предоставляет системе более мощные ресурсы, хотя и в ограниченном объёме. Таким образом система не теряет свою работоспособность даже в условиях интенсивной атаки и сохраняет полную автоматизацию процесса.

Ограничение количества запросов. Основной целю данного способа является предотвращение исчерпывания ресурсов системы одним источником, будь то злоумышленник или вышедший из-под контроля клиент. Данный способ легко обходится злоумышленникам благодаря осуществлению скоростных атак с разных IP-адресов.

Позитивная модель безопасности. Основой метода является создание эталонной модели легитимного поведения, при котором любой запрос, не соответствующий эталону, автоматически блокируется. Но разработчик может не учесть некоторые аспекты или ошибиться или слишком строгая

модель может блокировать легитимных пользователей. Помимо этого, злоумышленник может использовать украденные валидные учетные данные и делать запросы.

Использование средств фильтрации и анализа трафика на основе искусственного интеллекта и машинного обучения. На стадии обучения модели происходит анализ сотни параметров каждого запроса и сессий, после чего выявляются скрытые паттерны и корреляции. После обучения система начинает анализировать реальный трафик в режиме, близком к реальному времени, благодаря чему она становится более устойчивой и работоспособной [3]. Тем не менее, модель всё ещё может интерпретировать некоторые запросы как «внезапную популярность сервиса», а не как атаку. Поэтому данный метод рекомендуется комбинировать с другими.

Аутентификация и авторизация не могут полноценно решить проблему, связанную с атаками вида «отказ в обслуживании». Во-первых, злоумышленник может воспользоваться легитимными учётными данными из скомпрометированных баз данных. Во-вторых, злоумышленник может отправить множество запросов, тем самым перегрузив информационную систему [3].

Предлагается использование комбинированного метода на основе методов искусственного интеллекта с моделью прогнозирования количества запросов и нагрузки в зависимости от текущего межобъектного взаимодействия в распределённых информационных системах. Суть данного похода состоит в использовании анализа запросов и прогнозирования времени выполнения этих запросов. Если время стремится к критической точке, то запрос ограничивается согласно стандартным правилам модели защиты от угроз.

Таким образом, предложенное решение позволяет более эффективно защитить программные интерфейсы, в том числе и от угроз вида «отказ в обслуживании», и повысить общую кибербезопасность распределённых информационных систем.

Список литературы

1. Марков А.С. Важная веха в безопасности открытого программного обеспечения. Вопросы кибербезопасности. 2023, № 1(53), с. 2–12. DOI: http://dx.doi.org/10.21681/2311-3456-2023-1-2-12.

2. Милославская Н.Г., Толстой А.И. Управление информационной безопасностью. М.: НИЯУ МИФИ, 2020. – 536 с.

3. Когос К.Г., Финошин М.А. Перспективные подходы к обнаружению сетевых скрытых каналов. Безопасность информационных технологий, [S.l.], т. 28, № 2, с. 45–61, 2021. ISSN 2074-7136. DOI: http://dx.doi.org/10.26583/bit.2021.2.05. – EDN: QESDBW

УДК 004.056

А.Д. КОРЧУН, В.А. СТЕПАНОВА

МИРЭА – Российский технологический университет, Москва

СПОСОБЫ ПРОТИВОДЕЙСТВИЯ СОВРЕМЕННЫМ АТАКАМ ПО ОТКАЗУ В ОБСЛУЖИВАНИИ КЛАССА API-ФЛУД

Атаки API-флуд представляют собой особый тип распределённых атак отказа в обслуживании (DDoS), направленных на перегрузку интерфейсов прикладного программирования чрезмерным количеством запросов. В статье рассматриваются современные способы противодействия этим атакам с использованием комплексных мер, включая фильтрацию, аутентификацию, машинное обучение и масштабируемые инфраструктурные решения.

Введение

С ростом использования API в цифровых сервисах и интеграциях возрастает и уязвимость перед атаками, направленными на исчерпание их ресурсов путем массовой генерации автоматизированных запросов – API-флуд. Такие атаки приводят к снижению производительности, перебоям в работе сервисов и потере доверия пользователей. Эффективная защита требует междисциплинарного подхода и современных технологий [1, 2].

Для обеспечения устойчивости API-инфраструктуры к флудингу предлагается многоуровневая система защиты, ключевые элементы которой рассматриваются ниже.

1. Ограничение темпа запросов (rate limiting). Реализация лимитов количества запросов от одного клиента за единицу времени помогает предотвратить перегрузку API. Современные системы способны автоматически регулировать пороговые значения исходя из анализа трафика и текущей загруженности [3].

2. Аутентификация и авторизация. Использование токенов доступа (OAuth, JWT) и строгих политик контроля прав доступа исключает возможность проведения массовых запросов с неавторизованных источников, повышая уверенность в легитимности трафика [2].

3. Антибот-защита и анализ поведения. Внедрение антибот-систем с машинным обучением позволяет распознавать аномальные паттерны, характерные для автоматизированных атак, в том числе замаскированный под обычный трафик флуд [4].

4. Использование облачных анти-DDoS решений. Облачные платформы с распределённой архитектурой обеспечивают фильтрацию

больших объёмов трафика на границе сети, препятствуя достижению критической нагрузки на API-серверы [3, 4].

5. Масштабирование и балансировка нагрузки. Техническое расширение ресурсов и использование балансировщиков позволяет гасить всплески нагрузки, минимизируя вероятность отказа в обслуживании [5].

6. Шифрование и безопасность передачи данных. Применение HTTPS и других средств защиты каналов связи предотвращает посреднические атаки (MITM) и обеспечивает конфиденциальность [2, 3].

Регулярный мониторинг логов, активное тестирование API на уязвимости и постоянное обновление политик безопасности позволяют своевременно выявлять и нейтрализовать угрозы. Однако вызовами остаются сложности с точной идентификацией легитимного и злонамеренного трафика, высокие вычислительные затраты и потребность в адаптивных алгоритмах.

Заключение

В связи со сложностью проблемы защита от API-флуд требует комплексного, интегрированного подхода, сочетающего технические, аналитические и организационные меры. Современные технологии, включая машинное обучение и облачные анти-DDoS, значительно повышают устойчивость сервисов к современным видам атак, обеспечивая стабильность работы и безопасность пользователей, однако не один из приведенных выше методов на данный момент не предоставляет полной защиты, проблема требует дальнейшего изучения, которое будет включать в себя рассмотрение возможности использования искусственного интеллекта для повышения уровня защиты.

Список литературы

1. Защита сервера от DDoS-атак: методы и рекомендации // StormWall.pro. – 21 сентября 2025. – URL: https://stormwall.pro/resources/blog/zashchita-servera-ot-ddos-atak (дата обращения: 15.10.2025).

2. Как защитить API от DDoS-атак // StormWall.pro. – 20 апреля 2025. – URL: https://stormwall.pro/resources/blog/kak-zashchitit-api-ot-ddos-atak (дата обращения: 15.10.2025).

3. API на передовой: методы противостояния кибератакам // SecurityVision.ru. – 16 октября 2024. – URL: https://www.securityvision.ru/blog/api-na-peredovoy-metody-protivostoyaniya-kiberatakam/ (дата обращения: 15.10.2025).

4. Защита от DDoS-атак: как предотвратить отказ в обслуживании // Kurshub.ru. – 30 июля 2025. – URL: https://kurshub.ru/journal/blog/ddos-ataki-pochemu-ih-slozhno-ostanovit-i-kak-zashhititsya/ (дата обращения: 15.10.2025).

5. Архитектурные решения по однонаправленной передаче данных – Грачев А.С. // na-journal.ru – URL: https://na-journal.ru/4-2023-informacionnye-tekhnologii/4766-arhitekturnye-resheniya-po-odnonapravlennoi-peredache-dannyh (дата обращения: 15.10.2025).

Направление

Информационная безопасность социотехнических систем

Руководители секции:

ДВОРЯНКИН С.В. – д.т.н., профессор НИЯУ МИФИ

ВАНИЧКИНА А.С. – к.филол.н., директор ИИН МГЛУ

МИНАЕВ В.А. – д.т.н., профессор МосУ МВД России

УДК 004.056, 130.2

П.Г. БЫЛЕВСКИЙ
Московский государственный лингвистический университет

КУЛЬТУРА ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ЦИФРОВЫХ СОЦИОТЕХНИЧЕСКИХ СИСТЕМ

Развитие общегражданской культуры информационной безопасности отвечает характеру угроз и рисков, сопутствующих универсализации электронно-вычислительных технологий и интернет-коммуникаций, сопутствующих рисков и средств защиты. Формирование единой цифровой социотехнической среды «интернета вещей» сопровождается превращением всех граждан в многочасовых вовлечённых пользователей и постоянные объекты генерации «больших данных». Безопасное использование цифровых интернет-сервисов требует соответствующей общегражданской культуры, соответствующей особенностям профессии, быта и в целом образа жизни каждого гражданина. Необходимы государственные меры по формированию у всех граждан понимания рисков интернет-коммуникаций, умения распознавать угрозы, владения средствами самозащиты и готовности их применять.

Становление общегражданской культуры информационной безопасности является закономерным результатом предшествующей эволюции средств самозащиты граждан от угроз в публичных интернет-коммуникациях. Развитие культуры информационной безопасности должно соответствовать характеру угроз и рисков, сопутствующих цифровым социотехническим системам, универсализации электронно-вычислительных технологий и интернет-коммуникаций. Формирование единой цифровой социотехнической среды «интернета вещей» сопровождается превращением всех граждан в многочасовых вовлечённых пользователей и постоянные объекты генерации «больших данных». Безопасное использование цифровых интернет-сервисов требует общегражданской культуры, соответствующей особенностям профессии, быта и в целом образа жизни каждого гражданина.

Каждый гражданин должен знать и уметь минимизировать риски, связанные с возможностями хозяев цифровых сервисов (социальных сетей, интернет-мессенджеров и «интернета вещей») дистанционно определять поведение и влиять на него. Важнейшими являются умения отбирать из массы сведений достоверные [1], критически относиться к сообщениям, способным повлечь массовые панические настроения. Особое значение приобретают системные профилактические меры: информирование граждан о существующих угрозах и о мерах по защите [2],

организация эффективного «всеобуча» в области защиты информации в цифровых социотехнических системах сбора, обработки, передачи и хранения данных о пользователях [3], повышение уровня правовой культуры и компьютерной грамотности граждан.

Общегражданский характер культуры информационной безопасности предполагает последовательный охват всех возрастных групп, чтобы обеспечивать безопасность несовершеннолетних в цифровом пространстве и цифровую социализацию школьника, нейтрализуя издержки самостоятельного и стихийного освоения ребенком интернет-ресурсов в качестве развлекательного контента. Включать формирование компетенции информационной безопасности в образовательный процесс рекомендуется средним специальным и высшим учебным заведениям, далее в систему непрерывного образования, повышения квалификации, профессиональной переподготовки. Развивать культуру информационной безопасности граждан должны помогать социальная реклама и сотрудники интернет-сервисов, способствуя преодолению клиентами незнания правил безопасности, обусловленного некритичным восприятием интернета как безопасного места. Необходимы совершенствование и реализация государственной политики формирования у всех граждан понимания рисков интернет-коммуникаций, умения распознавать угрозы, владения средствами самозащиты и готовности их применять.

Список литературы

1. Былевский П.Г. Библиографическая культура как фактор безопасности доверенного публичного Интернета // Обсерватория культуры. 2024. Т. 21. № 4. С. 358–366. DOI: 10.25281/2072-3156-2024-21-4-358-366. EDN: SRCWDS.

2. Завьялов А.Н. Интернет-мошенничество (фишинг): проблемы противодействия и предупреждения // Baikal Research Journal. 2022. Т. 13. № 2. С. 4. DOI: 10.17150/2411-6262.2022.13(2).36. EDN: SRVHGS.

3. Малюк А.А., Малюк З.П. Актуальные вопросы создания системы массового обучения культуре информационной безопасности // Безопасность информационных технологий. 2021. Т. 28. № 4. С. 6. DOI:10.26583/bit.2021.4.01. EDN: XHPGMY.

УДК 004.9

Р.А. ОЛИН[1], М.Е. ФЕДИНА[1], В.Ю. ЖИВЦОВ[2]

[1]Самарский национальный исследовательский университет им. академика С.П. Королева

[2]Самарский государственный медицинский университет

УГРОЗЫ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В МЕТАВСЕЛЕННЫХ

В работе рассматриваются угрозы информационной безопасности в метавселенных, возникающие вследствие интеграции современных технологий. Уделено внимание рискам компрометации личности и появлению новых векторов атак, предложены направления повышения устойчивости систем за счёт внедрения принципов нулевого доверия, децентрализованных механизмов аутентификации и локальной обработки чувствительных данных.

Метавселенные представляют собой цифровую среду с совокупностью взаимосвязанных виртуальных пространств, созданных на основе технологий виртуальной (VR), дополненной (AR) и смешанной реальности (MR), искусственного интеллекта, интернета вещей (IoT) и блокчейна, в которых пользователи взаимодействуют в режиме реального времени с окружающими объектами и друг с другом. Интеграция разнородных технологий существенно расширяет поверхность возможных кибератак.

Визуально-зрительная природа взаимодействия в VR/AR формирует особый класс угроз, обусловленных высокой степенью иммерсивности и интеграцией пользователей. Устройства уязвимы к атакам по сторонним каналам, позволяющим извлекать конфиденциальную информацию из сенсорных данных [1]. Кроме компрометации приватности, непосредственное воздействие на чувства и восприятие пользователя создает возможность физического вреда, атакующие могут вызвать тошноту или дезориентацию.

Платформы метавселенных аккумулируют большие объемы биометрических данных и поведенческих паттернов. Кроме прямых технических уязвимостей, существуют системные риски утечки аудио- и видеозаписей из частных виртуальных встреч.

Особое значение приобретает проблема достоверности личности, рождая ряд угроз, связанных с компрометацией профилей участников системы. Получив контроль над учетной записью, злоумышленник может

распоряжаться цифровыми активами, общаться с другими людьми от имени жертвы и совершать другие противоправные действия.

Развитие технологии дипфейков позволяет сравнительно легко через поддельный аватар осуществлять подмену личности. Традиционные протоколы аутентификации (пароли, двухфакторная авторизация) не способны защитить от маскировки под законного пользователя.

Среди технических угроз отмечаются удаленное выполнение кода, аппаратные уязвимости IoT-компонентов, атаки на сетевую инфраструктуру, отказ в обслуживании и уязвимости в протоколах. Эксплойты на уровне программных интерфейсов или драйверов видеоадаптера способны привести к атакам через специально сконструированные 3D-модели и скрипты внутри виртуальной сцены для исполнения произвольного кода на клиентском устройстве [2].

Анализ угроз информационной безопасности в метавселенных показывает сложный, многоуровневый ландшафт рисков, в связи с чем необходим комплексный подход обеспечения информационной безопасности. Изначальное проектирование систем должно предполагать механизмы разграничения доступа, минимизацию сбора чувствительных данных с преимущественно локальной обработкой и применение средств криптографии. Многофакторная аутентификация и верификация аватаров позволит подтверждать подлинность личности при виртуальном взаимодействии [3].

Также целесообразно внедрять принципы нулевого доверия и децентрализованные блокчейн-технологии, что обеспечит распределенную аутентификацию и защиту от компрометации единой точки отказа.

Разнообразие угроз в метавселенных требует особого подхода к оценке рисков с учётом стремительного развития технологий генеративного искусственного интеллекта.

Список литературы

1. Chow Y., Susilo W., Nguyen C. Visualization and Cybersecurity in the Metaverse: A Survey [Электронный ресурс] // J. Imaging. – 2023. – № 1. – DOI: 10.3390/jimaging9010011. – URL: https://www.mdpi.com/2313-433X/9/1/11 (дата обращения: 15.09.2025).

2. Dey K., Barradas D., Hakak S. A Systematic Literature Review on Fundamental Technologies and Security Challenges in the Metaverse Platforms [Электронный ресурс] – arXiv, 2025. – № 2510.09621. – URL: https://arxiv.org/abs/2510.09621 (дата обращения: 23.09.2025).

3. Laiz-Ibáñez H., Mendaña-Cuervo C., Carus-Candas J. The metaverse: Privacy and information security risks [Электронный ресурс] – International Journal of Information Management Data Insights, 2025. – Vol. 5, № 2. – DOI: 10.1016/j.jjimei.2025.100373. – URL:

УДК 519.72; 004.056.5

В.П. ЦВЕТОВ

Самарский национальный исследовательский университет им. академика С.П. Королева

О КВАНТОВОЙ МОДЕЛИ ИНФОРМАЦИИ

В докладе рассматривается энергетическая модель меры объема информации – бита.

Представление информации в виде данных определяет квантовый характер такой формы представления. Хотя информация определяется способом интерпретации структур данных, сами данные составляют основу информационного обмена. Данные передаются порциями, кратными числу символов используемого языка, в простейшем случае – языка над двоичным алфавитом, – битами. При моделировании процессов передачи, хранения и обработки информации в терминах энергетического обмена, одной из значимых характеристик подобных моделей становится количество энергии, затраченной на представление 1 бита данных, – E_b (Дж/бит). Моделирование процесса передачи данных опирается на понятия источника данных и его энтропии – H (бит); канала, его пропускной способности – C (бит/с) и скорости передачи – $R \leq C$ (бит/с) [1].

В качестве предельного случая модели системы передачи данных рассмотрим модель однофотонных излучателя/приемника с кодово-импульсной модуляцией на канальном уровне. Обозначим через ν частоту излучаемого кванта, в таком случае длительность импульса будет равна 1/ν, межсимвольные интервалы будем считать равными нулю. Будем кодировать 1 одиночным импульсом, 0 – отсутствием импульса. Так как энергия одного кванта излучения равна $h\nu$, где h – постоянная Планка, то в случае источника с наибольшей энтропией для передачи одного бита потребуется энергия, $E=h\nu/2$ (Дж). Для согласования размерностей обозначим $E_b=E/1$ бит (Дж/бит), $\nu_b=\nu/1$ бит (Гц/бит), тогда $E_b=h\nu_b/2$. При отсутствии помех, время на передачу одного бита данных будет равно длительности импульса $T=1/\nu$. Как и ранее обозначим $T_b=T/1$ бит (с/бит), тогда скорость передачи данных в такой системе будет равна $R=C=1/T_b$ (бит/с). Введем размерный множитель $\chi=1/\text{бит}^2$, тогда $\nu_b=\chi/T_b$, и $E_b=\chi\, hR/2=\chi\, hC/2$.

Если в системе применяются K-фотонные излучатель/приемник фотонов частоты ν, то, очевидно, $E_b=\chi\, KhR/2=\chi\, KhC/2$.

Если система использует n уровней квантования от 0 до n-1 с шагом квантования K, то при отсутствии помех и наибольшей энтропии источника $E_b=\chi\, K(n\text{-}1)hR/\log(n)=\chi\, K(n\text{-}1)hC/\log(n)$.

В рассматриваемой ситуации помехи на канале могут возникать при распространении квантов в поглощающих/излучающих средах, тогда $E_b=\chi\, K(n\text{-}1)hR/\log(n)\leq\chi\, K(n\text{-}1)hC/\log(n)$.

Полученные выражения определяют энергию, необходимую для передачи/приема 1 бита, которую можно назвать его динамической энергией. Наибольшее значение динамической энергии бита в рассматриваемой модели будет достигаться на планковском времени $T_p=t_p/1$ бит, при $R=1/T_p$. Под статической энергией бита, вероятно, следует понимать значение, вытекающее из принципа Ландауэра [2], и задаваемое формулой Шеннона-фон Неймана-Ландауэра $E_b^{\,0}=k_b\, T_0\,/\log(e)$, где k_b – постоянная Больцмана, e – основание натурального логарифма T_0 – температура идеального средства хранения, позволяющего различать два состояния.

