На правах рукописи

РУСАНОВ МИХАИЛ АНДРЕЕВИЧ

ИНТЕЛЛЕКТУАЛЬНЫЕ МЕТОДЫ ОБНАРУЖЕНИЯ НЕСАНКЦИОНИРОВАННЫХ ВТОРЖЕНИЙ В ОПЕРАЦИОННЫЕ СИСТЕМЫ

Специальность: 2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

Москва – 2026

Работа выполнена на кафедре информационных и математических инноваций института информационных технологий аккредитованного образовательного частного учреждения высшего образования «Московский финансово-юридический университет МФЮА».

Научный руководитель:	**Лапина Мария Анатольевна,** кандидат физико-математических наук, доцент
Официальные оппоненты:	**Кознов Дмитрий Владимирович,** доктор технических наук, доцент, профессор кафедры системного программирования Федерального государственного бюджетного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет» **Котенко Игорь Витальевич,** доктор технических наук, профессор, главный научный сотрудник Федерального государственного бюджетного учреждения науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук»
Ведущая организация:	Федеральное государственное автономное образовательное учреждение высшего образования «Южный федеральный университет»

Защита состоится «17» сентября 2026 г. в 16:00 на заседании диссертационного совета 24.1.120.01 при Федеральном государственном бюджетном учреждении науки Институте системного программирования им. В.П. Иванникова Российской Академии Наук по адресу: 115035, г. Москва, ул. Садовническая, д. 41, ст. 2.

С диссертацией можно ознакомиться в библиотеке и на сайте Федерального государственного бюджетного учреждения науки Института системного программирования им. В.П. Иванникова Российской академии наук.

Автореферат разослан «___»_____20__ г.

Ученый секретарь
Диссертационного совета 24.1.120.01,
кандидат физико-математических наук Турдаков Д.Ю.

Общая характеристика работы

Актуальность работы. В условиях цифровой трансформации возрастает сложность современных информационно-вычислительных систем, характеризующихся распределенной архитектурой, высокой интенсивностью обмена данными и увеличением объемов потоковой информации. Современные программные комплексы функционируют в гетерогенных вычислительных средах, включающих облачные платформы, распределенные сервисы и ресурсно-ограниченные вычислительные узлы, что накладывает дополнительные ограничения на методы интеллектуальной обработки системных событий.

Развитие распределённых облачных технологий и программной инфраструктуры организации глобально распределённой обработки данных приводит к росту объёмов потоковых данных о работе программных компонентов. Анализ таких потоковых данных в режиме реального времени с высокой точностью представляет собой сложную научно-практическую задачу.

Традиционные подходы, основанные на правилах, сигнатурах и шаблонах, недостаточно адаптивны и не учитывают связи между событиями, поэтому всё чаще применяются методы машинного обучения. Используются модели семейства трансформеров, позволяющие учитывать контекстные взаимосвязи между элементами последовательностей системных событий и предназначенных для анализа текстовых данных. Их применение позволяет выявлять аномалии в последовательности системных событий и обеспечивать устойчивость и защищенность системы. Вместе с тем их практическое применение остаётся ресурсоёмким, что делает актуальной оптимизацию потокового вывода, направленную на повышение пропускной способности, сокращение времени обработки и параллельную обработку данных без изменения архитектуры модели.

Значительный научный вклад в рассматриваемую область внесли следующие исследователи: А.В. Федорченко, И.В. Котенко, А.В. Хорошилов, И. Г. Сидоркина, С. В. Михалищев, A. L. Buczak, E. Guven, R. Sommer и V. Paxson, T. Vyšniūnas, D. Čeponis, N. Goranin, A. Čenys, F. Wang, Q. Weng, M. Zhang, Y. Shao, Z. Alomari, A. Makanju, Z. Li, Zhang S., Zhao P., An Z. и др.

Объектом исследования являются интеллектуальные системы анализа потоковых событий операционных систем семейства Linux в распределенной вычислительной среде.

Предмет исследования – интеллектуальные методы, модели и программные средства формирования модулей обнаружения и классификации несанкционированных вторжений в операционные системы семейства Linux.

Целью является разработка интеллектуальных методов и программных средств анализа аутентификационных данных и потоковых системных событий операционных систем семейства Linux, направленных на повышение эффективности обнаружения и классификации несанкционированных вторжений и угроз компрометации учетных данных.

Для решения поставленной научной задачи была произведена ее декомпозиция на ряд **частных задач:**

1. Анализ методов обработки и классификации системных событий операционных систем и выявление ограничений их применения в задачах обнаружения и классификации несанкционированных вторжений.

2. Разработка методов предварительной обработки, представления и интеллектуального анализа потоковых событий операционных систем семейства Linux, включая модели обнаружения и классификации вредоносной активности, а также выявления потенциально компрометируемых аутентификационных данных.

3. Разработка методов и программных средств параллельной потоковой обработки событий операционных систем семейства Linux, обеспечивающих снижение времени обработки данных без уменьшения точности обнаружения и классификации.

Методология и методы исследования включают использование методов машинного обучения, обработки естественного языка, анализа данных, теория вероятностей, математическая статистика, а также методы математического моделирования.

Соответствие паспорту научной специальности.

Область исследования соответствует паспорту специальности 2.3.5 – «Математическое и программное обеспечение вычислительных систем,

комплексов и компьютерных сетей» по следующим пунктам:

п. 4. Интеллектуальные системы машинного обучения, управления базами данных и знаний, инструментальные средства разработки цифровых продуктов.

п. 8. Модели и методы создания программ и программных систем для параллельной и распределенной обработки данных, языки и инструментальные средства параллельного программирования.

п. 9. Модели, методы, алгоритмы, облачные технологии и программная инфраструктура организации глобально распределенной обработки данных.

Научная новизна диссертационной работы:

1. Разработан интеллектуальный метод оценки вероятности компрометации аутентификационных данных операционных систем семейства Linux, отличающийся от существующих подходов использованием статистических характеристик случайности в сочетании с моделями машинного обучения.

2. Разработан интеллектуальный метод обработки потоковых событий операционных систем семейства Linux для обнаружения вредоносной активности, отличающийся от существующих методов применением контекстного представления событий на основе моделей трансформеров с предварительной нормализацией и семантическим упрощением событий, что позволяет повысить точность классификации атак и уменьшить время работы моделей.

3. Разработан метод параллельной потоковой обработки событий операционных систем семейства Linux, отличающийся от существующих подходов использованием адаптивного формирования вычислительных пакетов, динамического выравнивания длины входных последовательностей и группировки событий по вычислительной сложности, что обеспечивает уменьшение времени обработки и повышение пропускной способности интеллектуальных моделей анализа системных событий без изменения архитектуры модели и снижения точности обнаружения и классификации.

Теоретическая значимость результатов диссертационного исследования заключается в следующем**:**

1) Разработанный метод выявления потенциально компрометируемых аутентификационных данных позволяет уточнить оценки вероятности их

компрометации за счет интеллектуального анализа косвенных признаков.

2) Разработанный интеллектуальный метод анализа событий операционных систем семейства Linux на основе языковых моделей вносит вклад в развитие теоретических основ для обнаружения вредоносной активности и расширяет представления о их применимости для анализа сложных и модифицированных атак.

3) Разработанный метод параллельной потоковой обработки событий операционных систем семейства Linux позволяет уменьшить время их обработки за счет адаптивного формирования вычислительных пакетов, динамического выравнивания длины входных последовательностей и группировки событий по вычислительной сложности.

Практическая значимость результатов диссертационного исследования заключается в следующем**:**

1) Разработанный метод выявления потенциально компрометируемых аутентификационных данных позволяет уточнить вероятность их компрометации, что обеспечивает возможность его применения в подсистемах контроля защищенности, аудита безопасности и управления доступом.

2) Разработанный интеллектуальный метод анализа событий операционных систем семейства Linux на основе языковых моделей позволяет уменьшить время их обработки в среднем в 5 раз, что обеспечивает возможность его применения при создании программного обеспечения для центров мониторинга информационной безопасности.

3) Разработанный метод параллельной потоковой обработки событий операционных систем семейства Linux позволяет уменьшить время обработки в среднем в 2 раза и повысить пропускную способность интеллектуальных подсистем мониторинга информационной безопасности на 123% без снижения точности классификации, что обеспечивает возможность его применения при создании программного обеспечения для центров мониторинга информационной безопасности.

На защиту выносятся следующие положения:

1. Интеллектуальный метод оценки вероятности компрометации

аутентификационных данных операционных систем семейства Linux, основанный на вычислении статистических характеристик случайности в сочетании с моделями машинного обучения.

2. Интеллектуальный метод обработки потоковых событий операционных систем семейства Linux для обнаружения вредоносной активности, основанный на применении контекстного представления событий и моделей трансформеров с предварительной нормализацией и семантическим упрощением событий.

3. Метод параллельной потоковой обработки событий операционных систем семейства Linux с использованием адаптивного формирования вычислительных пакетов, динамического выравнивания длины входных последовательностей и группировки событий по вычислительной сложности.

Достоверность и обоснованность полученных в диссертационной работе теоретических результатов и формулируемых на их основе выводов подтверждается корректным и обоснованным применением классических методов исследования, строгими математическими доказательствами и результатами анализа эффективности реализации разработанных моделей и методов с использованием языка программирования Python. Результаты теоретического анализа согласуются с результатами проведенных экспериментов.

Публикации. По теме диссертации опубликовано 6 печатных работ, в том числе в изданиях, рекомендованных ВАК [1,2,5] и Web of Science и Scopus [3,4], а также получено 2 свидетельства о государственной регистрации программ для ЭВМ [7,8].

Личный вклад автора. Диссертационная работа представляет собой исследование автора, объединенное тематикой и методами исследования. Все выносимые на защиту результаты получены лично автором. Из совместных работ в диссертацию включены только те результаты, которые принадлежат непосредственно автору. В опубликованных совместных работах постановка и решение задач осуществлялись совместными усилиями соавторов при непосредственном участии соискателя. В работе [1] автором разработан интеллектуальный метод оценки вероятности компрометации аутентификационных данных, основанный на анализе статистических

характеристик случайности паролей и применении моделей машинного обучения для выявления потенциально компрометируемых паролей. В работах [2-6] автором разработан интеллектуальный метод обработки потоковых системных событий операционных систем семейства Linux, обеспечивающий обнаружение вредоносной активности и классификацию атак по тактикам MITRE ATT&CK на основе контекстного представления событий и трансформерных моделей, а также повышение вычислительной эффективности их потоковой обработки за счёт адаптивного формирования вычислительных пакетов и динамического выравнивания входных последовательностей. В государственных регистрациях программ для ЭВМ [7,8] автором выполнена разработка программных комплексов обнаружения и классификации вредоносной активности в системных событиях и приложениях на базе нейронной сети.

Апробация работы. Результаты данной работы докладывались на конференциях:

1. XIII Всероссийская научно-техническая конференция с международным участием «Актуальные проблемы информационной безопасности», 2025, Самара;

2. Международная весенняя конференция молодых учёных в области программной инженерии (SYRCoSE Software Engineering Colloquium), 2025, Пятигорск;

3. Всероссийская конференция с международным участием "Радиоэлектронные устройства и системы для инфокоммуникационных технологий (РЭУС-ИТ), 2026, Москва;

4. XVII Международная научно-практическая конференция имени Олега Борисовича Макаревича «Современные методы, средства и технологии защиты информации – 2026», 2026, Таганрог.

5. Международная конференция «AITHD-2025: International Conference on Artificial intelligence as a Technologies of Human Development», 2025, Ставрополь.

Внедрение результатов. Результаты, полученные в рамках данной работы, были внедрены в подсистемы мониторинга событий информационной безопасности в АО «Газпромбанк» (акт о внедрении 302-3/12/26 от 05.05.2026).

Структура диссертации. Полный объем диссертации составляет 130

страниц, включая 30 рисунков и 20 таблиц. Список литературы содержит 108 наименований.