Внесение в предложенную модель неопределенности за счет контролируемого изменения помехообразующих параметров среды распространения квантов может оказаться целесообразным при разработке вычислительных устройств, реализующих вероятностные алгоритмы, и аналоговых устройств, моделирующих стохастические процессы. Подобные устройства, как правило, находят полезное применение в областях, которые связаны с эффективной обработкой и защитой информации [3].

Список литературы

1. Шеннон К. Работы по теории информации и кибернетике. – М.: Изд. Иностр. лит., 1963. – 830 с.

2. Rolf Landauer. Irreversibility and heat generation in the computing process / IBM Journal of Research and Development, vol. 5, pp. 183–191, 1961.

3. Nielsen, Michael A.; Chuang, Isaac L. (2000). Quantum Computation and Quantum Information (1st ed.). Cambridge: Cambridge University Press. ISBN 978-0-521-63503-5.

УДК 004.056

А.Ж. НИЗАМОВ[1], А.М. НИЗАМОВА[2]

[1]*Финансовый университет при Правительстве России, Москва*
[2]*Национальный исследовательский ядерный университет «МИФИ», Москва*

СОЦИОТЕХНИЧЕСКИЙ ПОДХОД К ЗАЩИТЕ ПЕРСОНАЛЬНОЙ ИНФОРМАЦИИ С ПРИМЕНЕНИЕМ ИСКУССТВЕННОГО ИНТЕЛЕКТА

Искусственный интеллект (ИИ) стал неотъемлемой частью цифрового общества. ИИ обрабатывает массивы данных, в том числе персональных и биометрических, что делает задачу обеспечения приватности крайне важно [1, 2].

Дополнительную угрозу представляют когнитивные воздействия, создаваемые ИИ. Которые могут включать вредоносные закладки, скрытые алгоритмы сбора данных или механизмы влияния на мышление пользователей.

Введение

Искусственный интеллект (ИИ) ускоряет анализ и принятие решений в образовании, здравоохранении и управлении, но повышает риски нарушения приватности. К персональным данным относятся сведения, позволяющие идентифицировать человека, включая биометрические и поведенческие.

Использование ИИ из недружественных юрисдикций создаёт угрозы: возможные закладки (backdoors), скрытая телеметрия и когнитивные манипуляции через искажённые ответы и рекомендации.

Постановка задачи

Основная цель исследования – определить методы и подходы к защите персональной информации, а также выработать критерии выбора безопасных систем для анализа данных. Для этого необходимо решить следующие задачи:

1. Определить угрозы, связанные с использованием зарубежных ИИ-платформ, включая риск утечек и когнитивных воздействий.
2. Описать методы защиты персональной информации на основе отечественных технологий.
3. Установить, какие ИИ целесообразно использовать для получения количественных данных и проведения вычислительных экспериментов.
4. Провести сравнительную оценку ИИ-платформ по скорости отклика и надёжности защиты данных.

Пути решения задачи

Для решения предлагается использовать комплексный подход:

1. Применение отечественных ИИ-систем. Использование российских решений (например, ЯндексGPT, SberAI, ruGPT, Kandinsky, GigaChat) снижает риски.

2. Выбор ИИ для расчётов и анализа данных.

Для проведения статистических расчётов целесообразно использовать:

– Python-библиотеки с интегрированными моделями машинного обучения (scikit-learn, NumPy, PyTorch);

– отечественные решения на базе СберКлауд и Яндекс Облака;

– локальные развертывания моделей LLM для автономной аналитики без передачи информации в сеть.

3. Методы защиты персональных данных.

– Применение принципов privacy by design (встроенная защита данных при проектировании);

– Federated learning – распределённое обучение без передачи исходных данных;

– Гомоморфное шифрование для безопасных вычислений над зашифрованной информацией.

4. Оценка оперативности и надёжности ИИ.

Эффективность ИИ для защиты информации определяется не только уровнем безопасности, но и скоростью реакции на запрос. Средняя оперативность отечественных ИИ (ЯндексGPT, GigaChat) составляет 2–6 с при простых запросах и до 20 с при сложных.

Заключение

Использование ИИ должно сопровождаться созданием безопасной цифровой среды, где персональные данные защищены от утечек, а когнитивное воздействие сведено к минимуму. Реализация принципов privacy by design, внедрение отечественных ИИ- и обучение специалистов —условия информационной независимости России.

Список литературы

1. Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации. Цифровая безопасность. URL: https://digital.gov.ru/ru/activity/directions/1003./

2. Stanford Institute for Human-Centered Artificial Intelligence (HAI). Privacy in an AI Era: How Do We Protect Our Personal Information? — Stanford University, 2023. URL: https://hai.stanford.edu/news/privacy-ai-era-how-do-we-protect-our-personal-information.

УДК 004.056

А.С. ЖИДКОВ

Национальный исследовательский университет «Высшая школа экономики», Москва

ПРЕДОТВРАЩЕНИЕ УТЕЧКИ РЕЧЕВОЙ ИНФОРМАЦИИ ЧЕРЕЗ ДАННЫЕ АКСЕЛЕРОМЕТРА СМАРТФОНА

В работе рассмотрены угрозы несанкционированной утечки речевой информации в современных смартфонах через сигнал акселерометра. Показано, что используя данные акселерометра можно распознать речевые сигналы, воспроизводимые смартфоном, с достаточно высокой точностью. Предложен алгоритм предобработки сигналов акселерометра, позволяющий предотвратить утечку речевой информации. На примере обработки данных нескольких моделей смартфонов показана практическая эффективность предложенного алгоритма.

Исследования последних лет показали [1, 2], что данные акселерометра современного смартфона могут быть использованы для несанкционированного восстановления речевой информации, которая воспроизводится динамиком смартфона. Для минимизации риска утечки речевой информации через акселерометр в существующих работах предлагается ограничение частоты дискретизации сигналов акселерометра до 100 Гц или ниже. Однако такое ограничение зачастую приводит к ухудшению потребительских свойств мобильных устройств, так как в ряде приложений для повышения точности требуются частоты дискретизации до 400 Гц, а иногда выше.

В настоящей работе изучена модель тракта динамик-акселерометр смартфона и на ее основе предложен адаптивный алгоритм компенсации паразитного сигнала, приводящего к утечке речевой информации. Предложенный алгоритм представлен на рис. 1 и представляет собой адаптивный фильтр наименьших средних квадратов с безинерционным нелинейным преобразователем логистического вида на входе; адаптация фильтра проводилась в моменты неподвижности устройства.

Для распознавания речевой информации по сигналам акселерометра использовалась открытая модель распознавания речи Whisper [3] дообученная на синтетических данных, в качестве которых использовался датасет TIMIT [4] пропущенный через модель тракта динамик-акселерометр нескольких смартфонов (всего 13395 записей). Для оценки качества восстановления набор из 118 тестовых фраз воспроизводился на смартфонах трех моделей при громкости 80% и одновременно с этим записывались данные акселерометра с максимально возможной частотой

дискретизации. Качество восстановления оценивалось по критерию вероятности ошибки на слово (WER). Результаты распознавания для трех моделей смартфонов без компенсации и с предложенной схемой компенсации приведены в табл. 1.

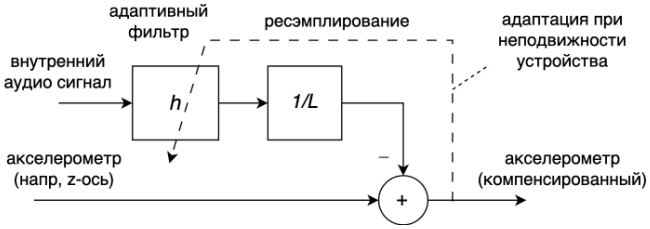

Рис. 1. Предложенный алгоритм компенсации паразитного сигнала

Таблица 1. Результаты распознавания речевых сигналов

Смартфон	F_{acc}	WER без компенсации	WER с компенсацией
Samsung A52	861 Гц	29.4%	89.9%
Honor 50	1000 Гц	46.5%	84.6%
Samsung S24Ultra	474 Гц	55.7%	88.9%

В [5] указано, что большинство людей понимают общий контекст разговора при WER<80%. Наши результаты показывают, что без какой-либо компенсации во всех исследованных смартфонах можно восстановить речь по данным акселерометра с WER равном 56% или ниже. Однако при использовании предложенного метода компенсации WER возрастает до 84% и выше, что существенно минимизирует риск утечки информации через сигнал акселерометра.

Список литературы

1. R. De Prisco, A. De Santis, D. Malandrino, R. Zaccagnino. An improved privacy attack on smartphones exploiting the accelerometer, Journal of Information Security and Applications, Vol. 75, 2023, DOI: https://doi.org/10.1016/j.jisa.2023.10347.

2. Жидков А. С. Восстановление речи по данным акселерометра смартфона на базе вокодерной модели речевых сигналов //Материалы научно-технической конференции «Микроэлектроника и информатика» – 2025, с. 228. – ISBN 978-5-7256-1062-8.

3. A. Radford, et al. Robust speech recognition via large-scale weak supervision //40th International Conference on Machine Learning (ICML'23), Vol. 202. pp. 28492–28518.

4. J. Garofolo, et al. TIMIT Acoustic-phonetic Continuous Speech Corpus. Linguistic Data Consortium, 1992, DOI: https://doi.org/10.35111/17gk-bn40.

5. S. Boovaraghavan, H. Zhou, M. Goel, Y. Agarwal. 2024. Kirigami: Lightweight Speech Filtering for Privacy-Preserving Activity Recognition using Audio //Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 8, 1, Article 36, March 2024, DOI: https://doi.org/10.1145/3643502.

УДК 004.056

Е.С. АНИСИМОВ

Финансовый университет при Правительстве Российской Федерации, Москва

ПОДХОДЫ К ОПРЕДЕЛЕНИЮ КАНАЛА ПЕРЕДАЧИ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ СТЕГАНОГРАФИЧЕСКИХ ПРЕОБРАЗОВАНИЙ

В настоящее время наблюдается отсутствие единого подхода к определению канала передачи данных с использованием стеганографических преобразований. В докладе представлены основные подходы с обозначением положения данного канала по отношению к скрытым каналам и каналам скрытой связи.

Стеганографические преобразования используются как для защиты цифрового контента от НСД и неправомерного использования, так и для сокрытия некоторого взаимодействия между сторонами, как правило нарушающего политику безопасности или законодательство.

В качестве одного из разделов стеганографии часто выделяют цифровые водяные знаки (ЦВЗ). Поскольку при формировании ЦВЗ использование стеганографических преобразований сводится к разовому действию, не предполагающему передачу данных с полезной смысловой нагрузкой, то в отличие от случая применения стеганографии для сокрытия взаимодействия, речь не идёт о полноценном канале передачи данных.

Такой канал передачи данных допустимо называть стеганографическим в соответствии с его природой (стеганоканал).

Таким образом, под стеганоканалом будем понимать канал взаимодействия с использованием стеганографических преобразований. Подобный подход к пониманию стеганоканала представляется, в частности, в работе [1].

Альтернативная формулировка стеганоканала через методы и средства создания стеганографической связи между двумя узлами представлена в работе [2]. Там же отмечена относительная новизна использования на тот момент термина стеганоканал (стегоканал в оригинале источника), поскольку в ранних работах для его обозначения употреблялось словосочетание «стеганографическая система».

Близкими к стеганоканалу терминами являются: скрытый канал передачи данных, подпороговый канал, нетрадиционный информационный канал, канал скрытой связи и канал скрытых коммуникаций. Характер их упоминания наряду с применением стеганографических преобразований различный: в качестве синонима

стеганоканала; в роли частного случая стеганоканала; или же в статусе обобщающего класса каналов, частным случаем которого являются стеганоканалы. В англоязычных источниках среди основных терминов используются: steganographic channel, hidden channel, hidden communication channel, covert channel, subliminal channel.

Допустимо отметить некоторые предлагаемые в различных источниках подходы к отношению стеганоканала со связанными классами каналов:

1. Стеганоканал – частный случай нетрадиционного информационного канала (используются специальные, не криптографические, преобразования передаваемых данных) [3, 4].
2. Стеганоканал – частный случай скрытого канала [5].
3. Скрытый и подпороговый каналы – частные случаи стеганоканала [1].
4. Предлагается иной термин – скрытый стеганографический канал [6].
5. Скрытый канал – отдельный термин, не связанный со стеганоканалами [7].
6. Стеганоканал – частный случай канала скрытой связи [8, 9].

Таким образом, представлены основные подходы к определению стеганоканала, среди которых наиболее признанным и предпочтительным является подход, указанный в пунктах 5 и 6.

Список литературы

1. Варновский Н.П., Голубев Е.А., Логачёв О.А. Современные направления стеганографии. Математика и безопасность информационных технологий. Материалы конференции в МГУ 28–29 октября 2004 г. М.: МЦНМО, 2005, с. 32–64.

2. Макаренко С.И. Эталонная модель взаимодействия стеганографических систем и обоснование на ее основе новых направлений развития теории стеганографии. Вопросы кибербезопасности. 2014. № 2 (3). С. 24–32.

3. Базовая модель угроз безопасности персональных данных при их обработке в информационных системах персональных данных (выписка). ФСТЭК России.

4. Приказ Федеральной налоговой службы от 21 декабря 2011 г. № ММВ-7-4/959@ «Об обеспечении безопасности персональных данных при их обработке в автоматизированных информационных системах налоговых органов».

5. Солодуха Р.А. Концепция формирования системы противодействия стеганографическим каналам в компьютерных сетях органов внутренних дел. Вестник Воронежского института МВД России. 2021, № 1, с. 131–142.

6. Авсентьев О.С., Бутов В.В., Цыганов К.А. Функциональная модель процесса реализации угроз безопасности информации с использованием скрытых стеганографических каналов внешним нарушителем. Безопасность информационных технологий. 2025, № 2, с. 83–101.

7. ГОСТ Р 53113.1-2008. Информационная технология. Защита информационных технологий и автоматизированных систем от угроз информационной безопасности, реализуемых с использованием скрытых каналов. Часть 1. Общие положения.

8. Голубев Е.А. Стеганографические технологии – новое направление защиты информации. T-Comm: Телекоммуникации и транспорт. 2012. № 6, с. 49–53.

9. Семёнов К.П., Зайцев П.В. Оценки стойкости стеганографических систем и условия их достижения. Информационная безопасность регионов. 2014. № 3 (16), с. 29–34.

УДК 004.056

А.С. МАМОНТОВ[1], В.Л. ЕВСЕЕВ[1,2], А.С. БУРАКОВ[1], В.С. МАМОНТОВ[3]

[1]Московский государственный лингвистический университет
[2]Национальный исследовательский ядерный университет «МИФИ», Москва
[3]ООО СК «Сбербанк страхование», Москва

АВТОМАТИЗИРОВАННЫЙ КОНТРОЛЬ КОДА С ПОМОЩЬЮ СТАТИЧЕСКОГО АНАЛИЗА В QUALITY GATE: ОБНАРУЖЕНИЕ УЯЗВИМОСТЕЙ И БЛОКИРОВКА НЕБЕЗОПАСНОГО КОДА ДЛЯ ЗАЩИТЫ ПОЛЬЗОВАТЕЛЕЙ И СОЦИУМА

Целью работы является исследование инструментов статического анализа (SAST) в Quality Gates для предотвращения внедрения вредоносного программного обеспечения (ПО), уязвимых библиотек и фишинговых механизмов, угрожающих финансовой безопасности и приватности пользователей. Проанализированы методы выявления скрытых угроз: трояны в зависимостях, поддельные интерфейсы, эксплойты. Особое внимание уделено блокировке кода, нарушающего этические нормы, и интеграции SAST с анализом зависимостей (SCA) для исключения компонентов с критическими уязвимостями [1] Исследованы критерии защиты от социальной инженерии, устаревших библиотек, минимизирующие риски утечек данных и мошеннические схемы.

Введение

Современная разработка ПО требует этической ответственности разработчиков: уязвимости и фишинговые элементы угрожают пользователям и могут привести к потере данных. Интеграция SAST и SCA в Quality Gates позволяет не только выявлять классические уязвимости (XSS, SQL-инъекции), но и блокировать подозрительные алгоритмы и поддельные SDK, сохраняя доверие клиентов и защищая их от мошенничества [2].

Постановка задачи

Система автоматизированного контроля кода с помощью статического анализа в Quality Gate позволяет решать задачи:

– обнаружение уязвимостей в исходном коде: SAST-инструменты анализируют исходный код приложения без его выполнения, выявляя потенциальные уязвимости безопасности, такие как SQL-инъекции, межсайтовый скриптинг (XSS), небезопасная работа с данными и другие.

– автоматическая блокировка небезопасного кода: интеграция SAST в Quality Gates позволяет автоматически блокировать сборку и развертывание приложения, если обнаружены критические уязвимости. Это предотвращает попадание небезопасного кода в продуктивную среду и обеспечивает соблюдение политик безопасности организации;

– оценка рисков и приоритизация уязвимостей: SAST-инструменты предоставляют детальную информацию о типе, критичности и потенциальном влиянии обнаруженных уязвимостей. Это позволяет команде разработчиков эффективно расставлять приоритеты при устранении проблем безопасности и фокусироваться на наиболее критичных угрозах.

Пути решения задачи

Для решения задач автоматизированного контроля кода при разработке ПО целесообразно внедрять инструменты статического анализа безопасности приложений (SAST) в системы Quality Gates. После анализа существующих решений в области автоматизации процесса обнаружения уязвимостей определяются наиболее универсальные и эффективные инструменты. Основное внимание уделяется возможности интеграции с CI/CD, точности обнаружения уязвимостей, скорости анализа, поддержке различных языков программирования и стоимости внедрения.

По итогам анализа предпочтение в использовании целесообразно отдать: 1. SonarQube; 2. PT Application Inspector; 3. Checkmarx SAST.

Заключение

Системы автоматизированного контроля кода с помощью статического анализа, такие как SonarQube, PT Application Inspector и Checkmarx SAST, играют критическую роль в обеспечении кибербезопасности современных программных продуктов. Интеграция SAST-инструментов в Quality Gates позволяет автоматически выявлять уязвимости в исходном коде на ранних этапах разработки и блокировать внедрение небезопасного кода в продуктивную среду.

Список литературы

1. Марков А.С. Важная веха в безопасности открытого программного обеспечения. Вопросы кибербезопасности. 2023, № 1(53), с. 2–12. DOI: http://dx.doi.org/10.21681/2311-3456-2023-1-2-12.

2. Шинкарев А.А. Роль программного обеспечения с открытым исходным кодом в современной разработке корпоративных информационных систем // Компьютерные и информационные науки. – 2021. – № 1. DOI: 10.14529/ctcr210202

УДК 004.056

В.Л. ЕВСЕЕВ[1, 2], А.С. БУРАКОВ[2], А.С. МАМОНТОВ[2]

[1]*Национальный исследовательский ядерный университет «МИФИ», Москва*
[2]*Московский государственный лингвистический университет*

ПОВЫШЕНИЕ БЕЗОПАСНОСТИ ВЕБ-СЕРВИСОВ С ПОМОЩЬЮ СИСТЕМЫ АВТОРИЗАЦИИ АВТОМАТИЗИРОВАННЫХ КЛИЕНТОВ

Рассмотрена проблема автоматизированного бот-трафика в интернете. Обоснована актуальность данной проблемы, связанной с кибератаками, мошенничеством и искажением аналитики. Проанализированы традиционные методы идентификации легитимных ботов, такие как обратный DNS-запрос, и выявлены их недостатки. Предложено принципиально новое решение на основе системы обязательной авторизации автоматизированных клиентов с использованием криптографических токенов, обеспечивающее 100% точность идентификации доверенных ботов.

Введение

Современный интернет столкнулся с масштабной проблемой автоматизированного бот-трафика. Злонамеренные боты используются для организации DDoS-атак, кражи конфиденциальных данных, скальпинга, кликфрода и распространения спама, нанося значительный экономический ущерб [1]. В контексте безопасности социотехнических систем проблема бот-трафика требует комплексного подхода, учитывающего как технологические, так и организационные аспекты. В качестве базового метода противодействия часто предлагается идентификация легитимных ботов, таких как поисковые роботы, боты мониторинга, агрегаторы и архиваторы. Однако их присутствие осложняет задачу фильтрации вредоносного трафика. Злоумышленники могут легко подделать идентификатор User-Agent легального бота. Это требует применения более надежных методов верификации.

Текущий подход к решению проблемы

Наиболее распространенным методом верификации легитимных ботов является обратный DNS-запрос (rDNS) [2]. Данный метод позволяет проверить, зарегистрирован ли IP-адрес, с которого поступил запрос, на официального владельца бота. Например, IP-адрес настоящего Googlebot должен иметь PTR-запись, содержащую домен googlebot.com. Однако данный подход имеет существенные ограничения. Многие поставщики

полезных бот-услуг, такие как сервисы мониторинга, агрегаторы не публикуют полные и актуальные списки своих IP-адресов, а их rDNS-записи могут быть неинформативными. Кроме того, даже проверенный IP-адрес крупной компании может быть скомпрометирован и использован злоумышленником. Следовательно, проверки только по rDNS недостаточно для надежной сегментации трафика.

Предлагаемое решение

В качестве принципиально нового подхода к решению проблемы идентификации бот-трафика предлагается внедрение системы обязательной авторизации для автоматизированных клиентов через механизм криптографических токенов. В отличие от пассивных методов анализа, данный подход предполагает активное управление доступом, при котором право на взаимодействие с сервером предоставляется только доверенным субъектам, прошедшим предварительную регистрацию [3]. Архитектура системы включает сервис регистрации и выдачи токенов, шлюз аутентификации и модуль валидации. Протокол взаимодействия реализуется через верификацию поставщика бот-услуг, получение ботом криптографического токена и обязательное включение этого токена в заголовки HTTP каждого запроса.

Заключение

Таким образом, предложенная система активной авторизации представляет собой принципиально новый подход к решению проблемы идентификации бот-трафика. Использование криптографических токенов позволяет перейти от пассивной фильтрации к активному управлению доступом и обеспечивает надежную защиту веб-сервисов от несанкционированного автоматизированного трафика.

Список литературы

1. Москаленко А.А., Лапонина О.Р., Сухомлин В.А. Разработка приложения веб-скрапинга с возможностями обхода блокировок //Современные информационные технологии и ИТ-образование. – 2019. – Т. 15. – №. 2. – С. 413–420.

2. Андриченко Н.А., Лубова Е.С. Что такое DNS и как это работает? //Евразийский научный журнал. – 2018. – №. 4. – С. 41–43.

3. Макаров Д.А. Механизм авторизации с использованием технологии JWT //Теория и практика современной науки. – 2020. – №. 1 (55). – С. 474–476.

УДК 004.056

В.М. НИКИТЮК
Научный руководитель – к.т.н., доцент В.Л. ЕВСЕЕВ
Национальный исследовательский ядерный университет «МИФИ», Москва

ПРИМЕНЕНИЕ DLP ДЛЯ ПРЕДОТВРАЩЕНИЯ УТЕЧЕК ЧЕРЕЗ СОЦИОТЕХНИЧЕСКИЕ КАНАЛЫ: РОЛЬ ФИШИНГА И ЧЕЛОВЕЧЕСКОГО ФАКТОРА

В докладе рассмотрена проблема утечек данных через социотехнические каналы в корпоративных сетях, где фишинг выступает триггером компрометации. Основное внимание уделено социальным аспектам защиты (обучение, культура репортинга, поведенческие интервенции) с использованием DLP как вспомогательного технического контура для предотвращения эксфильтрации данных после успешной социальной манипуляции.

Введение

Фишинг остаётся наиболее массовым инструментом социального воздействия на сотрудников, инициируя компрометацию учётных данных и запуская цепочку эксфильтрации данных. Исследования показывают, что более 60% утечек связаны с действиями инсайдеров, причём психологические и организационные факторы (стресс, неудовлетворённость, недостаток норм безопасности) существенно влияют на уязвимость персонала к социальным манипуляциям. Традиционные технические меры улавливают угрозы на поздних стадиях, когда данные уже скомпрометированы. В связи с этим актуальна разработка социотехнического подхода, объединяющего поведенческие интервенции с минимально достаточными DLP-настройками для раннего предупреждения и блокирования утечек после фишингового «триггера» [1].

Постановка задачи

Задача состоит в разработке практико-ориентированного социотехнического фреймворка, который связывает поведенческие и организационные интервенции (просвещение, культура репортинга, поведенческие «подталкивания») с DLP-контекстными правилами для предотвращения утечек данных на ранних стадиях. Необходимо формализовать связку «фишинговые тактики убеждения → поведенческие индикаторы – DLP-контекстные правила» и разработать метрики оценки, ориентированные на снижение человеческой уязвимости и последующего ущерба от утечек.

Пути решения задачи

Социальный контур. Фишинг эксплуатирует социальные эвристики (авторитет, срочность), снижая критичность восприятия пользователей. Для противодействия внедряются регулярные тренинги в формате микро-обучения, сценарные симуляции фишинга с безнаказательным репортингом, поведенческие «подталкивания» в интерфейсе почты (превентивные подсказки и маркеры риска внешних отправителей).

Организационный контур. Разрабатываются ясные политики обработки данных, линия быстрого уведомления о подозрительных письмах, обратная связь сотруднику по результатам инцидентов и поощрения за бдительность. Психологическая безопасность репортинга снижает барьеры к сообщению об ошибках и подозрениях, предотвращая развитие инцидента [2].

Технический контур. DLP рассматривается как «страховка второго эшелона» после человека и процесса. Применяются контентно-контекстные правила для e-mail, веб-каналов и облачных хранилищ, включая семантический анализ текста и вложений для выявления конфиденциальной информации. Внедряются шаблоны документов и контент-метки для блокирования или задержки отправки с уведомлением пользователя и службы ИБ. Правила учитывают контекст (рабочее время, тип файла, новизна получателя) и реализуют мягкие, объяснимые блокировки для сохранения доверия [3].

Выводы

Повышение устойчивости к фишингу достигается, прежде всего, через работу с мотивацией, нормами и поведением сотрудников, а правильно настроенная DLP закрывает оставшееся окно эксфильтрации, превращая единичную ошибку в управляемый инцидент без утечки. Социотехническое объединение осознанности, психологически безопасного репортинга и семантических DLP-политик создаёт «двойной барьер», который уменьшает вероятность инцидента и масштаб его последствий без избыточного давления на пользователей.

Список литературы

1. Gomes V., Reis J. Social Engineering and the Dangers of Phishing // 2020 15th Iberian Conference on Information Systems and Technologies (CISTI). – 2020. – DOI: 10.23919/CISTI49556.2020.9140445.

2. Ruohonen J., Saddiqa M. What Do We Know About the Psychology of Insider Threats? // arXiv preprint arXiv:2407.05943. – 2024. – DOI: 10.48550/arXiv.2407.05943.

3. Alhindi H., Traore I., Woungang I. Preventing Data Leak through Semantic Analysis // Internet of Things. – 2019. – Vol. 14. – P. 100073. – DOI: 10.1016/j.iot.2019.100073.

УДК 004.056

А.Д. АБХАЗИ[1], В.И. КОРОЛЁВ[2]

[1]Московский государственный лингвистический университет
[2]ФИЦ «Информатика и управление» РАН, Москва

О МЕТОДЕ ПРИМЕНЕНИЯ МНОГОФАКТОРНОГО ПОДХОДА В ПОСТРОЕНИИ ИНТЕГРИРОВАННЫХ СИСТЕМ ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Современные ИТ-инфраструктуры представляют собой сложные распределённые системы, включающие множество взаимосвязанных компонентов. Эффективная защита таких систем требует комплексного подхода, способного учитывать взаимное влияние технических и организационных факторов. В этой связи особое значение приобретают методы, позволяющие описывать совокупную эффективность механизмов защиты не изолированно, а как результат взаимодействия множества факторов.

Многофакторный подход, применяемый в инженерии, экономике и системном анализе, доказал свою эффективность при решении задач, где конечный результат определяется влиянием множества переменных. Так, в эконометрике он используется для анализа зависимости макроэкономических показателей от внешних факторов; в теории надёжности – для оценки устойчивости технических систем; в управлении качеством – для анализа влияния технологических параметров на итоговый продукт. Эти примеры демонстрируют универсальность метода, который может быть адаптирован и к задачам обеспечения информационной безопасности.

В контексте системы обеспечения информационной безопасности (СОИБ) многофакторный подход позволяет рассматривать систему защиты как совокупность взаимодействующих факторов, каждый из которых характеризует реализацию определённой функции безопасности. К таким функциям относятся обнаружение атак, фильтрация трафика, корреляция событий, предотвращение утечек, криптографическая защита, другие функции. Итоговая эффективность системы определяется не только наличием отдельных функций, но и степенью их согласованности.

Использование многофакторной модели позволяет проводить структурированный анализ вклада каждого компонента в общую защищённость, выявлять дублирующие функции, оценивать значимость различных средств и оптимизировать их сочетание. Такой подход

обеспечивает переход от фрагментарного внедрения средств защиты к системному синтезу, в рамках которого архитектура СОИБ формируется на основе приоритетов, определяемых моделью угроз и политикой безопасности.

Кроме того, многофакторная функция может применяться для построения моделей корреляции индикаторов атак и инцидентов. В этом случае факторами выступают признаки сетевого трафика, показатели энтропии и результаты классификации, отражающие состояние защищённости инфраструктуры. Построение моделей корреляции позволяет унифицировать обработку данных из различных источников и сегментов – от корпоративных ИТ-сетей до промышленных систем управления [1].

Рассмотрение процессов обеспечения безопасности через многофакторную модель открывает возможности для формализации взаимосвязей между техническими, организационными и аналитическими компонентами защиты. Это создаёт основу для объективной оценки эффективности решений различных вендоров и выбора оптимальных комбинаций средств защиты с точки зрения покрытия угроз, совместимости и стоимости [2].

Многофакторный метод обладает рядом преимуществ: он модульный, оценочный и гибкий. Применение данного подхода в информационной безопасности способствует переходу от субъективных экспертных оценок к формализованным моделям принятия решений.

Заключение: многофакторный подход, успешно зарекомендовавший себя в инженерных и экономических областях, может быть эффективно применён в информационной безопасности. Его использование обеспечивает системное представление архитектуры защиты, количественную оценку взаимодействия средств и формализует процесс выбора оптимальных решений.

Список литературы

1. Королёв В.И., Абхази А.Д. Анализ интеграционных возможностей средств защиты от сетевых атак систем интенсивного использования данных // Системы высокой доступности. 2025. Т. 21. No 3. С. 5−20. DOI: https://doi.org/ 10.18127/j20729472-202503-01

2. Королев В.И. Эволюция и современные тенденции защиты автоматизированных информационных систем с сетевой ИТ-инфраструктурой / В.И. Королев, А.Д. Абхази // Вестник РГГУ. Серия: Информатика. Информационная безопасность. Математика. – 2024. – № 4. – С. 58–80. – DOI 10.28995/2686-679X-2024-4-58-80. – EDN SVMVGJ

УДК 316.774:004.9

А.С. КОЛОМИНА

Московский университет МВД России им. В.Я. Кикотя

РАСПРОСТРАНЕНИЕ ДИСКРЕДИТИРУЮЩЕЙ ИНФОРМАЦИИ: ОБЗОР МОДЕЛЕЙ

Систематизированы подходы к построению моделей распространения дискредитирующей информации. В качестве решения задач предлагается системно-динамическое моделирование на основе имитационной платформы AnyLogic.

Введение

Социальные медиа выступают средой для реализации дискредитирующих кампаний применительно к организациям и конкретным индивидам. В статье систематизированы современные подходы к математическому моделированию распространения дискредитирующей информации.

Обзор подходов к моделированию

1. *Моделирование инсайдерских угроз*. Для анализа мотивационных факторов инсайдеров применяется системно-динамическое моделирование, отражающее характеристики ожидания, чувства достижения и недовольства [1]. В таких моделях важен комплексный учет триады переменных «человеческий фактор – технологическая среда – организационные механизмы». В то же время существующие модели слабо учитывают связь внутренних угроз с внешними скоординированными кампаниями осуществления дискредитации.

2. *Агент-ориентированное моделирование*. Подход формализует два типа социального влияния [2]: нормативное влияние – стремление к конформности под давлением группы; информационное влияние – основанное на доверии к источнику.

Динамика мнения агентов описывается дифференциальными уравнениями, учитывающими социальное влияние соседей во времени. Однако модель демонстрирует ограниченные возможности для исследования полномасштабных дискредитирующих кампаний, отличающихся нелинейными связями.

3. *Стохастические модели распространения информации*. Классическая модель Дейли-Кендала делит популяцию на три группы: невовлеченные, распространители дискредитирующей информации и

прекратившие ее распространять. Современные разработки объединяют модели распространения слухов и модели динамики эпидемий [3]. Они полезны для прогнозирования масштаба распространения дискредитирующей информации, но не объясняют социально-психологические механизмы вовлечения в процесс дискредитации.

4. *Системно-динамическое моделирование.* Российскими исследователями [4–6] разработаны комплексные модели распространения манипулятивного контента. Базовая модель описывается системой дифференциальных уравнений, связывающих ключевые переменные: количество лиц, способных принять идею дискредитации, и количество принявших такую идею.

Заключение

Аналитический обзор выявил разнообразие подходов к моделированию процессов дискредитации: агент-ориентированные модели раскрывают механизмы социального влияния; стохастические модели прогнозируют макродинамику процессов; системно-динамические модели учитывают нелинейные обратные связи; модели инсайдерских угроз объединяют машинное обучение с системной динамикой. Наиболее перспективной основой для построения комплексной модели распространения дискредитирующей информации представляется системно-динамический подход в средах типа AnyLogic.

Список литературы

1. Yang S.- C., Wang Y.- L. System Dynamics Based Insider Threats Modeling // Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications. 2015. Vol. 6, No 3. – P. 61–81. DOI: 10.22667/JOWUA.2015.09.31.061.

2. Alf H. Identifikation von Influentials in virtuellen sozialen Netzwerken. Berlin: Technische Universität Berlin, 2018. – 210 p.

3. Nekovee M., Moreno Y., Bianconi G., Marsili M. Theory of Rumour Spreading in Complex Social Networks // Physics: Statistical Mechanics and its Applications. 2007. Vol. 374, № 1. – P. 457–470. DOI: 10.1016/j.physa.2006.07.017.

4. Минаев В.А., Вайц Е.В., Грачева Ю.В. Имитационные эксперименты с моделью информационно-психологических воздействий на массовое сознание // Безопасность информационных технологий. 2017. Т. 24. № 2. – С. 61–71.

5. Минаев В.А., Сычев М.П., Вайц Е.В., Бондарь К.М. Системно-динамическое моделирование сетевых информационных операций // Инженерные технологии и системы. 2019. Т. 29, № 1. – С. 20–39. DOI: https://doi.org/10.15507/2658-4123.029.201901.020-039.

6. Минаев В.А., Корячко А.В., Коломина А.С. Системно-динамическая модель распространения дискредитирующей информации в социальных медиа // Вестник РГРТУ. 2025. № 93. С. 100–109.

УДК 004.838

К.М. БОНДАРЬ[1], В.С. ДУНИН[1], В.А. МИНАЕВ[2]
[1]Дальневосточный юридический институт МВД России им. И.Ф. Шилова, Хабаровск
[2]Московский университет МВД России им. В.Я. Кикотя

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ ПРИ РЕШЕНИИ ЗАДАЧ ПРОТИВОДЕЙСТВИЯ КИБЕРМОШЕННИЧЕСТВУ

Рассматриваются возможности решения оперативно-служебных задач органов внутренних дел (ОВД) на основе применения эффективных методов, тактических и технологических приемов искусственного интеллекта (ИИ), направленных на обеспечение профилактики, раскрытия, расследования кибермошенничеств, уровень которых в последние годы стал критическим.

Введение

Современное состояние кибермошенничества свидетельствует о массовом использовании киберпреступниками технологий машинного обучения, генеративного ИИ, создающих предельно качественное психологического онлайн-воздействие на население страны. Используя ИИ, мошенники проводят новые формы кибератак, заражения устройств пользователей вредоносным программным обеспечением, телефонного мошенничества, применения нейронных сетей для генерации в масштабе реального времени фейковых аудио- и видеоматериалов. Это нередко нарушает эмоциональный баланс жертв, притупляет их бдительность, приводя к масштабным финансовым и материальным потерям граждан и организаций. Критический уровень проблемы отмечен на мартовской коллегии МВД России 2025 г. Президентом РФ В.В. Путиным и Министром внутренних дел РФ В.А. Колокольцевым, указавшим на «недопустимый размер ущерба» для населения от данных преступных деяний. Сам факт их стремительного роста приравнен к угрозе национальной безопасности страны [1].

Пути применения технологий ИИ

Основными направлениями противодействия указанной угрозе является обеспечение оперативной, процессуальной, экспертной и профилактической деятельности ОВД [2] технологиями ИИ. Важно также государственное реагирование путем принятия Федерального закона по созданию государственной информационной системы противодействия правонарушениям, совершаемым с использованием информационно-коммуникационных технологий (ИКТ) [3].

Для решения задач противодействия кибермошенничеству ключевыми направлениями выступают:

1. *Организационно-методическое обеспечение.* Для «проигрывания» типовых ситуаций совершения преступлений, относимых к кибернетическим, весьма перспективно в методическом плане использование имитационного моделирования. Анализ модельной статистики в ее сопоставлении с реальными данными о киберпреступлениях позволяет обосновывать линии поведения пользователей и формировать эффективные методические рекомендации по противодействию отмеченным видам преступных деяний.

2. *Системы эмоционально-психологической поддержки* жертв кибернападения. Задача инструментария ИИ при этом – не допустить длительного нахождения потенциальной жертвы в условиях возникшего эмоционально-психологического стресса. Смысл состоит в том, чтобы атакуемый не успел «запустить» опасные для него психологические механизмы, которые работают на восприятие легенд «оригинальных» обстоятельств, излагаемых кибермошенниками. На это нацелены антифрод-системы ряда финансовых организаций (например, банки СБЕР, ВТБ, Тиньков Мобайл) с интерфейсом опций защитного режима.

3. *Противодействие новым угрозам*. Связано с интеллектуальным поиском уязвимостей программных продуктов правоохранительными структурами, ответственными за организацию противодействия киберпреступности. Результаты такой аналитической оценки должны лечь в основу формирования рекомендаций по профилактике новых киберугроз как для организаций, так и индивидуальных пользователей, подвергающихся кибермошенничеству и поэтому являющихся для правоохранительных органов категориями, требующими приоритетной киберзащиты.