Основное содержание работы

Во введении обоснована актуальность темы диссертации, сформулированы цель и задачи работы, выбраны объект и предмет исследования, показаны научная новизна, практическая и теоретическая ценность полученных результатов, приведены основные положения, выносимые на защиту.

В первой главе выполнен анализ существующих методов обнаружения и классификации вредоносной активности по данным системных событий операционных систем семейства Linux, методов выявления потенциально компрометируемых аутентификационных данных, а также подходов к повышению эффективности интеллектуальной обработки потоковых системных событий.

Показано, что традиционные системы обнаружения вторжений, основанные на сигнатурных и эвристических правилах, обладают ограниченной способностью к выявлению новых и модифицированных атак. Установлено, что современные методы машинного обучения и обработки естественного языка позволяют повысить качество анализа системных событий операционных систем семейства Linux за счет использования контекстной информации и формирования семантических представлений событий. Вместе с тем существующие решения ориентированы на применение ресурсозатных моделей, что затрудняет их использование в условиях обработки больших потоков системных событий и в ресурсно-ограниченных вычислительных средах.

Установлено, что большинство применяемых на практике решений основано на использовании фиксированных требований парольной политики, словарных проверок и сравнении с базами ранее скомпрометированных паролей. Проведенное исследование показало, что такие подходы не позволяют учитывать структурные особенности паролей и степень их статистической случайности. Определены характеристики, позволяющие количественно описывать свойства случайности паролей, включая количество задействованных алфавитов, относительное количество уникальных символов, частоту смены алфавитов, долю специальных

символов и числовых символов, а также особенности расположения символов на клавиатуре.

Во второй главе разработан интеллектуальный метод оценки вероятности компрометации аутентификационных данных в ОС Linux, основанный на статистических характеристиках случайности паролей и моделях машинного обучения. В качестве признаков использованы длина пароля (Len), количество алфавитов (AC), доля уникальных символов (Uniq), специальных символов (S) и цифр (N), среднее расстояние между символами на клавиатуре (Dist), а также частота смены алфавитов (Freq). Эти характеристики отражают разнообразие символов, повторяемость, структуру пароля и пространственные закономерности пользовательского ввода.

Для обучения и оценки признаков сформирована выборка из двух классов: потенциально компрометируемых и надёжных паролей. Первый класс получен из базы rockyou.txt после фильтрации по длине, числу алфавитов и допустимым символам и составил около 385 тыс. записей. Второй класс сформирован с помощью сервисов Passwordsgenerator, Passwords-Generator и Kaspersky; из около 4 млн сгенерированных паролей после фильтрации и выравнивания классов также отобрано 385 тыс. записей. Итоговая выборка составила около 770 тыс. паролей.

Для исключения избыточности признакового пространства выполнен корреляционный анализ. Установлено, что информационная энтропия (E), доли символов верхнего (U) и нижнего (L) регистров имеют высокую корреляцию с другими характеристиками, поэтому не включены в итоговый набор признаков. Коррелирующие признаки были исключены из итогового признакового пространства (рис. 1).

Анализ показал статистически значимые различия между признаками надёжных и потенциально компрометируемых паролей (рис. 2–4), однако пересечение их распределений ограничивает эффективность решений на основе отдельных пороговых значений.

Оценка вероятности компрометации аутентификационных данных должна выполняться не по отдельным характеристикам, а на основе их совместного анализа. При этом взаимосвязи между признаками имеют нелинейный характер и

зависят от их совокупного влияния на структуру пароля.

Рисунок 1 – Матрица корреляции признаков паролей

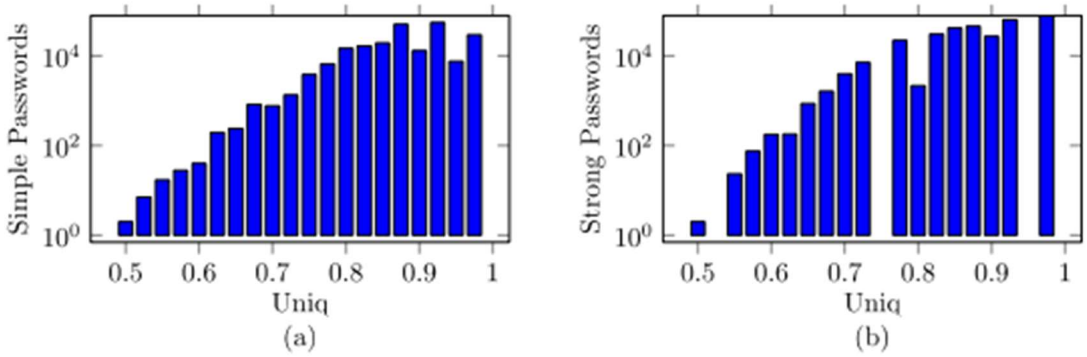


Рисунок 2 – Распределение паролей по количеству уникальных символов: (a) простые пароли; (b) надежные пароли

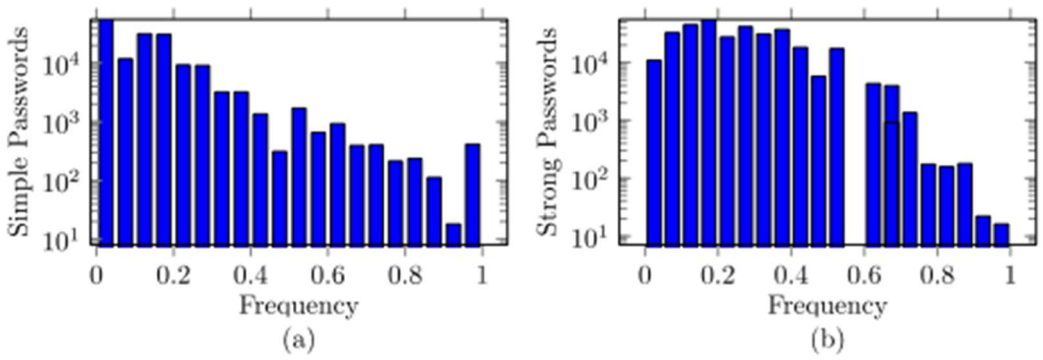


Рисунок 3 – Распределение паролей по частоте изменения алфавита: (a) простые пароли; (b) надежные