Список литературы

1. Путин поручил улучшить использование ИИ для расследования преступлений // URL: https://www.rbc.ru/rbcfreenews/ 65a7a3359a794761a6913056 (дата обращения: 09.09.2025).

2. Искусственный интеллект: противодействие киберпреступности: Монография / В.А. Минаев [и др.], под ред. В.А. Минаева, К М. Бондаря. Хабаровск: ДВЮИ МВД России имени И.Ф. Шилова, 2025. – 256 с.

3. О создании государственной информационной системы противодействия правонарушениям, совершаемым с использованием информационных и коммуникационных технологий, и о внесении изменений в отдельные законодательные акты Российской Федерации: Федеральный закон от 01.04.2025 № 41-ФЗ. Доступ из справочно-правовой системы «КонсультантПлюс».

УДК 004.056.5

В.А. МИНАЕВ[1], Д.И. СУРЖИК[1], О.Р. КУЗИЧКИН[2]

[1]Московский университет МВД России им. В.Я. Кикотя,
[2]Московский государственный университет им. Н.Э. Баумана

МОДЕЛЬ НАРУШИТЕЛЯ В СИСТЕМЕ ФАЗОМЕТРИЧЕСКОГО МОНИТОРИНГА ОБЪЕКТА КИИ

Рассматривается математическая модель нарушителя объекта критической информационной инфраструктуры (КИИ) в системе фазометрического мониторинга, связанного с анализом фазовых изображений искусственного поля в подповерхностной части геологической среды. Сигнал нарушителя представляется в виде сложения детерминированной кусочно-линейной функции, описывающей подповерхностную активность, и стохастической компоненты с нормальным распределением. Модель позволяет осуществлять раннее обнаружение, классификацию и прогнозирование подповерхностной активности нарушителя.

Введение

Обеспечение физической безопасности объектов КИИ требует создания многоуровневых систем защиты, способных обнаруживать различные сценарии нарушений, включая скрытое проникновение через подповерхностное пространство. Существующие системы периметровой безопасности часто не способны детектировать подобные угрозы на ранних стадиях их развития.

Описание модели

В работе предлагается подход к обнаружению нарушителя, основанный на регистрации и анализе фазовых параметров искусственного многофазового поля, формируемого в подповерхностной части геологической среды [1]. Изменения фазовой структуры поля, вызванные активностью нарушителя, описываются выражением:

$$\Delta\varphi(x,y,z,t) = \varphi^*(x,y,z,t) - \varphi_0(x,y,z,t),$$

где $\Delta\varphi(x,y,z,t)$ – аномальное отклонение, содержащее информацию о нарушителе, $\varphi^*(x,y,z,t)$ — текущее фазовое изображение, $\varphi_0(x,y,z,t)$ – фазовое изображение «нормального» состояния.

На основе анализа экспериментальных данных мониторинга предложена аддитивная модель фазового сигнала нарушителя:

$$s(t) = s_{НКЛФ}(t) + n(t),$$

где $s_{НКЛФ}(t)$ – детерминированная составляющая на основе переключающей непрерывной кусочно-линейной функции (НКЛФ) [2], моделирующая характерные фазы развития подповерхностной активности, $n(t)$ – стохастическая составляющая (как правило, нормально распределенная), моделирующая фоновые помехи и шумы.

Оптимальные параметры модели определяются путем минимизации итерационными методами (например, методом сопряженных градиентов) среднеквадратического отклонения (СКО) между экспериментальными данными $s_{эксп}(t)$ и модельным представлением $s_{мод}(t,\theta)$ по вектору θ:

$$J(\theta)=\sqrt{\frac{1}{T}\int_0^T\left[s_{эксп}(t)-s_{мод}(t,\theta)\right]^2 dt}\to\min_{\theta}.$$

Разработанная модель может быть интегрирована в информационно-аналитическую систему безопасности объекта КИИ, где осуществляется:

– обнаружение нарушителя по правилу $\int_0^T\left[s(t)-s_{НКЛФ}(t)\right]^2 dt<\gamma$ и принятие гипотезы $H_1: s(t)=s_{НКЛФ}(t)+n(t)$ или $H_0: s(t)=n(t)$;

– классификация типа активности (подкоп, деформация грунта, естественная геодинамика) на основе идентифицированных параметров θ;

– прогнозирование развития события путем экстраполяции модельной функции на интервал прогноза Δt: $s(t+\Delta t)=s_{НКЛФ}(t+\Delta t)+n(t+\Delta t)$.

Заключение

Экспериментальная апробация модели показала эффективность обнаружения проникновений на ранних стадиях, возможность классификации типа подповерхностной активности и достоверность краткосрочного прогнозирования ее развития.

Список литературы

1. Суржик Д.И. Регистрация геодинамических процессов фазометрическим методом в системах локального геоэкологического мониторинга / Д.И. Суржик, О.Р. Кузичкин, Г.С. Васильев и др. // Фундаментальные и прикладные проблемы техники и технологии. – 2021. – № 5(349). – С. 109–118.

2. Курилов И.А. Методы анализа радиоустройств на основе функциональной аппроксимации / И.А. Курилов, В.В. Ромашов, Е.А. Жиганова, Д.Н. Романов, Г.С. Васильев, С.М. Харчук, Д.И. Суржик // Радиотехнические и телекоммуникационные системы. – 2014. – № 1(13). – С. 35–49.

УДК 519.635

В.А. МИНАЕВ[1], И.С. МАНЖУЛА[2]

[1]Московский университет МВД России им. В.Я. Кикотя
[2]Дальневосточный юридический институт МВД России им. И.Ф. Шилова, Хабаровск

МОДЕЛИРОВАНИЕ ТЕПЛОМАССОПЕРЕНОСА В СОЛНЕЧНЫХ КОЛЛЕКТОРАХ ПРЯМОГО ПОГЛОЩЕНИЯ ДЛЯ ОБЕСПЕЧЕНИЯ УСТОЙЧИВОСТИ КРИТИЧЕСКОЙ ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЫ

В работе сформулирована математическая модель тепломассопереноса в солнечном коллекторе прямого поглощения (СКПП) – ключевом элементе автономного теплоснабжения объектов КИИ. Модель представлена системой дифференциальных уравнений в частных производных для температуры, скорости и давления с учётом начальных и граничных условий, отражающих физику течения, теплообмена и поглощения солнечного излучения.

Нанотехнологии играют всё более значимую роль в обеспечении устойчивости КИИ. Особенно перспективны наножидкости, которые в СКПП одновременно поглощают солнечное излучение и транспортируют тепло [1].

Для КИИ надёжное теплоснабжение в полевых или автономных условиях критично: оно обеспечивает стабильную работу электронного оборудования, систем жизнеобеспечения и защиты данных. СКПП позволяют автономно поддерживать необходимый температурный режим в мобильных пунктах управления, узлах связи или серверных модулях в удалённых или труднодоступных районах, где отсутствует централизованное энергоснабжение. Таким образом, применение наножидкостей в СКПП – это не только технологическое усовершенствование, но и вклад в укрепление информационного суверенитета Российской Федерации.

Рассмотрим область $D = \{0 < x < L, \quad 0 < y < H\}$, где L и H – длина и высота СКПП, через которую течет наножидкость и на которую сверху в вертикальном направлении падает солнечный свет.

Процессы тепломассопереноса в СКПП моделируются системой дифференциальных уравнений в частных производных, описывающей распределение температуры T, вектора скорости $\mathbf{V}$ и давления p рабочей жидкости, дополненной соответствующими начальными и граничными

условиями, отражающими физические особенности течения, теплообмена и поглощения солнечного излучения [2]:

$$\nabla\left(k\nabla T\right)-c_v\mathbf{V}\nabla T-\nabla\mathbf{I}=0,\quad (x,y)\in D, \tag{1}$$

$$\nabla\left(\mu\nabla\mathbf{V}\right)-\rho\left(\mathbf{V}\nabla\right)\mathbf{V}-\nabla p=0, \tag{2}$$

$$\nabla\mathbf{V}=0, \tag{3}$$

$$\left(k\frac{\partial T}{\partial y}-\sigma_0\left(T-T_e\right)+\alpha I_\theta\right)\Big|_{y=0}=0,\left(k\frac{\partial T}{\partial y}+\sigma_H\left(T-T_e\right)\right)\Big|_{y=H}=0,\ x\in[0,L], \tag{4}$$

$$T\big|_{x=0}=T_{in},\quad y\in[0,H], \tag{5}$$

$$\mathbf{V}\big|_{x=0}=\mathbf{V}_0,\quad \mathbf{V}\big|_{x=L}=\frac{\partial\mathbf{V}}{\partial x},\quad y\in[0,H], \tag{6}$$

$$\mathbf{V}\big|_{y=0}=0,\quad \mathbf{V}\big|_{y=H}=0,\quad x\in[0,L], \tag{7}$$

$$\int_D p\left(x\right)dx=0, \tag{8}$$

где $\mathbf{I}=(0,-I)$, I – интенсивность светового потока солнечного излучения; T_e – температура окружающей среды; σ_0, σ_H – коэффициенты теплоотдачи на нижней и на верхней границах СКПП; α - коэффициент, принимающий значения 0 и 1; ρ – плотность жидкости, μ – кинематическая вязкость, $\theta=\text{const}$ – коэффициент отражения солнечного излучения от нижней границы, $\theta\in[0,1]$.

Список литературы

1. Tyagi, H. Predicted Efficiency of a Low-Temperature Nanofluid-Based Direct Absorption Solar Collector / H. Tyagi, P. Phelan, R. Prasher // Journal of Solar Energy Engineering. – 2009. – Vol. 131, iss. 4.

2. Вихтенко Э.М., Манжула И.С. Программный комплекс для проведения математического моделирования интенсивности солнечного излучения внутри наножидкостного солнечного коллектора прямого поглощения // Перспективы науки. 2024. № 5 (176). С. 105–111.

УДК 004.491.22

В.А. МИНАЕВ[1], А.О. ФАДДЕЕВ[1], Ю.В. БРОНЕНКОВА[2]

[1]*Московский университет МВД России им. В.Я. Кикотя*
[2]*Академия управления МВД России, Москва*

ПРОГНОЗИРОВАНИЕ ФИШИНГОВЫХ АТАК НА ОСНОВЕ РЕКУРРЕНТНОЙ НЕЙРОННОЙ МОДЕЛИ

В работе использована рекуррентная нейронная сеть для прогнозирования фишинговых атак. Точность краткосрочного прогноза в среднем составила 6%. Таким образом, прогноз, выполненный на основании *LSTM*-модели, обладающей механизмами постепенного забывания прошлой информации и замещения ее актуальной, значительно превышает точность моделей, базирующихся на аппарате обычной экстраполяции.

Введение

Рассмотрены фишинговые атаки, являющиеся доминирующими киберпреступлениями в Российской Федерации [1–3]. При оценках их изменения во времени использованы модели, применяемые для описания динамических рядов, у которых значение $Y(t_i)$ определяется значениями $Y(t)$ в предыдущие моменты времени $t_{i-1}, t_{i-2}, \ldots$, а именно:

$$Y(t_i) = f\left(Y(t_{i-1}), Y(t_{i-2}), \ldots\right) + \varepsilon(t_i), \quad (1)$$

где $e\ (t_i)$ – случайная составляющая.

Методика прогнозирования

Степень статистической связи между последовательностями $Y(t_1)$, $Y(t_2)$, …, $Y(t_n)$ и $Y(t_{1+l})$, $Y(t_{2+l})$, …, $Y(t_{n+l})$ определялась с помощью значений его автокорреляционной функции. Проверка значимости отдельного коэффициента автокорреляции производилась на основании критерия стандартной ошибки, а значимость всей совокупности коэффициентов автокорреляции динамического ряда – на основании Q-критерия Бокса-Пирса.

Для анализа структуры динамического ряда фишинговых атак использовалась аддитивная модель вида:

$$Y(t_i) = q(t_i) + \varepsilon(t_i), i = 1, 2, \ldots, n, \quad (2)$$

где детерминированная составляющая $q(t_i)$ включает в себя трендовую $g(t_i)$, сезонную $s(t_i)$ и периодическую $p(t_i)$ компоненты.

Задача прогнозирования представляет собой получение, на основе имеющихся последовательных данных о фишинговых атаках, информации о количестве атак подобного вида на заранее определённом временном

интервале упреждения. В качестве такого временного интервала выбран первый квартал 2025 г.

Для получения прогнозной информации использована рекуррентная нейронная сеть *LSTM* (*Long Short-Term Memory*).

Обратимся к результатам моделирования динамики фишинговых атак на основе использования рекуррентной нейронной сети *LSTM* (рис. 1).

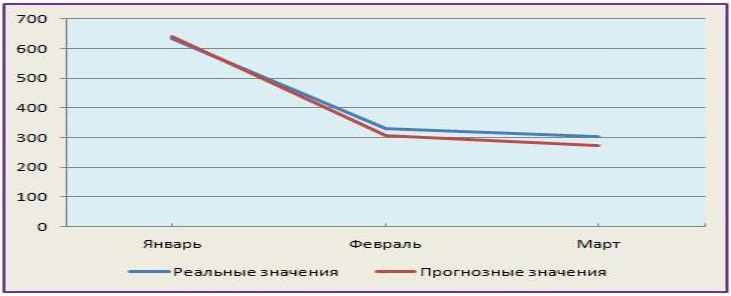

Рис. 1. Результаты моделирования фишинговых атак

Программное исполнение модели выполнено на языке *Python*. Верхней линией на рис. 1 обозначены фактические данные исходного динамического ряда фишинговых атак за первые три месяца 2025 г., нижней линией – значения количества фишинговых атак, полученные на основании прогнозных оценок по *LSTM*-модели.

Точность прогноза в среднем составила 6%, то есть прогнозные и реальные данные о количестве фишинговых атак в первый квартал 2025 г. весьма близки между собой.

Заключение

Развиваемый в статье подход, ориентированный на использование нейросетевых моделей типа *LSTM*, позволяет получить достаточно точный краткосрочный прогноз фишинговых атак в стране, позволяя перейти ко второму этапу – имитационному моделированию в сфере противодействия компьютерным атакам в Российской Федерации.

Список литературы

1. Лунеев В.В. Преступность XX века: мировые, региональные и российские тенденции: Монография. 2-е изд., переработанное и дополненное. М.: Норма, 2021. – 912 с.

2. Авсентьев А.О., Винокуров С.А., Минаев В.А. и др. Основы кибербезопасности: Учебник. М.: ГУРЛС МВД России, 2024. – 408 с.

3. Андреев А.А., Бондарь К.М., Минаев В.А. Терроризм и экстремизм: моделирование информационного противодействия: Монография. Хабаровск: РИО ДВЮИ МВД России, 2020. – 248 с.

УДК 004.93:004.056

А.С. ВАНИЧКИНА, Н.О. АНТОНЮК

Московский государственный лингвистический университет

МЕТОДЫ ОБНАРУЖЕНИЯ АУДИО-ДИПФЕЙКОВ НА ОСНОВЕ ГЛУБОКОГО ОБУЧЕНИЯ

Рассмотрена проблема обнаружения синтетических аудиозаписей, созданных с использованием технологий искусственного интеллекта. Обоснована актуальность данной проблемы, связанной с угрозами для финансового сектора, систем голосовой биометрии и распространением дезинформации. Проанализированы традиционные методы идентификации аудио-дипфейков на основе ручных признаков (MFCC, GTCC) и выявлены их недостатки. Предложено использование современных подходов на основе глубоких нейронных сетей с применением сверточных и рекуррентных архитектур, обеспечивающих автоматическое извлечение признаков и повышенную адаптивность к новым типам атак.

Введение

Современные технологии синтеза речи достигли такого уровня качества, что синтетический голос практически неотличим от подлинного. Злонамеренные аудио-дипфейки используются для организации мошеннических схем, обхода систем биометрической аутентификации, манипулирования общественным мнением и дестабилизации политической обстановки, нанося значительный экономический и репутационный ущерб [1]. В качестве базового метода противодействия часто предлагается использование акустических признаков для классификации. Однако простые методы легко обходятся современными алгоритмами генерации. Злоумышленники могут применять постобработку для сокрытия артефактов синтеза. Это требует применения более надежных методов обнаружения.

Текущий подход к решению проблемы

Наиболее распространенным методом обнаружения аудио-дипфейков является анализ ручных признаков с применением классических алгоритмов машинного обучения [2]. Данный метод позволяет извлекать мел-частотные кепстральные коэффициенты (MFCC), спектральные характеристики и другие акустические признаки для последующей классификации с помощью SVM или Random Forest. Например, SVM-классификатор способен достигать точности до 99% на известных типах атак. Однако данный подход имеет существенные ограничения. Ручное

проектирование признаков требует глубокой экспертизы, а производительность значительно снижается на зашумленных данных и новых, ранее не встречавшихся типах атак. Следовательно, использования только классических методов машинного обучения недостаточно для надежного обнаружения аудио-дипфейков в реальных условиях.

Предлагаемое решение

В качестве принципиально нового подхода к решению проблемы обнаружения аудио-дипфейков предлагается применение методов глубокого обучения с использованием сверточных и рекуррентных нейронных сетей. Данный подход предполагает автоматическое извлечение сложных иерархических представлений из спектрограмм аудиосигналов [3]. Архитектура системы включает модуль предобработки для формирования мел-спектрограмм, сверточные слои для извлечения пространственных признаков и рекуррентные блоки для моделирования временных зависимостей. Протокол обнаружения реализуется преобразованием аудио в частотно-временное представление, извлечение глубоких признаков нейронной сетью и классификацию с использованием специализированных функций потерь, таких как OC-Softmax, обеспечивающих разделение подлинной и синтетической речи в пространстве признаков.

Заключение

Таким образом, предложенная система обнаружения на основе глубокого обучения представляет собой принципиально новый подход к решению проблемы идентификации аудио-дипфейков. Использование сверточных и рекуррентных нейронных сетей позволяет перейти от ручного проектирования признаков к автоматическому извлечению сложных паттернов и обеспечивает надежную защиту от синтетических аудиозаписей с достижением EER менее 2% на стандартных наборах данных.

Список литературы

1. Almutairi Z., Elgibreen H. A review of modern audio deepfake detection methods: Challenges and future directions // Algorithms. – 2022. – Vol. 15. – N. 5. – Art. 155.

2. Chen T., Kumar A., Nagarsheth P., et al. Generalization of audio deepfake detection // Proceedings of the Odyssey 2020: The Speaker and Language Recognition Workshop. – 2020. – P. 132–137.

3. Jung J.W., Heo H.S., Tak H., et al. AASIST: Audio anti-spoofing using integrated spectro-temporal graph attention networks // Proceedings of the IEEE ICASSP. – 2022. – P. 6367–6371.

УДК 004.056

Д.П. ВОРОНЦОВА
Национальный исследовательский ядерный университет «МИФИ», Москва

ЗАЩИТА ПЕРСОНАЛА ОБЪЕКТОВ КИИ ОТ ДЕСТРУКТИВНЫХ ИПВ МУЛЬТИМЕДИЙНОГО КОНТЕНТА СОЦИАЛЬНЫХ СЕТЕЙ

Целью исследования является разработка подхода к защите персонала объектов критической информационной инфраструктуры от информационно-психологических воздействий мультимедийного контента. Предложено создание программного изделия, обеспечивающего анализ визуальных, вибрационных и акустико-речевых характеристик контента. Реализация подхода позволит повысить устойчивость персонала к деструктивному влиянию и минимизировать риски ошибок при работе с информацией.

Современное развитие критической информационной инфраструктуры (КИИ) сопровождается ростом числа атак, направленных не только на технические системы, но и на персонал, обеспечивающий их функционирование. За последние годы накоплен значительный массив исследований, посвящённых информационно-психологическим угрозам. Так, Lazer и соавторы показали, что ложные сообщения распространяются в цифровой среде значительно быстрее достоверных, формируя устойчивые искажения восприятия [1]. Wardle и Derakhshan разработали концепцию «информационного расстройства», согласно которой деструктивный контент влияет на эмоциональную сферу и снижает способность человека к критическому анализу [2]. Исследования Chesney и Citron раскрывают риски, связанные с использованием технологий deepfake для дискредитации и манипулирования общественным мнением [3].

Несмотря на активное развитие этого направления, большинство работ сосредоточено на массовом потребителе информации. При этом специфические условия работы персонала КИИ, высокая ответственность, информационная насыщенность и стрессовые факторы, практически не учитываются. Существующие методы защиты ориентированы на фильтрацию текстовых сообщений и не включают анализ синтетического мультимедийного контента.

Появление технологий генерации аудио- и видеоматериалов (deepfake, синтетическая речь, искусственные аватары) создаёт качественно новые угрозы. Воздействие осуществляется не только на уровне сознания, но и через подсознательные механизмы восприятия (мимику, тембр, ритм,

интонацию). Это делает персонал КИИ особенно уязвимым, поскольку даже кратковременное эмоциональное воздействие может повлиять на точность профессиональных решений. Поэтому возникает необходимость в системном решении, способном выявлять признаки психологического влияния и формировать адаптивные меры защиты.

В рамках исследования предлагается создание программного изделия, обеспечивающего комплексный анализ мультимедийного контента социальных сетей с целью оценки его потенциального деструктивного воздействия. Основу решения составляют три взаимодополняющих вектора анализа:

– визуальный, связанный с оценкой мимических реакций и микровибраций лица;

– вибрационный, включающий анализ динамических изменений изображения;

– акустико-речевой, направленный на исследование интонации, тембра и спектральных характеристик голоса для определения эмоционального состояния.

Объединение этих направлений в единую модель с применением методов машинного обучения и многомодального анализа данных позволит выявлять закономерности между типом мультимедийного воздействия и эмоциональной реакцией человека.

Таким образом, исследование направлено на переход от фрагментарных методов анализа информации к созданию целостной системы защиты персонала КИИ. Разрабатываемое программное изделие позволит не только своевременно выявлять потенциально опасный контент, но и формировать рекомендации по его нейтрализации и обучению сотрудников методам противодействия психологическому давлению. Реализация предложенного подхода обеспечит повышение устойчивости персонала, снижение вероятности ошибок, вызванных информационно-психологическим влиянием, и станет основой для дальнейшего развития технологий комплексной защиты КИИ.

Список литературы

1. Дэвид М. Дж. Лазер и др. Наука о фейковых новостях. Science 359, с. 1094–1096 (2018). DOI: 10.1126/science.aao2998.

2. Wardle C., Derakhshan H. Information Disorder: Toward an Interdisciplinary Framework. – Council of Europe, 2017. – [Электронный ресурс]. URL: https://rm.coe.int/information-disorder-report/168076277c.

3. Робберт Чесни, Даниэль Ситрон. (2019). Глубокие фейки: надвигающаяся угроза конфиденциальности. California Law Review. https://doi.org/10.15779/Z38RV0D15J.

УДК 004.056

М.А. ЛАПИНА, А.Р. БАГАУТДИНОВА
Северо-Кавказский федеральный университет, Ставрополь

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ДЕТЕКЦИИ ДЕЗИНФОРМАЦИИ В МЕДИАСРЕДЕ

Аспектом исследования является разработка веб-сервиса для автоматического обнаружения фейковых текстов с использованием корпуса ТАЙГА (Текстовый аналитический инструмент глубинного анализа). Описывается архитектура сервиса, применяемые модели классификации, а также методы обработки естественного языка (NLP) для повышения точности детекции.

Актуальность и проблематика

Современное информационное пространство столкнулось с угрозой распространения дипфейков – синтетического медиаконтента, созданного с помощью искусственного интеллекта. Дипфейки способны реалистично имитировать изображения, видео, аудио и тексты, что делает их мощным инструментом дезинформации, манипуляции общественным мнением, шантажа и мошенничества. Особую опасность представляет их распространение через веб-платформы, где миллионы пользователей могут стать жертвами недостоверного контента. Традиционные методы верификации зачастую не справляются с обнаружением дипфейков, что требует разработки новых подходов на основе машинного обучения и анализа цифровых артефактов [1].

Основной целью является разработать и проанализировать методы машинного обучения для обнаружения дипфейков и фейковых новостей, включая создание веб-сервиса для автоматической детекции недостоверного контента. Были изучены типы дипфейков и методов их создания, оценена эффективность алгоритмов машинного обучения, проведено экспериментальное тестирование моделей на релевантных наборах данных, разработан практический инструмент для анализа текстов и медиафайлов.

Методология и подходы

Исследование основано на применении современных методов машинного обучения, включая сверточные нейронные сети (CNN), рекуррентные нейронные сети (RNN), трансформеры и гибридные архитектуры. Для анализа текстовой информации использовались модели серии RoBERTa, дообученные для детекции AI-генерации [2].

Эксперименты проводились на наборе данных FaceForensics++, содержащем подлинные и поддельные видео, созданные с использованием различных методов генерации дипфейков. Для обработки естественного языка (NLP) применялись методы лингвистического анализа и верификации источников.

Результаты исследования

Разработан веб-сервис на основе фреймворка Streamlit, интегрированный с моделями roberta-base-openai-detector и roberta-large-openai-detector для обнаружения текстов, сгенерированных нейросетями. Сервис позволяет анализировать входные тексты, определять вероятность их принадлежности к AI-генерации и визуализировать результаты. Проведенные эксперименты показали, что модель RoBERTa-large демонстрирует высокую точность в задачах детекции, особенно при анализе длинных и связных текстов. Для видеоанализа наилучшие результаты показали гибридные архитектуры, сочетающие CNN и RNN, с точностью до 97% на данных FaceForensics++.

Заключение

Наиболее эффективными методами борьбы с дипфейками и фейковыми новостями являются подходы, основанные на глубоком обучении и комплексном анализе метаданных, цифровых артефактов и лингвистических паттернов [3]. Разработанный веб-сервис демонстрирует практическую применимость моделей RoBERTa для автоматической детекции AI-генерации текста. В будущем необходимо продолжать совершенствование алгоритмов, адаптируя их к новым методам генерации дипфейков, а также развивать образовательные и правовые механизмы противодействия цифровым угрозам.

Список литературы

1. Masood M. et al. Deepfakes generation and detection: State-of-the-art, open challenges, countermeasures, and way forward //Applied intelligence. – 2023. – Т. 53. – №. 4. – С. 3974–4026.

2. Hancock J. T., Bailenson J. N. The social impact of deepfakes //Cyberpsychology, behavior, and social networking. – 2021. – Т. 24. – №. 3. – С. 149–152.

3. Nguyen T. T. et al. Deep learning for deepfakes creation and detection: A survey //Computer Vision and Image Understanding. – 2022. – Т. 223. – С. 103525.

УДК 004.056

Ф.А. ОЩЕНКО
Научный руководитель – д.т.н., профессор С.В. ДВОРЯНКИН
Национальный исследовательский ядерный университет «МИФИ», Москва

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ ЦИФРОВЫХ ЗВУКОВЫХ ОТПЕЧАТКОВ ДЛЯ РАСПОЗНАВАНИЯ И РЕКОНСТРУКЦИИ РЕЧИ В УСЛОВИЯХ ШУМОВ

Введение

Повышение точности автоматического распознавания речи в условиях сильных акустических помех является актуальной задачей в контексте развития человеко-машинных интерфейсов, систем безопасности и управления в таких сложных средах, как «умный город», транспорт и промышленное производство. Одним из перспективных направлений решения этой проблемы выступает интеграция технологии цифровых аудиоотпечатков (аудиофинтерпринтинга) в традиционные конвейеры распознавания речи для создания гибридных шумоустойчивых систем. Данное исследование посвящено оценке возможности и разработке методов применения компактных звуковых сигнатур для реконструкции и идентификации речевых фрагментов в зашумленных условиях, что направлено на преодоление ключевого ограничения классических подходов – значительной деградации качества при работе в реальных акустических средах.

Основная часть

Повышение устойчивости систем распознавания речи в шумной среде достигается за счет интеграции технологии аудиофинтерпринтинга, которая позволяет идентифицировать речевые фрагменты по их компактным спектральным отпечаткам. Ключевыми элементами реализованной гибридной системы являются:

- Спектральные пики (ориентиры), извлекаемые из спектрограммы сигнала, как наиболее устойчивые к аддитивному шуму и искажениям признаки.
- Алгоритм формирования отпечатка, аналогичный Shazam, основанный на хешировании пар спектральных пиков с учетом их частот и временных задержек.

- Модуль сопоставления, который сравнивает хеши входного сигнала с базой эталонных отпечатков для поиска чистого речевого фрагмента.

Вместе с тем, традиционные алгоритмы аудиофинтерпринтинга сталкиваются с рядом проблем при обработке именно речевых сигналов: высокая вариативность произношения, короткая длина сопоставляемых фрагментов (менее 1 с) и вычислительная сложность в реальном времени.

Недостаточная точность сопоставления в сложных акустических условиях может привести к серьезным последствиям: ошибкам в распознавании команд, снижению надежности голосовых интерфейсов и деградации производительности систем безопасности и управления.

В связи с этим, для повышения эффективности технологии в задачах реконструкции речи предлагается использовать следующие модификации:

1. Адаптивное шумоподавление на этапе предобработки сигнала.
2. Оптимизацию параметров Гауссова сглаживания и пороговой фильтрации для управления плотностью спектральных пиков.
3. Комбинирование признаков, таких как спектральные пики и мел-частотные кепстральные коэффициенты (MFCC), для повышения дискриминативности отпечатка.
4. Использование механизма голосования по временным сдвигам при сопоставлении для надежной верификации совпадений.

Заключение

Интеграция технологии аудиофинтерпринтинга демонстрирует потенциал для повышения точности и устойчивости систем распознавания речи в условиях акустических помех.

Список литературы

1. Дворянкин С.В., Зенов А.Е., Устинов Р.А., Дворянкин Н.С. Кодирование изображений спектрограмм для обеспечения переменной скорости передачи аудиоданных с сохранением качества их звучания. Безопасность информационных технологий. – 2021.

2. Будко Павел Александрович, Будко Никита Павлович, Виноградено Алексей Михайлович. Способы повышения помехоустойчивости в автоматизированных системах контроля. Системы управления, связи и безопасности. 2020.

3. Ладыгин Павел Сергеевич. Подход к созданию цифровых отпечатков аудиофайлов с использованием анализа нотной последовательности. Символ науки. 2017.

УДК 004.056

В.Б. ПРАХОВ[1], С.В. ДВОРЯНКИН[2]

[1]Московский государственный лингвистический университет
[2] Национальный исследовательский ядерный университет «МИФИ», Москва

ЗАЩИТА КОНФИДЕНЦИАЛЬНЫХ ПЕРЕГОВОРОВ В НЕПОДГОТОВЛЕННЫХ МЕСТАХ ПРОВЕДЕНИЯ

Представлен мобильный комплекс для защиты конфиденциальных переговоров (до 8 участников) в неподготовленных помещениях и на открытых пространствах. Комплекс обеспечивает защиту от утечки речевой информации по прямому акустическому каналу за счет комбинации пассивных (звукопоглощающий шлем, маскирующий речевой аппарат) и активных (речеподобная помеха, цифровое шифрование) методов. Использование беспроводных технологий связи (Bluetooth, Wi-Fi) обеспечивает защищённость мобильность и быстрое развертывание системы.

Довольно часто на практике встает вопрос о средствах и способах ЗРИ в местах, изначально не оборудованных спец. техническими средствами акустозагражденияРИ от утечки по прямому акустическому каналу, в то время как необходимость обсуждения значимых важных тем с коллегами явно присутствует. Это может быть и в салоне автомобиля, самолете, на аэродроме, в лесу, в ресторане и других открытых и закрытых площадках.

Преимуществом рассматриваемого мобильного комплекса является возможность ведения переговоров, защищенных от перехвата в зоне ограниченного пространства вокруг головы участника. Это достигается:

– сокрытием и маскировкой шлемом «речеобразующего аппарата» субъекта конфиденциальных переговоров для исключения возможности «чтения по губам» на расстоянии со стороны злоумышленника (ЗЛ);

– использованием мягких звукопоглощающих материалов с коэффициентом звукопоглощения (0,7–0,85), что в значительной степени «заглушает» распространение РС в окружающее пространство;

– применением вне головной зоны генератора речеподобной помехи (РПП) из-за её эффективности [1], доказанной рядом исследований [2];

– преобразованием защищаемого сигнала в цифровую формулу с последующим наложением криптоалгоритмов;

– применением встроенного микрофона с криптокодером внутри для обеспечения конфиденциальности переговоров;

– использованием специальных «наушников» для приема речевого сигнала переговорщиков и исключения влияния речеподобной помехи;

– использованием аналога toolkita Raspberry Pi для реализации дополнительных функций активной акустозащиты;

– использованием технологий «Bluetooth», «Wi-fi» или «Фемтосота» в качестве канала сетевой связи между другим мобильными комплексами участников конфиденциальных переговоров.

Пример работы комплекса для 2-х человек представлен на рис.1.

Рис. 1. Схема работы мобильного комплекса обеспечения конфиденциальности переговоров для 2-х участников

Мобильный комплекс рассчитан на шесть-восемь участников. Отмечается «быстроразворачиваемость» мобильного комплекса, совместимость работы со звуковыми отражающими экранами, удобная реконфигурация пассивной и активной составляющих комплекса под конкретные угрозы со стороны ЗЛ.

Благодаря применению мобильного комплекса снимается дискомфорт для участников конфиденциальных переговоров от использования РПП, повышается защищенность РИ от перехвата и взлома, нет необходимости в применении дорогостоящей спец. аппаратуры, проверки и аренды изначально оборудованных ею помещений конфиденциальных переговоров.

В качестве направления дальнейшего исследования можно рассмотреть оценку эффективности защиты речевой информации от утечки по акустовибрационному каналу.

Список литературы

1. Дворянкин С.В. Дворянкин Н.С., Устинов Р.А. Речеподобная помеха, стойкая к шумоочистке, как результат скремблирования защищаемой речи – Вопросы кибербезопасности 2022, № 5 (51). С. 14–27.

2. Прахов В.Б. Исследование влияния акустических помех на защищенность речевой информации – В сборнике: Кибернетика и информационная безопасность «КИБ-2024». Сборник научных трудов Второй Всероссийской научно-технической конференции. Москва, 2024. С. 150–151.

УДК 004.056

А.М. АЛЮШИН[1, 2], С.В. ДВОРЯНКИН[1, 2]

[1]*Национальный исследовательский ядерный университет «МИФИ», Москва*
[2]*Московский государственный лингвистический университет*

ОЦЕНКА ВЛИЯНИЯ ПАРАМЕТРОВ БИНАРНЫХ СОНОГРАММ НА РАЗБОРЧИВОСТЬ РЕЧИ

В работе представлены результаты экспериментального исследования влияния различных параметров и алгоритмов построения бинарных спектрограмм на разборчивость синтезированной речи. В качестве объективной метрики оценки использовался расширенный кратковременный объективный индекс разборчивости (ESTOI).

Задача маркирования речевой информации через визуальные описания представляет значительный интерес для области информационной безопасности [1]. Особую актуальность приобретают методы, основанные на бинарных спектрограммах, где сжатие данных до двух пикселей позволяет эффективно внедрять маркеры в графические представления речевых сигналов. Однако процесс бинаризации неизбежно приводит к потере критически важных для восприятия речевых компонентов. Целью исследования является определение оптимальных параметров бинаризации, обеспечивающих сохранение разборчивости речи при использовании в системах маркирования аудиоданных.

Эксперименты проводились на основе речевого корпуса RUSLAN. Исходные спектрограммы обрабатывались с применением различных алгоритмов бинаризации, которые можно разделить на две группы:

1. Пороговые методы, основанные на частотной ширине гармоники. Исследовался параметр максимальной ширины гармоники (в пикселях при фиксированной высоте спектрограммы в 512 точек). Тестировались значения от 1 до 7 пикселей.

2. Формулы преобразования яркости. Апробировались различные алгебраические и логарифмические преобразования исходных значений яркости пикселей.

Качество каждого варианта оценивалось с помощью метрики ESTOI, которая демонстрирует высокую корреляцию с результатами субъективного тестирования разборчивости [2–4]. Значения ESTOI интерпретировались в соответствии со стандартной шкалой разборчивости: 45–59% (низкая),

60–74% (средняя), 75–84% (высокая). Результаты для формулы равномерного разбиения приведены в табл. 1.

Таблица 1. Результаты экспериментов

Максимальная толщина гармоники в отн. ед.	**Результат ESTOI**	**Описание**
7	0.80	Высокая разборчивость, хорошо понимаются все слова
6	0.79	
5	0.79	
4	0.78	
3	0.74	
2	0.61	Основной смысл понятен, некоторые слова требуют усилий
1	0.46	Понимается только общий смысл фраз

Проведенное исследование демонстрирует, что для практической реализации алгоритмов синтеза речи по бинарным спектрограммам рекомендуется использовать ширину гармонических компонент на уровне 4 пикселей, что обеспечивает высокую разборчивость при оптимальных вычислительных затратах, так как, несмотря на то, что наилучший результат (ESTOI = 0.80) был достигнут при ширине в 7 пикселей, уменьшение ширины с 7 до 4 пикселей не приводит к значительной деградации разборчивости. Полученные результаты могут применяться при аудиомаркирования документированной информации, где важна помехоустойчивость и высокое качество распознавание.

Список литературы

1. Минаев В.А., Дворянкин С.В., Алюшин А.М. Методы биомаркирования защищаемых объектов. Информация и безопасность. 2023, т. 26, № 3, с. 321–328. ISSN 1682-7813. DOI: 10.36622/VSTU.2023.26.3.016. – EDN: TGJKIC.

2. Jensen J., Taal C.H. An Algorithm for Predicting the Intelligibility of Speech Masked by Modulated Noise Maskers. IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2016, vol. 24, no. 11, pp. 2009–2022. DOI: 10.1109/TASLP.2016.2585878.

3. Taal C.H., Hendriks R.C., Heusdens R., Jensen J. An Algorithm for Intelligibility Prediction of Time-Frequency Weighted Noisy Speech. IEEE Transactions on Audio, Speech, and Language Processing. 2011, vol. 19, no. 7, pp. 2125–2136. DOI: 10.1109/TASL.2011.2114881.

4. Taal C.H., Hendriks R.C., Heusdens R., Jensen J. A Short-Time Objective Intelligibility Measure for Time-Frequency Weighted Noisy Speech. In: 2010 IEEE International Conference on Acoustics, Speech and Signal Processing. 2010, pp. 4214–4217. DOI: 10.1109/ICASSP.2010.5495701.

УДК 004.056

Г.П. ГАВДАН

Национальный исследовательский ядерный университет «МИФИ», Москва

ПРОБЛЕМЫ И НАПРАВЛЕНИЯ РАЗВИТИЯ СИСТЕМ ЗАЩИТЫ РЕЧЕВОЙ ИНФОРМАЦИИ

Стремительное развитие речевых технологий на основе искусственного интеллекта (ИИ), включая распознавание, синтез и анализ речи, создает новые киберугрозы и вызовы в области защиты информации (ЗИ). В работе систематизированы ключевые проблемы защиты речевой информации и перспективные направления развития защитных механизмов.