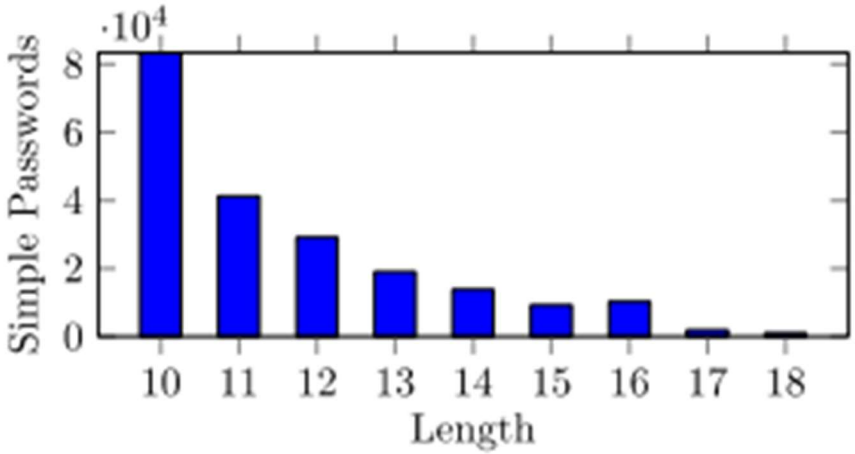


Рисунок 4 – Распределение простых паролей по длине

Для реализации разработанного метода предложено использование моделей машинного обучения, позволяющих выявлять зависимости между статистическими характеристиками случайности и оценивать вероятность компрометации аутентификационных данных по совокупности признаков.

Разработан **интеллектуальный метод обнаружения вредоносной активности по потоковым системным событиям ОС Linux**, основанный на контекстном представлении событий, моделях трансформеров, предварительной нормализации и семантическом упрощении данных (рис. 5). Метод реализуется двумя компонентами: SIEM-системой, выполняющей сбор, агрегацию и корреляцию событий от узлов инфраструктуры, и программным комплексом, осуществляющим их интеллектуальный анализ и классификацию. Для формирования контекста события объединяются в последовательности в пределах заданного временного окна и рассматриваются как единый объект анализа, что позволяет учитывать связи между событиями без их изолированной обработки.

На первом этапе выполняется нормализация системных событий Linux, направленная на снижение вариативности входных данных и удаление параметров, не влияющих на семантику действий. Она включает приведение текста к нижнему регистру, удаление лишних пробелов, Normalization Form KC (NFKC)-нормализацию, исключение нерелевантных блоков и параметров, а также замену идентификаторов процессов, IP-адресов, доменов, хостов и хешей статичными токенами. Пути к файлам кодируются метками `<path_N>` с сохранением словаря соответствий.

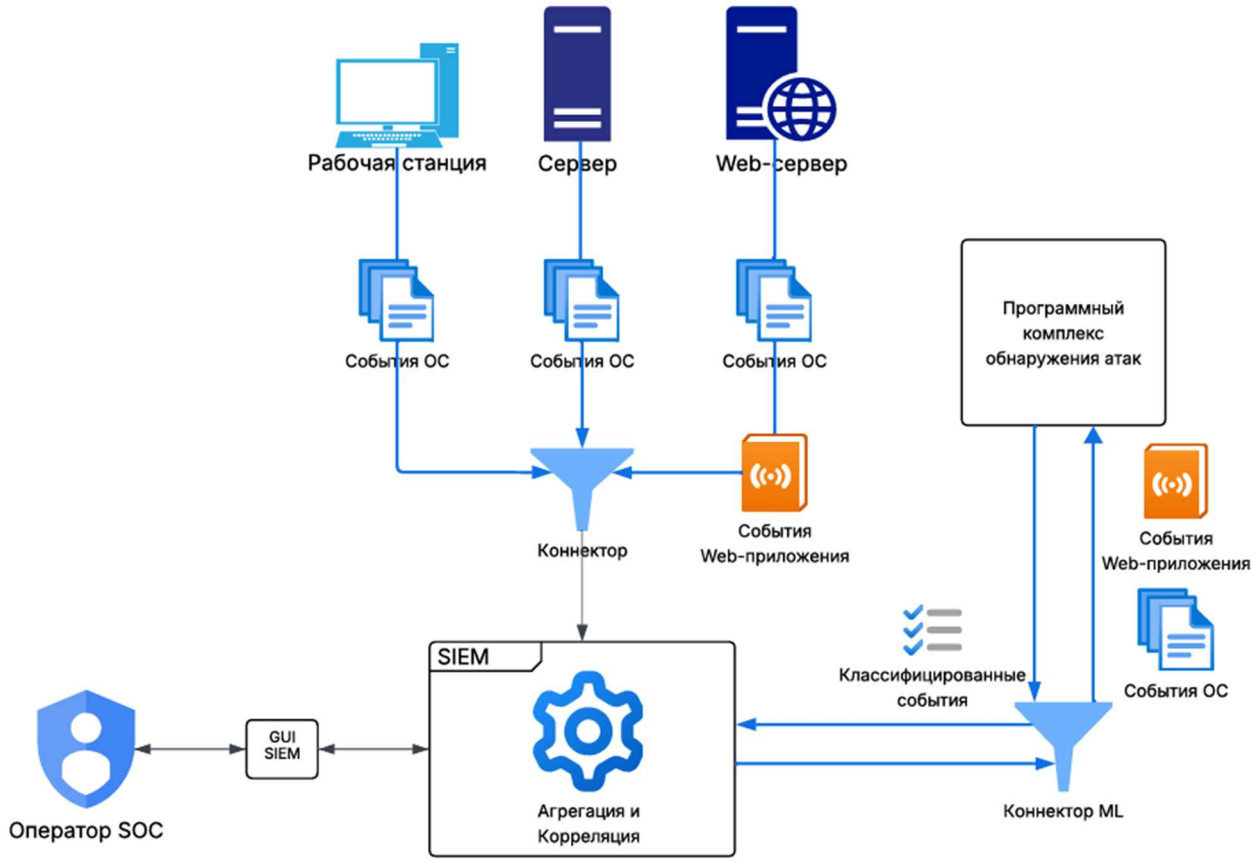


Рисунок 5 – Схема метода обнаружения вредоносной активности по потоковым системным событиям ОС Linux

После обработки события объединяются в последовательности в пределах временного окна и передаются в трансформерную модель бинарной классификации для выявления вредоносного контекста. Для снижения вычислительной сложности используется дистилляция знаний. При обнаружении атаки последовательность дополнительно анализируется второй трансформерной моделью, выполняющей многометочную классификацию по тактикам MITRE ATT&CK. Метод объединяет нормализацию, семантическое упрощение, контекстное представление событий, обнаружение вредоносной активности и последующую классификацию атак.