Введение

Широкое внедрение искусственного интеллекта в речевые технологии кардинально трансформирует человеко-машинное взаимодействие, стремясь сделать его не только точным, но и контекстуально осмысленным. Однако эта эволюция создает новые векторы для кибератак, предоставляя злоумышленникам мощные инструменты на основе тех же технологий ИИ. В связи с этим, задача обеспечения безопасности речевых данных и систем их обработки приобретает критическую актуальность, требуя опережающего развития методов и средств защиты информации.

Ключевые проблемы защиты речевой информации

Анализ современных исследований [1–3] позволяет выделить следующие системные проблемы:

1. *Уязвимости моделей машинного обучения*: речевые системы на основе ИИ уязвимы к атакам.
2. *Риски конфиденциальности и этики:* сбор и обработка биометрических голосовых данных создают риски слежения.
3. *Зависимость от импортных технологий:* использование зарубежных платформ и фреймворков создает риски к защите данных.
4. *Высокие требования к вычислительным ресурсам:* реализация сложных алгоритмов защиты и шифрования речи в реальном времени.

Перспективные направления развития защитных механизмов

Для противодействия указанным угрозам развиваются следующие технологические направления, сущность и примеры которых систематизированы в табл. 1.

Таблица 1. Направления развития систем ЗИ в речевых технологиях

№	Направления развития	Сущность и примеры защитных механизмов
1	Упреждающий анализ угроз	Глубокий семантический анализ для выявления скрытых угроз; прогнозирование моделей поведения нарушителей; обнаружение аномалий в речевых потоках.
2	Активная защита на основе Generative AI	Генерация защищенных речевых сигналов, устойчивых к искажению; создание систем для проверки подлинности голоса (AI-верификаторы); разработка методов цифрового водяного знака для синтезированной речи.
3	Мультимодальная аутентификация	Комбинирование голосовой биометрии с анализом лицевых образов и поведенческих паттернов для повышения надежности идентификации.
4	Конфиденциальность данных	Применение методов деидентификации голоса; использование гомоморфного шифрования для обработки данных без их расшифровки; внедрение блокчейн-технологий для обеспечения неизменности голосовых логов.
5	Обеспечение доступности и целостности	Разработка систем речевого оповещения, устойчивых к шумовым и помеховым воздействиям; создание инструментов для лиц с нарушениями речи.

Проведенный анализ демонстрирует, что развитие речевых технологий неразрывно связано с эскалацией киберугроз.

Заключение

Дальнейшие исследования необходимо сконцентрировать на создании отечественных защищенных речевых платформ, способных функционировать в условиях импортозамещения. Для обеспечения безопасности необходима комплексная стратегия, включающая:

1. разработку адаптивных алгоритмов защиты;
2. внедрение архитектур безопасности «по умолчанию»;
3. создание нормативно-правовой базы (НПБ), регламентирующей сбор, обработку и использование биометрических голосовых данных.

Список литературы

1. Перспективные направления применения технологий искусственного интеллекта при защите информации / Р.В. Мещеряков, С.Ю. Мельников, В.А. Пересыпкин, А.А. Хорев // Вопросы кибербезопасности. – 2024. – № 4(62). – С. 2–12. – DOI 10.21681/2311-3456-2024-4-02-12. – EDN GJWQWP (дата обращения: 25.09.2025).

2. Дураковский, А.П. Эволюция и направления развития технологий маскирования конфиденциальных речевых сообщений / А.П. Дураковский, С.В. Дворянкин, Н.С. Дворянкин // Вопросы кибербезопасности. – 2024. – № 5(63). – С. 58–66. – DOI 10.21681/2311-3456-2024-5-58-66. – EDN AOBKXM (дата обращения: 29.09.2025).

3. Дворянкин, С.В. Лабораторный комплекс «Образный анализ и защита речевой информации в интеллектуальных системах» / С.В. Дворянкин // Приборы и системы. Управление, контроль, диагностика. – 2023. – № 6. – С. 66–72. – DOI 10.25791/pribor.6.2023.1419. – EDN IMXOAK (дата обращения: 29.09.2025).

УДК 004.056

С.В. ДВОРЯНКИН

Финансовый университет при Правительстве Российской Федерации, Москва
Московский государственный лингвистический университет

ПРОБЛЕМНЫЕ ВОПРОСЫ И РЕШЕНИЯ В ОБЕСПЕЧЕНИИ БЕЗОПАСНОСТИ РЕЧЕВОЙ ИНФОРМАЦИИ

Анализируется современное состояние и перспективы развития речевых технологий в контексте систем безопасности, управления и связи. Поднимаются актуальные вопросы защиты речевой информации, в т.ч. необходимость применения ИИ и моделирования речеподобных сигналов для тестирования систем. Особое внимание уделяется роли искусственного интеллекта в создании перспективных и анализе существующих систем защиты речевой информации.

В настоящее время речевые технологии претерпели революционные изменения благодаря развитию искусственного интеллекта и цифровой обработки сигналов. Компьютерная телефония, смартфоны, голосовые интерфейсы и ассистенты, а также системы голосовой биометрической идентификации стали неотъемлемой частью инфраструктуры безопасности, связи и управления. Однако широкое внедрение этих технологий в разных областях и приложениях создает новые вызовы в области защиты речевой информации (ЗРИ) [1].

Безопасность речевых данных в существующих приложениях.

Компьютерная телефония встраивается в системы безопасности как канал передачи тревожных сообщений и средство оперативной связи. В корпоративных системах IP-телефония с шифрованием и аутентификацией обеспечивает надежную защиту голосовой связи.

Голосовые ассистенты на базе ИИ (Siri, Alexa, Алиса) используются в системах "умного дома" и промышленной автоматизации, что требует защиты от несанкционированного доступа и речевых атак.

Биометрическая идентификация по голосу примется в банковской сфере и системах контроля доступа. Современные технологии создания голосовых клонов с помощью нейросетей представляют серьезную угрозу таким системам. Вопрос решается, но медленно [2, 3].

Угрозы защите речевой информации. Основные угрозы ЗРИ сегодня включают: перехват и запись речевых сообщений; создание лживых речевых команд с помощью ИИ; манипуляцию голосовой биометрии; мониторинг речевого трафика и др. Отсюда необходима объективная оценка эффективности систем ЗРИ через их тестирование [1].

Речеподобные сигналы для тестирования систем защиты. Для экономичного и качественного тестирования систем ЗРИ предлагается использование синтетических речеподобных сигналов (РПС), которые имитируют основные акустические параметры речи (форманты, частота основного тона, обертона и др.), позволяют проводить массовое тестирование систем ЗРИ, снижают затраты на создание новых систем и тестовых баз данных (речевых корпусов) и могут генерироваться с заданными параметрами для моделирования различных атак

Нейросетевые технологии в моделировании систем защиты позволяют автоматически генерировать тестовые РПС, имитировать атаки на системы ЗРИ, оценивать устойчивость и эффективность защитных механизмов, адаптировать системы защиты к новым угрозам.

Предложена модель системы защиты РИ на основе: генеративно-состязательных сетей GAN – для создания РПС; рекуррентных нейросетей – для анализа речевых паттернов; методов аугментации данных – для расширения тестовых выборок: блока речепреобразования – для моделирования приложения, реализуемого в тестируемой системе ЗРИ.

Эксперименты показали:

– использование РПС ускоряет тестирование систем ЗРИ до 40%;

– комбинирование реальной и синтетической речи даёт лучшие результаты.

Заключение

Современные речевые технологии требуют новых подходов к защите информации. Применение речеподобных сигналов и искусственного интеллекта позволяет: экономично тестировать системы защиты, быстро адаптироваться к новым угрозам, обеспечивать качественную защиту речевых коммуникаций.

Перспективные направления исследований нацелены на разработку методов обнаружения глубоких голосовых подделок и создание самообучающихся систем ЗРИ.

Список литературы

1. Paracha, A., Arshad, J., Farah, M. et al. Machine learning security and privacy: a review of threats and countermeasures. EURASIP J. on Info. Security 2024, 10 (2024). https://doi.org/10.1186/s13635-024-00158-3.

2. Borrelli, C., Bestagini, P., Antonacci, F. et al. Synthetic speech detection through short-term and long-term prediction traces. EURASIP J. on Info. Security 2021, 2 (2021). https://doi.org/10.1186/s13635-021-00116-3.

3. Евсюков М.В., Путято М.М., Макарян А.С., Немчинова В.О. Методы защиты в современных системах голосовой аутентификации. // Прикаспийский журнал: управление и высокие технологии, № 3 (59), 2022, с. 84–92.

УДК 004.056

С.В. ДВОРЯНКИН[1,2], А.С. ВАНИЧКИНА[2]

[1]Финансовый университет при Правительстве Российской Федерации, Москва
[2]Московский государственный лингвистический университет

РЕЧЕВАЯ ИНФОРМАЦИЯ, РЕЧЕВОЙ И РЕЧЕПОДОБНЫЙ СИГНАЛЫ

Определены ключевые для информационной безопасности (ИБ) понятия речевой информации (РИ), речевых (РС) и речеподобных сигналов (РПС). РИ рассматривается как содержательная составляющая речевого сообщения, извлекаемая из РС – носителя. Показано, что инвариантной основой РИ является гармоническая структура РС, сохраняющаяся при его передаче. На этом принципе основаны методы защиты РИ с использованием РПС – сигналов, близких к речевым по физическим свойствам, но зачастую не несущих смысловой нагрузки.

Введение. Вопрос определения и извлечения речевой информации (РИ) из речевого сигнала (РС) остается актуальной и сложной задачей, особенно в контексте информационной безопасности (ИБ). Информация в звуке – это инвариантная структура гармонических составляющих, сохраняющаяся несмотря на изменения формы РС при его передаче [1].

Речевой сигнал и его инвариантная структура. РИ определяется параметрами ее носителя – РС. Сложный звуковой сигнал, включая речь, на коротком интервале может быть представлен суммой узкополосных сигналов (микрогармоник) по Гильберту:

$$S(t) = \sum_{k=0}^{K} A_k(t)\cos\big(wt + \varphi_{0k}(t)\big). \quad (1)$$

Из трех параметров микрогармоники (амплитуда, частота, фаза) инвариантной к трансформациям при передаче является именно частота. Это свойство лежит в основе спектрального анализа и синтеза РС с использованием преобразований Фурье, Фурье-Гаусса и др.

Таким образом, хотя форма волны РС может существенно изменяться, в нем сохраняется инвариантная гармоническая структура – частотные треки локальных максимумов элементарных микрогармоник (МГ). Эта структура обеспечивает надежную передачу смысла речевого сообщения.

Для визуализации этого подхода используются спектрограммы:

- Узкополосные сонограммы – лучше отображают гармоническую структуру (обертоны голоса на вокализованных участках).
- Широкополосные сонограммы – лучше визуализируют формантную структуру (резонансные частоты речевого тракта).

Совместно гармоническая и формантная структуры, присутствующие на динамических спектрограммах, полностью определяют разборчивость и узнаваемость речи (рис. 1). Из голосовой базы диктора по одним только трекам МГ можно восстановить формантную структуру и смысл РС.

На рис. 1 РС представлен как совокупность наложения формантной и гармонической структуры. При помощи разных алгоритмов инверсии по изображению спектрограммы можно синтезировать РС или РПС с заданными свойствами, создавая нужную оценку фазовой составляющей.

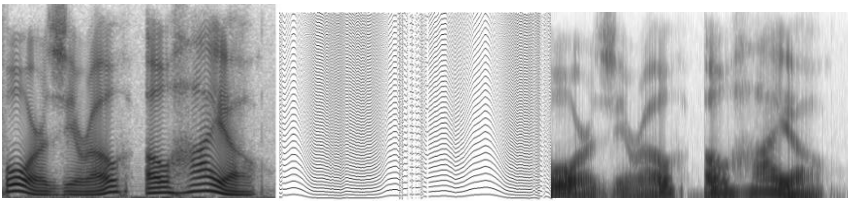

Рис. 1. Спектрограммы РС: узкополосная полутоновая (слева); бинарная гармоническая структура (центр); широкополосная п/тоновая.

РПС сигналы. Тогда к РПС будем относить такие сигналы, которые, как и РС, можно представить в виде совокупности гармонической структуры и квазиформантной огибающей во всем речевом диапазоне частот. Т.е. РПС – это сигналы, близкие к речевым по акустическим или структурным свойствам, что позволяет передавать и хранить их так же, как и обычный РС, при этом они могут и не нести смысловой нагрузки.

Заключение. Таким образом, инвариантными элементами любого звукового сигнала, включая РС и РПС, являются элементарные микрогармоники. Структура частот этих МГ (треки МГ), если она не разрушается при передаче, служит лучшим инвариантным (информационным) отражением сигнала [1]. Этот принцип неявно используется при слуховом восприятии и явно в инструментальных технологиях спектрального анализа, обработки и защиты РИ.

Разработанные в последние годы технологии инверсии спектрограмм позволяют создавать и реконструировать РПС со спектром, соответствующим п/тоновому изображению любого содержания с высоким качеством, даже без информации о первоначальной фазе.

Список литературы

1. Женило В.Р. Информация, звук и преобразование Фурье-Гаусса // Сб. трудов XV Межд. конф. «Информатизация и информационная безопасность правоохранительных органов». М.: 2006.

УДК 004.932:004.056

А.А. БОБЫЛЕВА, М.Т. ЧОЧИЕВА

МИРЭА – Российский технологический университет, Москва

СПОСОБЫ ОБНАРУЖЕНИЯ ЦИФРОВЫХ СЛЕДОВ НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ В ТЕКСТЕ

В работе исследованы современные подходы выявления характерных признаков, которые позволяют отличить тексты, созданные с помощью искусственных нейросетевых алгоритмов. В статье рассмотрены достоинства и недостатки этих подходов, а также пути решения основных проблем. Результаты демонстрируют возможность повышения точности определения искусственного происхождения текста, что важно для проверки подлинности информации.

В настоящее время искусственные нейросетевые алгоритмы достигли такого уровня развития, что способны генерировать тексты, практически неотличимые от созданных человеком. Это порождает фундаментальную проблему проверки подлинности информации в академической, медийной и деловой сферах. Возникает критическая необходимость в исследовании и разработке надежных подходов для выявления характерных признаков, или «цифровых следов». Задача данной работы – исследовать современные методы детекции, рассмотреть их достоинства и недостатки, а также предложить пути решения существующих проблем для повышения точности определения искусственного происхождения текста.

Современные методы обнаружения ИИ-текстов можно классифицировать по двум основным парадигмам: анализ на основе статистических метрик и использование моделей машинного обучения.

Подход, основанный на статистическом анализе, является одним из первых и наиболее изученных. Он базируется на гипотезе о том, что процесс генерации текста языковой моделью оставляет статистически измеримые следы, отличные от тех, что свойственны человеческой речи. Ключевыми метриками здесь выступают перплексия (мера неопределенности текста) и распределение вероятностей слов согласно закону Ципфа. Тексты, созданные ИИ, часто имеют более низкую перплексию и более предсказуемую структуру. К достоинствам данного подхода относятся вычислительная эффективность и интерпретируемость результатов. Однако его главный недостаток — снижение точности при работе с текстами, сгенерированными продвинутыми LLM, которые лучше аппроксимируют статистические свойства естественного языка [1].

Второй, более превалирующий подход, основан на моделях машинного обучения, в частности, на тонкой настройке трансформерных архитектур, таких как BERT и RoBERTa. Эти модели обучаются на больших сбалансированных наборах данных, содержащих как человеческие, так и машинные тексты, и функционируют как бинарные классификаторы. Основное достоинство этого метода – высокая точность детекции, поскольку модель учится распознавать сложные, нелинейные паттерны и семантические артефакты генерации. К недостаткам относятся высокие вычислительные затраты на обучение и эксплуатацию, а также уязвимость к состязательным атакам, когда текст незначительно изменяется для обмана детектора [2].

Современные методы обнаружения цифровых следов в текстах включают технологию «цифровых водяных знаков», которая встраивает скрытый сигнал в процесс генерации текста и позволяет точно определить источник с помощью статистического теста. Этот подход устойчив к изменениям текста, но требует стандартизации и поддержки от разработчиков моделей. Другой инновационный метод – zero-shot детекторы, например DetectGPT, которые на основе анализа специфических статистических свойств текста способны выявлять искусственно сгенерированные фрагменты без предварительного обучения. Оба подхода показывают высокую эффективность в обнаружении нейросетевого происхождения текста [3].

Проведенный анализ демонстрирует, что задача обнаружения ИИ-сгенерированного текста представляет собой сложную научную проблему, находящуюся в состоянии постоянного развития – своеобразной «гонки вооружений» между генеративными и детективными моделями.

Список литературы

1. Соколов И.В. Статистические метрики и их применение для детекции машинной генерации текста // Вопросы кибернетики. – 2023. – № 2 (45). – С. 89–102 (дата обращения: 20.10.2025).

2. Петров А.Н., Смирнова Е.К. Классификация текстов с использованием трансформерных архитектур BERT и RoBERTa для задач аутентификации авторства // Искусственный интеллект и принятие решений. – 2022. – № 4. – С. 56–68 (дата обращения: 20.10.2025).

3. Григорьев М.С. Гибридные модели для повышения устойчивости детекторов искусственного текста к состязательным атакам // Доклады Академии наук. – 2024. – Т. 488, № 1. – С. 11–15 (дата обращения: 20.10.2025).

УДК 004.056

А.В. ТОКАРЬ, А.Е. ЯКОВЛЕВА

МИРЭА – Российский технологический университет, Москва

НАДЁЖНЫЕ СПОСОБЫ ПРОТИВОДЕЙСТВИЯ ВИШИНГУ В СОВРЕМЕННЫХ УСЛОВИЯХ

В качестве центральной темы работы выступает такая форма мошенничества, как «вишинг» («voice phishing»). Основной целью является систематизация мер противодействия ему. В статье проводится анализ тактик злоумышленников и предлагаются практические рекомендации для защиты пользователей.

Среди различных фишинговых атак отдельного внимания заслуживает вишинг – голосовой фишинг. Подделка номеров (spoofing), использование синтезированного голоса с помощью AI-систем и сценариев социального воздействия позволяют злоумышленникам обмануть даже подготовленных сотрудников и добиться немедленной передачи кодов, паролей или одобрения транзакций [1].

Вишинг реализуется как целенаправленная кампания, обычно имеющая три стадии: сбор данных, разработка сценария взаимодействия и непосредственное проведение. Сбор данных включает анализ профилей в социальных сетях, изучение утечек баз данных и даже автоматизированный парсинг открытых источников. Это позволяет составить «портрет» адресата и выбрать наиболее правдоподобную маску (сотрудник банка, техподдержки, сотрудник органов, родственник в беде и т. п.). На следующем этапе подготовленный сценарий прорабатывается с учётом возможных реакций и создаётся набор реплик и контрреплик. Наконец, сам звонок – момент применения психологических триггеров: срочность, страх, авторитет, обращение к эмпатии, которые направлены на снижение критичности суждений и ускорение принятия неверного решения [2].

Стандартные меры защиты, ориентированные на анализ телефонных номеров, оказываются недостаточными, так как искусственно сгенерированный голос может воспроизводить индивидуальные интонации, тембр и речевые особенности конкретного человека. Также расследование вишинга осложнено техническими барьерами: международная VoIP-маршрутизация, анонимные шлюзы и подмена номеров препятствуют трассировке. Тем не менее вишинговые атаки всё же имеют определённые закономерности. Эффективный ответ на угрозу вишинга должен опираться на интегрированную стратегию, объединяющую технические (табл. 1) и организационные контрмеры (табл. 2).