Разработан **метод параллельной потоковой обработки системных событий ОС Linux**, основанный на формировании независимых вычислительных пакетов, динамическом выравнивании последовательностей и группировке событий по сложности (рис. 6). Распределение пакетов между параллельными вычислительными потоками позволяет одновременно обрабатывать несколько

групп событий и повышает эффективность использования вычислительных ресурсов.

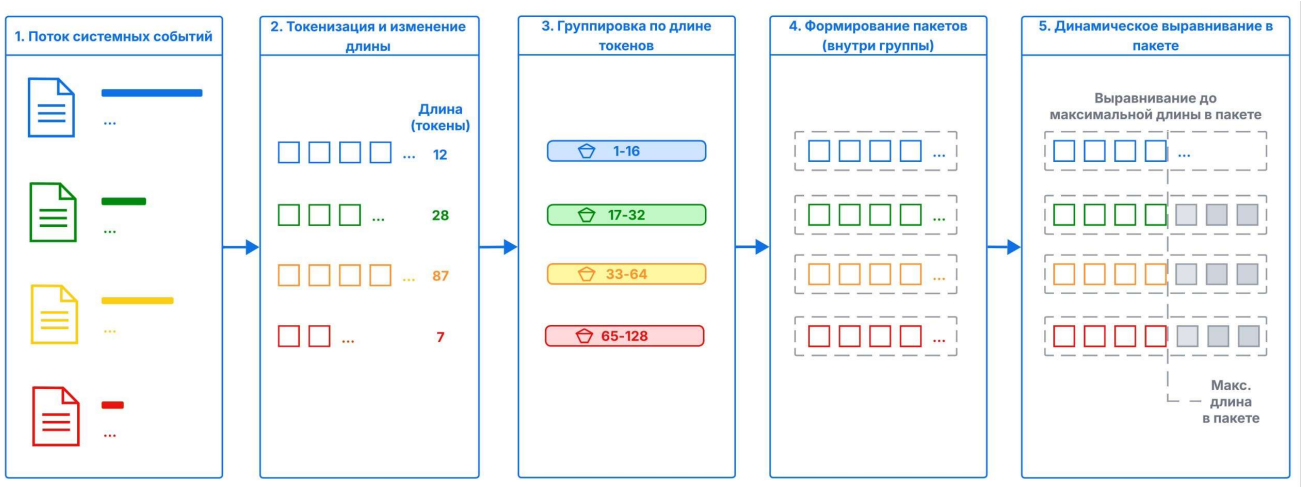


Рисунок 6 – Схема метода параллельной потоковой обработки системных событий

После предварительной обработки и токенизации события преобразуются в последовательности токенов и объединяются в вычислительные пакеты, что снижает накладные расходы на подготовку данных и запуск вычислений. Выравнивание выполняется по максимальной длине последовательности внутри каждого пакета, что уменьшает объём вычислений для модели трансформер. Дополнительно последовательности группируются по вычислительной сложности перед распределением между потоками, что снижает дисбаланс нагрузки и повышает эффективность параллельной обработки.

В третьей главе для практической проверки разработанного метода оценки вероятности компрометации аутентификационных данных реализован программный комплекс, интегрируемый в ОС Linux и предназначенный для автоматического контроля создаваемых пользователями паролей.

Интеграция выполнена через подсистему Pluggable Authentication Modules (PAM), позволяющую подключать дополнительные механизмы проверки паролей без изменения исходного кода ОС и стандартных утилит. Вызов модуля анализа осуществляется из стека `password` с помощью `pam_exec`: при смене пароля введённое значение передаётся через стандартный поток ввода в модуль, где обрабатывается в соответствии с разработанным методом. Схема интеграции с PAM представлена на рис. 7.

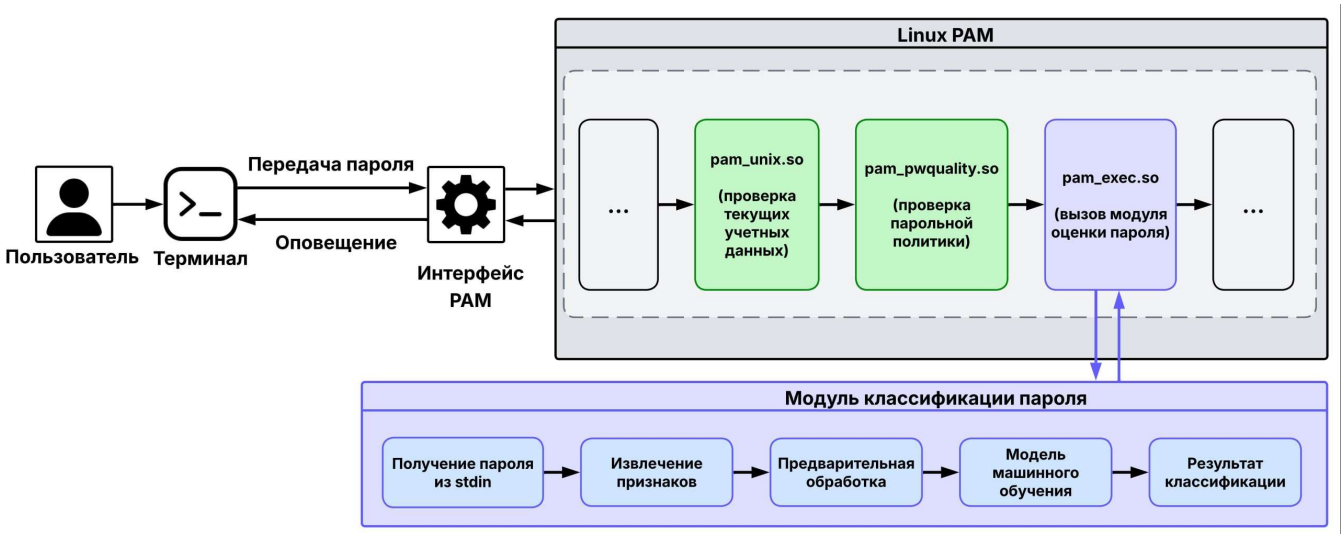


Рисунок 7 – Интеграция программного комплекса с модулем PAM

Разработанный модуль встраивается в стандартный стек обработки паролей Linux PAM и выполняется совместно с другими модулями, участвующими в процедуре изменения аутентификационных данных. После прохождения встроенных проверок учетных данных и требований парольной политики управление передается модулю pam_exec, который инициирует вызов разработанного модуля классификации паролей.

Полученный от пользователя пароль передается в модуль классификации через стандартный поток ввода, после чего выполняются его предварительная обработка, извлечение признаков и классификация средствами модели машинного обучения. Результат анализа возвращается обратно в подсистему PAM в виде кода завершения, на основании которого принимается решение о продолжении или прекращении процедуры изменения пароля и формируется соответствующее уведомление пользователю.

После получения пароля выполняется проверка его соответствия минимальным требованиям парольной политики, расчет статистических характеристик случайности и формирование признакового описания (алг. 1).

На следующем этапе осуществляется расчет признаков, используемых для оценки вероятности компрометации аутентификационных данных. Для анализируемого пароля вычисляются длина пароля, количество задействованных алфавитов, относительное количество уникальных символов, доля специальных символов, доля числовых символов, среднее расстояние между символами по их

расположению на клавиатуре и частота смены используемых алфавитов (алг. 2).

Алгоритм 1. Предварительная обработка паролей

Вход: p
Выход: p_{valid}, R, Len, AC
$MinL = 8$
$MaxL = 18$
$Len \leftarrow size(p)$
Если $Len < MinL$ **или** $Len > MaxL$
    $R \leftarrow PAM\_REJECT$
    Возвратить R
Конец условия
$AC \leftarrow calculate\_AC(p)$
Если $AC < 3$
    $R \leftarrow PAM\_REJECT$
    Возвратить R
Конец условия
$p_{valid} \leftarrow p$
$R \leftarrow PAM\_SUCCESS$
Возвратить p_{valid}, R, Len, AC

Алгоритм 2. Формирование признакового описания

Вход: p, Len, AC
Выход: F
$Uniq \leftarrow calculate\_Uniq(p)$
$N \leftarrow calculate\_N(p)$
$S \leftarrow calculate\_S(p)$
$Dist \leftarrow calculate\_Dist(p)$
$Freq \leftarrow calculate\_Freq(p)$
$F \leftarrow \{Len, AC, Uniq, N, S, Dist, Freq\}$
Возвратить F

Проведенные эксперименты показали, что наилучшее качество классификации обеспечивает алгоритм Decision Tree, который был выбран в качестве основы вычислительного модуля оценки вероятности компрометации аутентификационных данных. Результаты сравнения метрик качества моделей представлена в таблице 1.

Таблица 1 – Классификации паролей

Алгоритм	Тип пароля	TP	FP	TN	FN	Точность	Cohen's Kappa	Время, с
XGBoost	0	0.97	0.03	0.98	0.02	0.97	0.94	6.56
	1	0.98	0.02	0.97	0.03			
Random forest	0	0.97	0.03	0.98	0.02	0.97	0.94	13.70
	1	0.98	0.02	0.97	0.03			
Naive Bayes	0	0.32	0.61	1.00	0.00	0.69	0.39	0.58
	1	1.00	0.00	0.32	0.61			
Decision Tree	0	0.97	0.03	0.98	0.02	0.98	0.96	14.03
	1	0.98	0.02	0.97	0.03			

Качество разработанного классификатора оценивалось с использованием матрицы ошибок (таблица 2) и стандартных метрик бинарной классификации. Полученная модель обеспечила точность классификации 97,40%, коэффициент ошибок 2,60% и значение коэффициента Cohen's Kappa 0,957, что свидетельствует о высокой эффективности предложенного подхода к выявлению потенциально компрометируемых аутентификационных данных. Дополнительная оценка качества выполнялась на основе ROC-анализа рис. 8.

Таблица 2 – Матрица ошибок модели классификации паролей

Истинный класс	Предсказан 0	Предсказан 1	Точность, %
0	107839	2883	97,40
1	1611	98285	98,39

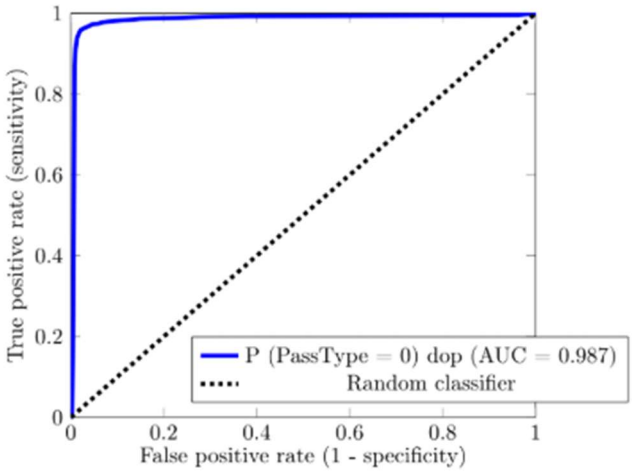


Рисунок 8 – ROC-кривая классификатора паролей

Площадь под ROC-кривой составила 0,987, что подтверждает высокую разделяющую способность классификатора и эффективность использования

статистических характеристик случайности совместно с машинным обучением для выявления потенциально компрометируемых паролей, включая случаи формального соответствия парольной политике.