Таблица 1. Технические меры защиты

Мера	Описание
Фильтрация вызовов	Аналитика аномалий голосового трафика
Проверка идентификаторов (STIR/SHAKEN)	Стандарт верификации вызывающего номера в ряде юрисдикций
Блокировка мошеннических шлюзов	Чёрные списки известных атакующих
Ограничение голосовых подтверждений	Введение двуканальной аутентификации для критичных операций

Таблица 2. Организационные меры защиты

Мера	Описание
Минимизация привилегий сотрудников	Выдача только необходимых прав
Сценарные учения и тренировки	Обучение распознаванию вишинговых звонков
Просветительская работа	Обучение уязвимых групп, пожилых людей, финансовых сотрудников
Фиксация подозрительных вызовов	Сохранение данных для служб безопасности и правоохранительных органов

Ситуацию усугубляет низкая сообщаемость о преступлениях. Одна из актуальных задач – это улучшение механизмов обмена информацией между операторами связи, финансовыми институтами и правоохранительными органами, а также развитие международных стандартов и практик, позволяющих ускорять блокировку и расследование подобных инцидентов. Хотя вишинг останется угрозой, защита от него достижима. Именно внимательность, базовые правила безопасности и элементарная осведомлённость становятся самым надёжным щитом против вишинга [2].

Список литературы

1. Пекарева В.В. Вишинг и смишинг как нескончаемые угрозы информационной безопасности / В.В. Пекарева // Скиф. Вопросы студенческой науки. – 2024. – № 1(89). – С. 80–83. – EDN LEZYXD.

2. Шинкарук И.П. Вишинг как одна из распространенных форм мошеничества в современном мире / И.П. Шинкарук // За нами будущее: взгляд молодых ученых на инновационное развитие общества: Сборник научных статей 3-й Всероссийской молодежной научной конференции, Курск, 03 июня 2022 года. Том 2. – Курск: Юго-Западный государственный университет, 2022. – С. 285–289. – EDN UKVEWT.

УДК 004.056

П.А. МАКРЕЦОВА, Н.А. СОРОКИН

МИРЭА – Российский технологический университет, Москва

СОВРЕМЕННЫЕ ТРЕНДЫ ПРОТИВОДЕЙСТВИЯ СОЦИАЛЬНОЙ ИНЖЕНЕРИИ

В статье рассматриваются современные тенденции в области защиты от социальной инженерии, акцентируя внимание на ключевых мерах, обеспечивающих устойчивость организаций к манипулятивным атакам. В условиях цифровой трансформации и роста киберпреступности противодействие социальной инженерии требует не только технических решений, но и системного подхода, охватывающего организационные, технологические и поведенческие аспекты.

Введение

Социальная инженерия – это метод манипуляции людьми с целью получения конфиденциальной информации, доступа к системам или совершения действий, выгодных злоумышленнику [1]. В отличие от традиционных кибератак, основанных на эксплуатации технических уязвимостей, социальная инженерия нацелена на «человеческий фактор» – самое уязвимое звено в цепи информационной безопасности. Согласно отчётам Verizon Data Breach Investigations Report за минувший 2024 год, более 71% инцидентов компрометации данных связаны с человеческими ошибками, вызванными фишингом, предтекстингом и другими техниками социальной инженерии.

Новые тренды и практики в социальной инженерии

Атаки через мессенджеры и корпоративные чаты представляют растущую угрозу, чему способствует их повсеместное использование в деловой среде.

В текущем году зафиксированы случаи создания мошенниками фальшивых аккаунтов в платформе Telegram, где они, выдавая себя за руководителей, убеждали сотрудников раскрыть конфиденциальные данные или перевести средства на мошеннические счета под предлогом «срочных проектов». Эффективность подобных атак повышается за счет эксплуатации эмоциональных триггеров.

В периоды кризисов повышенная тревожность провоцирует импульсивные действия, что подтверждается фишинговыми кампаниями во время пандемии COVID-19 и в 2022 году. Психологический механизм

основан на подавлении критического мышления, заставляя жертву действовать в обход процедур верификации.

Дополнительным фактором риска является развитие технологий глубокой подделки (deepfake) и синтеза голоса. Доступность этих инструментов позволяет злоумышленникам создавать убедительные фальсификации.

В условиях отсутствия абсолютно надежных методов детекции ключевой мерой противодействия становится развитие киберграмотности персонала и внедрение строгих регламентов подтверждения распоряжений.

Заключение

Современный ландшафт киберугроз демонстрирует стремительную эволюцию методов социальной инженерии, особенно в условиях широкого внедрения мессенджеров в корпоративную коммуникацию и бурного развития генеративного искусственного интеллекта.

Атаки через платформу Telegram и другие чат-платформы, усиленные использованием deepfake-технологий и синтезированного голоса, становятся всё более изощрёнными и трудно обнаружимыми.

При этом злоумышленники целенаправленно эксплуатируют человеческий фактор – в частности, эмоциональные состояния, такие как страх, тревога и ощущение срочности, – чтобы обойти даже технически надёжные системы защиты.

Таким образом повышение эффективности средств противодействия социальной инженерии является весьма актуальным в настоящий момент.

Список литературы

1. The Art of Deception: Controlling the Human Element of Security / K. Mitnick, W. L. Simon. – Indianapolis: Wiley Publishing, Inc., 2002. – 320 p. – ISBN 0-7645-4280-X.

2. Тимофеев К. Социальная инженерия: новые тренды и тактики. [Электронный ресурс] // IT World. – 24.12.2024. – URL: https://www.it-world.ru/security/ 25wpdi0jvj1cgkkwg gwwc44os0wc8gg.html (дата обращения: 08.10.2025).

3. Нохрин М. Социальная инженерия: методы атак и как защититься [Электронный ресурс] // SpectrumData. URL: https://spectrumdata.ru/blog/proverka-soiskatelya/chto-takoe-sotsialnaya-inzheneriya-i-kak-ot-nee-zashchititsya/ (дата обращения: 08.10.2025).

УДК 004.056

Н.С. БУБНОВ, К.А. ЕРМАЧКОВА

МИРЭА – Российский технологический университет, Москва

СОВРЕМЕННЫЕ ТРЕНДЫ АНОНИМИЗАЦИИ ЦИФРОВЫХ СЛЕДОВ ЛИЧНОСТИ

Целью работы является анализ современных методов анонимизации цифровых следов личности в контексте выявления оптимальных подходов к обеспечению баланса между конфиденциальностью информации и её аналитической ценностью. В результате предлагается развитие гибридных систем анонимизации, устойчивых к современным угрозам деанонимизации при сохранении аналитической ценности.

Введение

В современных условиях цифровизации активность пользователей в сети порождает огромные массивы данных. Любые действия в интернете, включая поисковые запросы, публикации и взаимодействие с онлайн-сервисами, оставляют информацию, которая может быть использована как во благо, так и против самого пользователя. Без надлежащей защиты эти данные представляют угрозу конфиденциальности и могут быть использованы для установления личности и анализа поведения. Эффективное обезличивание персональных данных позволяет сохранять возможность анализа информации, минимизировав при этом риски нарушения приватности. Для обеспечения защиты применяются различные методы анонимизации данных, каждый из которых имеет собственные особенности и ограничения.

Изложение основного материала

Наиболее распространёнными методами анонимизации, по мнению М.А. Полтавцевой, В.В. Платонова и др., являются псевдонимизация, обобщение, подавление, шифрование, замена чувствительных данных токенами и использование синтетических данных [1].

Псевдонимизация предполагает замену прямых идентификаторов, таких как имена или номера документов, на искусственные обозначения или коды, что позволяет сохранить структуру данных и их пригодность для анализа. Обобщение и подавление направлены на снижение детализации информации, например, вместо точной даты рождения указывается только год или возрастная группа, а отдельные чувствительные поля полностью удаляются. Шифрование и

использование токенов позволяют преобразовывать данные в зашифрованный или замещённый вид, что предотвращает несанкционированный доступ к исходным значениям Использование синтетических данных предполагает генерацию искусственных наборов, статистически схожих с реальными, но не содержащих персональной информации [1].

Данные методы обеспечивают высокий уровень анонимности информации при сохранении ее аналитической ценности. При этом чрезмерное обезличивание данных может привести к потере информативности и ухудшению качества аналитических выводов. Оптимальный баланс между конфиденциальностью и полезностью данных позволяет достичь комбинирование перечисленных методов, включая их применение на разных уровнях защиты. Поэтому современные тренды анонимизации основаны на использовании гибридных систем, устойчивых к угрозам деанонимизации и обеспечивающих соотнесение глубины анонимизации с уровнем рисков и задачами по обработке данных.

Заключение

Современные подходы к анонимизации цифровых следов личности базируются на сочетании классических и инновационных методов обработки данных. Их применение позволяет эффективно защищать персональную информацию и одновременно сохранять возможность аналитической работы. Главной задачей в ближайшие годы станет формирование гибридных систем анонимизации, сочетающих криптографические методы, статистические модели и синтетические технологии. Такие решения обеспечат устойчивость к новым формам инструментов для нарушения конфиденциальности данных о цифровых следах и создадут основу для безопасного использования информационных систем в условиях активного развития искусственного интеллекта и больших данных.

Список литературы

1. Полтавцева, М.А. Методы анонимизации персональных данных / М.А. Полтавцева, В.В. Платонов, А.Ф. Супрун, П.В. Семьянов. // SAEC. 2024. № 2. URL: https://cyberleninka.ru/article/n/metody-anonimizatsii-personalnyh-dannyh (дата обращения: 15.10.2025).

УДК 004.056

В.А. ВОРОНЕНКО, И.К. ЗЕМСКОВ

МИРЭА – Российский технологический университет, Москва

СОВРЕМЕННЫЕ МЕТОДЫ ВИЗУАЛИЗАЦИИ ТЕКСТОВОЙ ИНФОРМАЦИИ В СФЕРЕ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

В статье рассматриваются современные методы визуализации текстовых данных в области информационной безопасности. Проанализированы функциональные возможности, гибкость и сложность внедрения готовых решений и открытых библиотек Python. На основе анализа предложены критерии выбора инструментов визуализации с учётом специфики и ресурсов организации.

Введение

Современные организации генерируют большие объёмы текстовых данных в области безопасности [1]. Их визуализация необходима для эффективного управления ИБ, так как традиционные текстовые методы не справляются с большими массивами информации.

Ключевые задачи визуализации – оперативное выявление аномалий и инцидентов, наглядное отображение связей между событиями и интерактивный анализ паттернов [2].

Решения разделяются на коммерческие продукты (SIEM, BI-платформы) и открытые Python-библиотеки для создания гибких инструментов [3].

Коммерческие решения для визуализации

Платформа Visiology – российская BI-система для анализа и визуализации данных с компонентами для интерактивных отчётов и высокопроизводительной обработки. Она позволяет визуализировать безопасность в реальном времени, тренды и географию угроз.

SIEM-системы объединяют управление информацией и событиями безопасности, предоставляя агрегацию данных, корреляцию событий, оповещения и визуализацию.

Grafana – открытая платформа для мониторинга и визуализации метрик, широко используемая в ИБ для создания интерактивных дашбордов и отслеживания событий.

Программные библиотеки Python

Matplotlib – базовая библиотека для статических графиков [4], применяемая для анализа временных рядов и распределений метрик.

Seaborn – высокоуровневая библиотека на основе Matplotlib, упрощающая статистическую визуализацию, полезна для анализа корреляций и представления матриц связей.

Plotly позволяет создавать интерактивные графики и веб-приложения, используется для интерактивных дашбордов и 3D-визуализаций сетевых данных.

NetworkX специализируется на создании и анализе графов [5], полезна для визуализации сетевых атак и топологии инфраструктуры.

Сравнительный анализ

SIEM – системы предлагают полный функционал, но требуют значительных вложений и имеют высокую сложность внедрения.

Visiology предоставляет мощные визуализации с профессиональной поддержкой, но требует настройки для ИБ-задач.

Открытые библиотеки Python гибки, бесплатны и имеют активные сообщества, но нуждаются в квалификации для разработки

Заключение

Выбор решения зависит от требований, ресурсов и экспертизы организации. Замещение коммерческих платформ открытыми библиотеками возможно, но требует тщательного планирования. Для небольших организаций предпочтительны открытые решения, для крупных – коммерческие с гибридным подходом

Список литературы

1. Милославская Н.Г., Толстой А.И., Бирюков А.А. Визуализация информации при управлении информационной безопасностью информационной инфраструктуры организации // Вопросы защиты информации. 2014. № 2. С. 32–48.

2. Котенко И.В., Коломеец М.В., Жернова К.Н., Чечулин А.А. Визуальная аналитика для информационной безопасности: области применения, задачи и модели визуализации // Вопросы кибербезопасности. 2021. № 4(44). С. 2–15.

3. Зеличенок И.Ю. Визуализация и предобработка событий информационной безопасности на основе данных CICIDS17 // Информационные технологии и телекоммуникации. 2021. Т. 9, № 4. С. 49–55.

4. Hunter J.D. Matplotlib: A 2D graphics environment // Computing in Science & Engineering. 2007. Vol. 9, No. 3. P. 90-95.

5. Hagberg A.A., Schult D.A., Swart P.J. Exploring network structure, dynamics, and function using NetworkX // Proceedings of the 7th Python in Science Conference. 2008. P. 11–15.

УДК 004.056

П.В. АЛЕКСАНДРОВА, В.А. АРХИПОВ, В.А. ЖУТЯЙКИНА

МИРЭА – Российский технологический университет, Москва

АНАЛИЗ ПРОБЛЕМ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ СОВРЕМЕННЫХ СТРИМИНГОВЫХ СЕРВИСОВ

В статье рассматриваются ключевые проблемы информационной безопасности стриминговых сервисов в 2025 г. Выявлены основные угрозы – фишинг, вредоносное ПО, DDoS-атаки и пиратство, приводящие к утечке данных и нарушению работы платформ. Отмечено, что развитие облачной инфраструктуры повышает риски из-за ошибок конфигурации и слабой защиты API. Предложены направления повышения безопасности: внедрение концепции Zero Trust и MFA, применение DLP и SIEM-систем, микросегментация сетей и использование ИИ для обнаружения аномалий и фишинга.

Введение

Современные стриминговые платформы (Netflix, YouTube, Twitch, Kinopoisk) все чаще становятся мишенью злоумышленников, стремящихся получить доступ к учетным записям пользователей и платежным данным. По данным «Лаборатории Касперского», за последние два года было зафиксировано 96 288 попыток атак, маскирующихся под названия стриминговых сервисов – это троянцы и шифровальщики, похищающие личную информацию, при этом троянцы остаются доминирующим типом угроз. Злоумышленники активно используют брендовые названия сервисов, чтобы распространять опасные файлы, создавая поддельные приложения и сайты [1].

Кроме вредоносного ПО, серьезную угрозу представляют DDoS-атаки на стриминговые платформы. В 2025 г. отмечен рост числа атак по протоколам туннелирования GRE и L7 OSI, которые не предусматривают шифрование трафика и аутентификацию, а также использование изощренных схем с атакой на несуществующие поддомены [2]. Эти тактики направлены на выведение серверов из строя и нарушение непрерывности трансляций.

Уязвимости облачной инфраструктуры и архитектурные риски

Большинство стриминговых сервисов используют гибридные или облачные среды для хранения контента и данных пользователей. Такие инфраструктуры часто становятся целью атак, направленных на

эксплуатацию неправильно настроенных прав доступа, слабых паролей или открытых S3-хранилищ. Ошибки конфигурации и отсутствие Zero Trust-контроля делают стриминговые экосистемы уязвимыми даже без прямого вмешательства человека.

Одним из наиболее эффективных инструментов злоумышленников остается фишинг – создание поддельных сайтов или электронных писем, имитирующих интерфейс стриминговой платформы. Через такие атаки мошенники получают пароли, токены авторизации и банковские данные подписчиков. Вариации вишинга и смишинга также активно применяются в связке с рекламными рассылками и акционными предложениями.

Проблема пиратства и обхода систем DRM (Digital Rights Management) остается значимой. Современные методы взлома позволяют воспроизводить защищенный контент через нелегальные зеркала или перехваты потоков. Это не только снижает доходы правообладателей, но и способствует распространению зловредов через фальшивые «бесплатные» версии сервисов.

Основными направлениями повышения уровня безопасности стриминговых сервисов являются интеграция концепции Zero Trust и многофакторной аутентификации для усиленной защиты учетных записей, внедрение DLP- и SIEM-систем для раннего выявления утечек данных, переход к архитектурам с микросегментацией и контролем облачных API, а также использование алгоритмов машинного обучения для обнаружения аномалий поведения пользователей и предотвращения фишинговых атак.

Заключение

Таким образом, проблемы информационной безопасности в стриминговых сервисах в 2025 г. связаны с сочетанием классических угроз (вредоносное ПО и фишинг) и современных рисков, обусловленных облачными архитектурами, масштабом аудитории и ростом кибермошенничества. Эффективное решение требует комплексного подхода – объединяющего технические, организационные и поведенческие меры защиты.

Список литературы

1. Стрим, клик, риск: как киберпреступники пользуются популярностью любимых стриминговых сервисов, фильмов, сериалов и аниме поколения Z [Электронный ресурс] // Лаборатория Касперского. – М.:2025. – URL: https://www.kaspersky.com/blog/gen-z-streaming-report-2025 (дата обращения: 21.10.2025).

2. DDoS-атаки сегодня: новые типы киберугроз в 2025 году [Электронный ресурс] // DDoS Guard. – М.: 2025. – URL: https://ddos-guard.ru/blog/ddos-ataka-segodnya-novye-tipy-kiberugroz (дата обращения: 21.10.2025).

УДК 004.056

Л.П. БОЖКО, Л.А. БОЧАРОВ

МИРЭА – Российский технологический университет, Москва

ПРИМЕНЕНИЕ ОБОРОНИТЕЛЬНОЙ ДИСТИЛЛЯЦИИ ДЛЯ ЗАЩИТЫ ИНФОРМАЦИОННЫХ СИСТЕМ ОТ НОВЫХ ВИДОВ КИБЕРАТАК

В статье рассматривается применение метода оборонительной дистилляции для защиты информационных систем, основанных на машинном обучении, от новых видов атак, в частности состязательных (adversarial) атак. Оборонительная дистилляция позволяет повысить устойчивость моделей искусственного интеллекта за счёт обучения на сглаженных метках, что снижает чувствительность к малым изменениям входных данных, используемым злоумышленниками для обмана систем.

В настоящее время информационные системы, основанные на технологиях искусственного интеллекта (ИИ) и машинного обучения, получили широкое распространение в различных областях – от финансов и здравоохранения до кибербезопасности и промышленности. Однако с ростом их значимости возросли и риски, связанные с устаревшими и новыми видами атак, среди которых особое место занимают состязательные (adversarial) атаки. Они представляют собой тщательно скоординированные попытки злоумышленников ввести минимальные искажения в исходные данные с целью введения модели в заблуждение. В данной статье рассматривается применение оборонительной дистилляции – перспективного метода повышения устойчивости моделей машинного обучения к такого рода атакам.

Изначально метод дистилляции был разработан в машинном обучении как способ передачи знаний от более большой и сложной модели к более компактной и простой модели при сохранении точности и обобщающей способности. Оборонительная дистилляция переосмысливает этот процесс применительно к задаче устойчивости: она предусматривает обучение модели на «смягчённых» метках, сгенерированных при повышенной температуре на этапе дистилляции. Это приводит к тому, что модель становится менее восприимчивой к незначительным вариациям во входных данных, которые злоумышленники обычно используют для атак.