Второй метод реализован как программный комплекс обнаружения и классификации вредоносной активности, включающий предварительную обработку событий, формирование контекстного представления, бинарное обнаружение и классификацию атак по тактикам MITRE ATT&CK. Для обучения использован Linux-APT-Dataset-2024, собранный в тестовой среде и агрегированный Wazuh. Он содержит 114 851 запись, из них 91 804 обычных и 23 046 вредоносных события, охватывающих повышение привилегий Linux, эксплуатацию уязвимостей, эмуляцию кейлоггера и кампании APT41, APT28, APT29 и Turla.

После нормализации и семантического упрощения события токенизировались с помощью WordPiece и передавались в трансформерные модели. Для бинарной классификации исследованы BERT-Base, DistilBERT, MobileBERT и TinyBERT; их сравнение выполнено по качеству классификации и времени обработки, результаты приведены в таблице 3.

Таблица 3 – Сравнение вариантов BERT для бинарной классификации вредоносной активности

Модель	Время, с	Precision		Recall		F1-score		Точность
		normal	malicious	normal	malicious	normal	malicious	
BERT-Base	486.31	0.99	1.00	0.99	0.97	1.00	0.98	0.99
DistilBERT	252.96	0.99	1.00	1.00	0.97	1.00	0.98	0.99
MobileBERT	691.80	0.99	1.00	1.00	0.97	1.00	0.98	0.99
TinyBERT-4L	127.75	0.87	1.00	1.00	0.37	0.93	0.54	0.88

Эксперименты показали, что BERT-Base обеспечивает наилучшее качество классификации системных событий, но требует значительных вычислительных ресурсов, тогда как компактные модели работают быстрее, уступая по точности. Для снижения вычислительной сложности без существенной потери качества применена дистилляция знаний: BERT-Base использовалась как модель-учитель, а TinyBERT — как модель-студент. В результате студент перенимал выявленные

учителем закономерности, что позволило получить более компактную и эффективную модель. Схема подготовки дистиллированной модели представлена на рисунке 9.

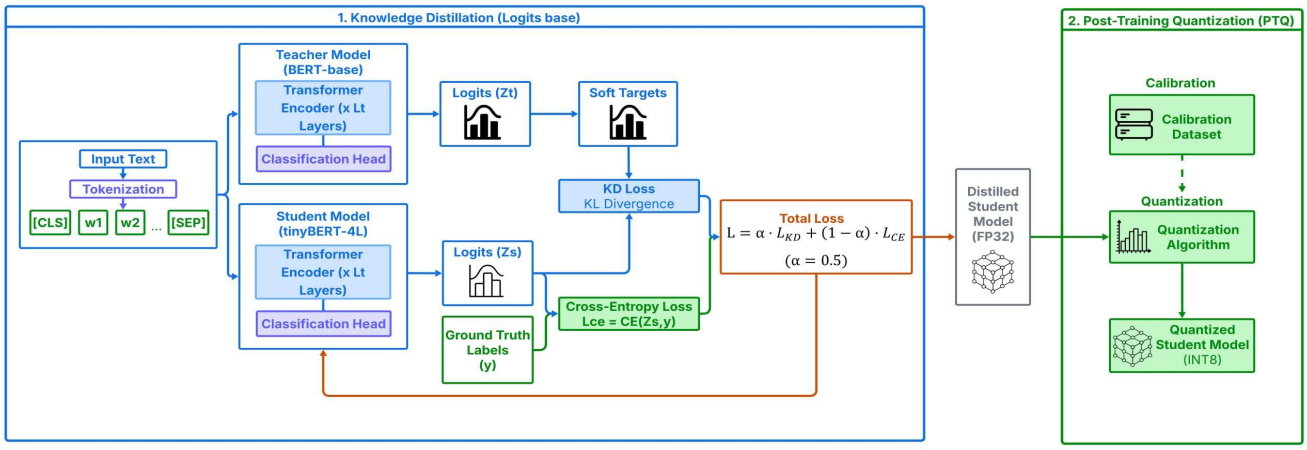


Рисунок 9 – Схема подготовки модели

Обучение модели выполнялось с использованием оптимизатора AdamW и функции потерь CrossEntropyLoss. Полученные показатели качества дистиллированной модели представлены в таблице 4.

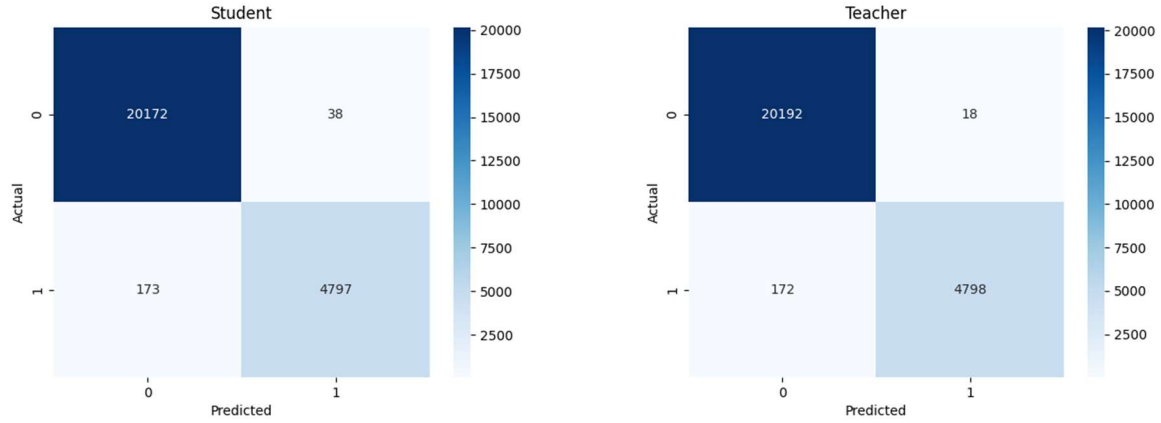
Таблица 4 – Метрики качества дистиллированной модели TinyBERT

Время, с	Precision		Recall		F1-score		Точность	PR AUC
	normal	malicious	normal	malicious	normal	malicious		
93.97	0.986	0.997	0.999	0.944	0.993	0.970	0.988	0.990

Анализ полученных результатов показал, что дистиллированная модель обеспечивает качество классификации, близкое к качеству исходной модели BERT, при существенно меньших вычислительных затратах. Матрицы ошибок модели-учителя и модели-студента представлены на рисунке 10.