Механизм работы оборонительной дистилляции состоит в следующем:

- Исходная модель, обученная на исходных данных, используется для генерации вероятностных меток (smoothed labels) с использованием параметра температуры.
- Новая модель обучается на этих вероятностных метках, что заставляет её формировать более сглаженные границы решений.
- В результате модель снижает чувствительность к мелким возмущениям, которые служат механизмом атак.

Использование оборонительной дистилляции особенно актуально для систем, где модели ИИ функционируют в условиях повышенного риска атак, например, системы распознавания образов, биометрическая аутентификация, обнаружение вторжений и пр. Такой защитный слой позволяет эффективно противодействовать новым видам атак, которые нацелены на эксплуатацию уязвимостей моделей ИИ, создавая искусственно искажённые входные данные.

Оборонительная дистилляция представляет собой эффективный метод повышения устойчивости моделей машинного обучения к новым видам атак, в частности, к состязательным атакам, направленным на искажение входных данных. Внедрение такого подхода в информационные системы значительно улучшает их защитные возможности, делая системы более надёжными и безопасными в условиях эволюции современных киберугроз. Комплексное использование оборонительной дистилляции вместе с другими защитными методиками позволит существенно снизить риски и обеспечить стабильность работы интеллектуальных систем в будущем.

Список литературы

1. Петров В.А., Иванов С.К. Методы защиты искусственного интеллекта от атак / Вестник кибербезопасности. – 2024. – № 12. – С. 45–53.
2. Смирнова Е.В. Состязательные атаки и методы их предотвращения в системах машинного обучения / Информационная безопасность. – 2025. – Т. 18, № 3. – С. 112–121.
3. Козлов Д.П. Оборонительная дистилляция: принципы и применение / Журнал современных технологий. – 2023. – Т. 7, № 4. – С. 33–41.
4. Иванова М.Н. Современные методы защиты интеллектуальных систем / Информационные технологии и безопасность. – 2025. – Т. 10, № 1. – С. 24–30.
5. Wang H., Liu Y. Defensive Distillation for Robustness Against Adversarial Examples / Journal of Machine Learning Research. – 2016. – Vol. 18, No. 65. – P. 1–10.
6. Пырков Н.С. Исследования уязвимостей информационной безопасности в сфере киберспортивной деятельности / Н.С. Пырков, А.Д. Лазовик, В.А. Шутов // Кибербезопасность: технические и правовые аспекты защиты информации: Сборник научных трудов по итогам III ежегодной национальной научно-практической конференции, Москва, 23–24 апреля 2024 года. – М.: МИРЭА., 2024. С. 122–125. – EDN SXZJPN.

УДК 004.056

С.Ж. САРБАЛИНА, А.А. УШАКОВА

МИРЭА – Российский технологический университет, Москва

СПОСОБЫ ЗАЩИТЫ ОТ УЯЗВИМОСТЕЙ БОЛЬШИХ ЛИНГВИСТИЧЕСКИХ МОДЕЛЕЙ

В статье рассматриваются современные типы уязвимостей больших лингвистических моделей (LLM), включая prompt injection, jailbreaking, извлечение тренировочных данных и методы отравления обучающих выборок. Проводится анализ механизмов реализации этих атак, оценка их влияния на безопасность и конфиденциальность обработки информации, а также предлагаются эффективные методики защиты, такие как промпт-фильтрация, контроль целостности данных и автоматизированный аудит запросов к LLM.

Введение

Исследования показывают, что LLM демонстрируют высокую способность к генерации естественного текста, но их архитектура и методы обработки данных всё ещё делают их уязвимыми для сложных атак. Даже последние версии, такие как GPT-4, несмотря на более совершенные фильтры, подвержены изощрённым техникам обхода, например, атаке «Deceptive Delight» – слиянию вредоносного контента с легитимным приведением к нелегальным результатам всего за два обращения. LLM используются для автоматизации поиска уязвимостей в коде и разработки инструментов защиты, однако их слабости часто провоцируют появление новых угроз – как при генерации кода, так и при анализе угроз. Было показано, что хотя модели показывают высокую точность в обнаружении простых уязвимостей (более 80% в простых задачах), при анализе сложного кода их точность значительно падает. Кроме того, модели склонны к высокому уровню ложноположительных срабатываний – до 63% даже у лучших моделей [1].

Особенности низкоинтенсивных DDoS-атак

Современные уязвимости больших языковых моделей (LLM) включают широкий спектр атак, начиная от манипуляций запросами (prompt injection) и обхода защитных мер (jailbreak), до атак на извлечение или утечку данных и внедрение бэкдоров через отравление данных (data poisoning) [2].

Основные векторы атак. Промпт-инъекции. Относятся к внедрению зловредных или манипулирующих инструкций в запрос, что позволяет получить нежелательный или некорректный ответ.

Jailbreak-атаки. Позволяют злоумышленнику обойти фильтры и запреты, внедрённые в модель, открывая доступ к несанкционированному функционалу или запрещённому содержимому.

Извлечение данных (data extraction attacks). Использование LLM с целью получения скрытых, обучающих или конфиденциальных данных, которые модель могла усвоить в процессе обучения, – например, персональных данных или корпоративных секретов [3].

Отравление данных (data poisoning/backdoor). Внедрение в обучающие датасеты вредоносных образцов, что может привести к закладке скрытых “бэкдоров” или занижению безопасности модели.

Состязательные атаки (adversarial attacks): Создание специально подобранных входных данных, провоцирующих неправильные или опасные ответы модели.

Традиционные подходы к обеспечению безопасности не всегда применимы к LLM из-за «чёрного ящика» их архитектуры и динамично меняющихся векторов атак. Это требует комплексных мер управления рисками, регулярного аудита данных и постоянного совершенствования защитных механизмов, включая фильтрацию промптов и многоуровневую проверку результатов [4].

Заключение

Современные подходы к повышению уровня ИБ требуют междисциплинарного подхода с применением классических и инновационных средств, что остаётся одной из самых актуальных тем на стыке искусственного интеллекта и кибербезопасности.

Список литературы

1. Байбара Б.В. Уязвимости больших языковых моделей: анализ и методы защиты // Информационная безопасность. – 2025. – № 2. – С. 32–44.

2. Безопасность и уязвимости больших языковых моделей // Научный вестник. – 2024. – 14 окт. – URL: https://nvjournal.ru/article/BEZOPASNOST_I_ UJaZVIMOSTI__BOLShIH_ JaZYKOVYH_MODELEJ/nvjournal

3. Герасименко Д.В. Извлечение тренировочных данных: риски и решения в контексте безопасности LLM // Киберинформатика. – 2024. – № 2. – С. 21–35.

4. Котенко И.В. Использование больших языковых моделей для поиска угроз кибербезопасности на основе методов глубокого обучения: анализ // Кибернетика. – 2024. – № 3. – С. 57–62.

УДК 004.056

Ю.В. НОГИНА

МИРЭА – Российский технологический университет, Москва

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И МАШИННОГО ОБУЧЕНИЯ ДЛЯ АВТОМАТИЗИРОВАННОГО РЕАГИРОВАНИЯ НА КИБЕРИНЦИДЕНТЫ И ИНЦИДЕНТЫ БЕЗОПАСНОСТИ

Современные методы кибербезопасности используют искусственный интеллект и машинное обучение для улучшения обнаружения и реагирования на киберинциденты. Эти технологии автоматизируют анализ данных, быстро выявляют угрозы и принимают решения по их нейтрализации. Автоматизация сокращает время реакции и снижает нагрузку на специалистов, повышая устойчивость систем безопасности.

Введение

Традиционные системы автоматизации в информационной безопасности основаны на фиксированных правилах и сигнатурах и хорошо справляются с известными угрозами, но малоприменимы против новых, быстро меняющихся кибератак. Искусственный интеллект и машинное обучение предлагают адаптивный подход с самообучением и выявлением неизвестных атак, что значительно укрепляет защиту и позволяет быстро реагировать на новые угрозы [1].

Использование искусственного интеллекта и машинного обучения для автоматизированного реагирования на киберинциденты

Современные системы класса SIEM (Security Information and Event Management), SOAR (Security Orchestration, Automation and Response) и NDR (Network Detection and Response) все активнее используют алгоритмы искусственного интеллекта для комплексного анализа огромных массивов данных в режиме реального времени [2]. Эти интеллектуальные системы эффективно выявляют подозрительные паттерны и аномалии в поведении пользователей, сетевом трафике и логах, обеспечивая глубокий и оперативный анализ текущей ситуации. Более того, применение предиктивной аналитики позволяет им прогнозировать потенциальные кибератаки, что способствует переходу организаций от традиционно реактивного к проактивному уровню защиты.

Интеграция ИИ в такие системы позволяет автоматически инициировать меры реагирования: от изоляции заражённых узлов до

блокировки доступа злоумышленников и уведомления операторов безопасности. Способность к адаптивному обучению снижает число ложных срабатываний и повышает эффективность работы центров реагирования. Это значительно сокращает время реакции, оптимизирует ресурсы и повышает устойчивость к киберугрозам.

Использование ИИ снижает нагрузку на команды безопасности, позволяя сосредоточиться на стратегическом управлении. Однако успешное внедрение требует качественных данных и регулярного обновления моделей для эффективности в быстро меняющейся среде угроз. Важно сочетать автоматизацию с экспертным анализом.

Основные вызовы – регулярное обновление моделей и поддержка качественных данных. ИИ активно применяется мировыми лидерами, такими как Microsoft с Defender Copilot, Google с Chronicle Security Operations, а также российскими решениями: Security Vision SOAR, Kaspersky Anti-APT и PT Sandbox. Рынок ИИ в кибербезопасности растёт миллиардами долларов с двузначным годовым темпом, подтверждая свою значимость для будущей защиты информации [3].

Заключение

Таким образом, искусственный интеллект и машинное обучение трансформируют подходы к кибербезопасности, обеспечивая проактивный режим работы систем. Они позволяют значительно ускорить обнаружение угроз, автоматизировать ответные меры и гибко адаптироваться к постоянно эволюционирующим атакам. Одновременно снижение нагрузки на специалистов даёт им возможность сосредоточиться на глубоких стратегических задачах. В условиях растущей сложности и масштабности киберугроз автоматизация на базе ИИ становится одним из ключевых факторов надежной и своевременной защиты цифровых ресурсов.

Список литературы

1. Русаков А.М., Бобырь-Бухановский А.И. Исследование интеллектуальных методов анализа журналов событий для обеспечения информационной безопасности // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и Технические Науки. 2025. – № 06/2. – С. 180–186 DOI 10.37882/2223-2966.2025.06-2.32.

2. Sheeraz M., Durad M.H., Paracha M.A., Mohsin S.M., Kazmi S.N., Maple C. Revolutionizing SIEM Security: An Innovative Correlation Engine Design for Multi-Layered Attack Detection // Sensors. – 2024. – т. 24, № 15. – С. 4901. – DOI: 10.3390/s24154901.

3. Анализ методов корреляции событий безопасности в SIEM-системах. Часть 2 / А.В. Федорченко, Д.С. Левшун, А.А. Чечулин, И.В. Котенко // труды СПИИ-РАН. – 2016. – № 6(49). – С. 208–225. – DOI 10.15622/sp.49.11. – EDN XHFTDR.

УДК 004.056

Е.Б. БРИСКОВА

МИРЭА – Российский технологический университет», Москва

АВТОМАТИЗИРОВАННОЕ ВЫЯВЛЕНИЕ ПРОТИВОРЕЧИЙ В НОРМАТИВНОЙ ДОКУМЕНТАЦИИ С ПОМОЩЬЮ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Предлагается подход к автоматизации процесса выявления скрытых семантических коллизий на основе комбинации методов обработки естественного языка (NLP) и машинного обучения. Рассматриваются этапы классификации нормативных предписаний и поиска противоречий с помощью трансформерных моделей. Разработанная методика направлена на повышение качества и скорости проведения нормативной экспертизы для органов регулирования и корпоративных служб комплаенс.

Введение

В области автоматизированного анализа нормативных текстов можно выделить три основных направления:

1. Использование синтаксических парсеров и методов NLP для извлечения субъектно-предикатных конструкций с последующей проверкой на противоречия с помощью формальной логики. Основной недостаток – низкая адаптивность к языковым вариациям и сложным юридическим оборотам.

2. Построение формальных онтологий предметной области, которые описывают взаимосвязи между терминами и понятиями.

3. Трансляция текстовых норм в формальные логические выражения. Пересечения выявляются путем автоматического логического вывода [1].

Таким образом, существующие решения либо излишне ригидны и не учитывают контекст, либо требуют непозволительно высоких затрат на предварительную настройку.

Предлагаемый подход

Для преодоления ограничений существующих методов предлагается гибридная методика, комбинирующая современные NLP-модели и алгоритмы машинного обучения для семантического анализа. Процесс включает следующие этапы:

1. Использование предобученных трансформерных моделей для, адаптированных для русского языка (таких как RuBERT, DeBERTa).

2. Применение алгоритмов классификации нормативных предписаний для категоризации выделенных фрагментов.

3. На этапе поиска потенциальных противоречий формируются пары норм, и их векторные представления сравниваются.

4. Система дообучается на корпусе текстов, где эксперты вручную разместили примеры конфликтующих и непротиворечивых пар норм.

5. Для наглядности результаты работы системы представляются в виде семантического графа, где узлы – это нормы, а связи – различные отношения (сходство, конфликт). Конфликтующие узлы подсвечиваются, что позволяет эксперту быстро сфокусироваться на проблемных участках документа [2].

Применимость и перспективы

Разработанный подход обладает значительным практическим потенциалом и может быть внедрен в следующие области:

1. Автоматизированная проверка новых законов и стандартов;

2. Помощь в актуализации и гармонизации нормативной базы, устранении устаревших и конфликтующих требований;

3. Обеспечение полноты и непротиворечивости внутренних политик безопасности и процедур, снижение регуляторных рисков;

4. Повышение прозрачности требований для создателей защищенных систем.

В перспективе методика может быть усилена за счет интеграции с системами логического программирования для более строгой формальной проверки.

Заключение

Таким образом автоматизация выявления противоречий в нормативной документации с помощью машинного обучения позволяет сократить время на экспертную проверку и повысить качество нормативно-правового обеспечения информационной безопасности. Применение современных моделей NLP и алгоритмов классификации делает возможным выявление скрытых коллизий, ранее доступных только узким специалистам. Дальнейшее развитие направления связано с расширением обучающих корпусов и практической интеграцией подобного инструментария в деятельность государственных органов и корпоративных структур.

Список литературы

1. Кульфединов Р.М. Сравнительный анализ современных методов автоматической обработки текстовых данных. Актуальные исследования. 2025. № 15-2 (250), с. 33–38.

2. Ермолатий Д.А., Быстров А.И. Повышение качества анализа и обработки правовой информации на основе применения методов машинного обучения. Современные наукоемкие технологии. 2023. № 4, с. 42–48.

УДК 004.056

Д.П. БАИАШВИЛИ, Д.Р. КАНТЕЕВА, Е.К. КОСТЫЛЕВА

МИРЭА – Российский технологический университет», Москва

СПОСОБ ПРОКТОРИНГА ЭЛЕКТРОННОЙ СИСТЕМЫ ТЕСТИРОВАНИЯ ЗНАНИЙ

В статье представлены способы прокторинга – системы контроля за проведением онлайн-тестирования, направленной на предотвращение мошенничества и объективную оценку знаний. Описаны основные виды прокторинга, их недостатки. Предложена новая мультимодальная система, сравнение с уже существующими.

Введение

Прокторинг верифицирует личность экзаменуемого, отслеживает отклонения от нормального поведения и рассчитывает вероятность нарушения. Система может в реальном времени использовать веб-камеру, микрофон и рабочий стол персонального компьютера [1]. Искусственный интеллект (ИИ) анализирует поведение экзаменуемого, фиксирует подозрительное поведение.

Имеющиеся решения

Автоматический прокторинг – это система, которая с помощью ИИ анализирует поведение сдающего. Синхронный прокторинг – это метод прокторинга, при котором обученный проктор наблюдает за экзаменуемым в режиме реального времени [2]. При асинхронном прокторинге, в свою очередь, запись экзамена с камеры и экрана сохраняется и просматривается проктором для выявления нарушений.

Таблица 1. Сравнение видов прокторинга

Виды прокторинга	Недостатки
Синхронный	Наиболее затратный из возможных методов внедрения системы онлайн-мониторинга за проведением экзаменов. Необходимость в работе сотрудников поддержки. В случае массовых экзаменов формат трудозатратен
Асинхронный	Проверка имеет выборочный характер, поскольку перепроверка видеозаписей будет только в том случае, если в заключении стоит низкая оценка доверия.
Автоматический	ИИ способен видеть только часть запрещенных действий, а также может ошибочно трактовать некоторые эпизоды

Существующие методики недостаточно эффективны против обхода систем прокторинга [3], необходимы более совершенные средства защиты.

Мультимодальная система

Для противодействия обходу систем прокторинга предлагается внедрение мультимодальной биометрической системы непрерывной аутентификации. Её ключевой принцип – постоянное подтверждение личности студента на протяжении всего экзамена.

Основу системы составляет одновременное использование двух каналов идентификации: распознавания лица на основе нейронных сетей и идентификации по голосу с помощью технологий, подобных x-векторной системе. Совместное использование данных с обоих каналов значительно повышает надёжность по сравнению с использованием каждого метода по отдельности, позволяя системе адекватно реагировать на неоднозначные ситуации.

Мультимодальная система прокторинга обеспечивает максимальную защиту от мошенничества благодаря двойной биометрической идентификации по лицу и голосу, что делает подмену личности практически невозможной. Непрерывный контроль исключает возможность подмены пользователя после начальной проверки. Автоматизация процесса минимизирует участие прокторов, которые привлекаются только для анализа подозрительных сессий. Высокая точность системы достигается совмещением двух каналов идентификации, что снижает количество ложных срабатываний - при проблемах с одним каналом система использует другой.

Заключение

Анализ показал ограниченную эффективность существующих систем прокторинга. Предлагаемая мультимодальная система на основе непрерывной биометрической аутентификации обеспечивает повышенную защиту от мошенничества за счет одновременной верификации по лицу и голосу. Этот подход позволяет сохранить надежность контроля при оптимизации затрат.

Список литературы

1. Холли Х. П. Хуанг и др "Handbook of Research on Remote Surveillance and Assessment in Educational Settings".

2. Startexam «Прокторинг 2025: как выбрать систему для безопасных онлайн-экзаменов» URL: https://www.startexam.ru/journal/likbez/proktoring-polnyy-gid-po-vyboru-sistemy-dlya-onlayn-ekzamenov-v-2025-godu/ (Дата обращения 04.10.2025).

3. Sber Developer «Как работает прокторинг на онлайн-экзаменах: контроль и проверка студентов в удобном формате» URL: https://developers.sber.ru/help/hr/proctoring (Дата обращения: 05.10.2025).

ИМЕННОЙ УКАЗАТЕЛЬ АВТОРОВ СТАТЕЙ

Подписано в печать 20.11.2025. Формат 60x84 1/16.
Печ. л. 14,0. Уч.-изд. 14,0. Тираж 160 экз.
Изд. №023-2. Заказ № 54

Национальный исследовательский ядерный университет «МИФИ»
Типография МИФИ
115409, Москва, Каширское ш. 31