Для дополнительного уменьшения вычислительных затрат к дистиллированной модели применялась процедура 8-битной пост-тренировочной квантизации, которая затрагивала линейные слои модели и не изменяла структуру трансформерной архитектуры. Полученные результаты представлены в таблице 5.

Применение квантизации позволило дополнительно уменьшить размер модели и сократить время обработки данных при сохранении показателей качества классификации практически на прежнем уровне.



(a) (b)

Рисунок 10 – Сравнение дистиллированных моделей TinyBERT: (a) матрица ошибок для модели-студента; (b) матрица ошибок для модели-учителя [источник: составлено автором]

Таблица 5 – Метрики качества квантованной модели TinyBERT

Время, с	Precision		Recall		F1-score		Точность
	normal	malicious	normal	malicious	normal	malicious	
52.45	0.986	0.997	0.999	0.944	0.993	0.970	0.988

Итоговое ускорение обработки обусловлено переходом от полноразмерных трансформерных архитектур к компактной дистиллированной модели. Таким образом, сравнение выполняется относительно решений, основанных на использовании базовых моделей семейства BERT для анализа системных событий, для которых характерны более высокие вычислительные затраты на этапе вывода.

Для классификации обнаруженной вредоносной активности по тактикам матрицы MITRE ATT&CK была подготовлена отдельная трансформерная модель, использующая аналогичный подход к представлению системных событий. Результаты многометочной классификации по тактикам атак представлены в таблице 6.

Многометочная классификация показала высокое качество определения большинства тактик MITRE ATT&CK, за исключением Collection, Credential Access и Command and Control. Снижение метрик связано с дисбалансом классов: указанные тактики составляют менее 0,2 % выборки, что ограничивает формирование устойчивых обучающих закономерностей. Вместе с тем бинарная классификация в большинстве случаев корректно относит эти последовательности

к вредоносным с точностью около 98 %. Полученные результаты подтверждают эффективность программного комплекса для обнаружения вредоносной активности и её последующей классификации по тактикам MITRE ATT&CK в потоковых системных событиях Linux. Третий метод направлен на повышение эффективности комплекса за счёт параллельной потоковой обработки.

Таблица 6 – Метрики качества многометочной классификации по тактикам атак

Тактика	Precision	Recall	F1-Score	AUC
Defense Evasion	0.9988	0.9902	0.9945	0.9993
Privilege Escalation	0.9974	0.9932	0.9953	0.9997
Initial Access	0.9120	0.9977	0.9529	0.9267
Discovery	0.9958	0.9977	0.9967	0.9994
Impact	1.0000	1.0000	1.0000	1.0000
Execution	0.9878	0.9099	0.9476	0.9852
Persistence	1.0000	1.0000	1.0000	1.0000
Reconnaissance	0.9685	0.9690	0.9647	0.9924
Lateral Movement	0.9733	0.9783	0.9758	0.9934
Credential Access	0.7500	0.4286	0.5455	0.9556
Collection	0.0000	0.0000	0.0000	0.0000
Command and Control	0.0000	0.0000	0.0000	0.0000

Алгоритм программной реализации (алг. 3) включает загрузку событий, токенизацию, оценку длины последовательностей, распределение по вычислительным группам, буферизацию, формирование пакетов и выполнение вывода модели. Используется CPU-ориентированная конфигурация с управлением числом потоков через `torch.set_num_threads()`, что позволяет задействовать параллельное выполнение операций линейной алгебры.

В отличие от пакетов фиксированного размера, разработанный механизм учитывает характеристики входного потока и динамически изменяет состав пакетов, обеспечивая более равномерную загрузку потоков и сокращение времени ожидания обработки последовательностей (алг. 4).

Алгоритм 3. Подготовка вычислительного пакета к выводу модели

Вход: $B = \{s_1, s_2, \dots, s_n\}$
Выход: Y
1: $maxLen \leftarrow \max\_length(B)$
2: **Цикл для каждого** $s_i \in B$ **выполнить**
3: $s_i \leftarrow pad(s_i, maxLen)$
4: **Конец цикла**
5: $X \leftarrow create\_tensor(B)$
6: $logits \leftarrow inference(B)$
7: $Y \leftarrow predict(logits)$
8: **Возвратить** Y

Динамическое выравнивание выполняется по максимальной длине последовательности внутри каждого пакета, а не по всему набору данных, что уменьшает количество служебных элементов заполнения. Дополнительно последовательности группируются по длине и предполагаемой вычислительной сложности перед формированием пакетов.

Алгоритм 4. Обработка вычислительного пакета

Вход: $A = \{s_1, s_2, \dots, s_m\}, BATCH\_SIZE, timeout$
Выход: $B = \{s_1, s_2, \dots, s_n\}$
1: **Цикл для каждого** $s_i \in A$ **выполнить**
2: $Len_i \leftarrow size(s_i)$
3: $G \leftarrow determine\_group(s_i, Len_i)$
4: $Buffer_G \leftarrow Buffer_G \cup \{s_i\}$
5: **Если** $size(Buffer_G) \geq BATCH\_SIZE$
6: $B \leftarrow B \cup Buffer_G$
7: Возвратить B
8: **Иначе если** $elapsed\_time(Buffer_G) \geq timeout$
9: $B \leftarrow B \cup Buffer_G$
10: Возвратить B
11: **Иначе**
12: *Ожидать поступления новых событий*
13: **Конец условия**
14: **Конец цикла**
15: **Возвратить** B

Для оценки эффективности метода проведено экспериментальное исследование производительности при обработке потоковых системных событий. Основными метриками выступали время обработки и пропускная способность системы; результаты представлены на рис. 11.

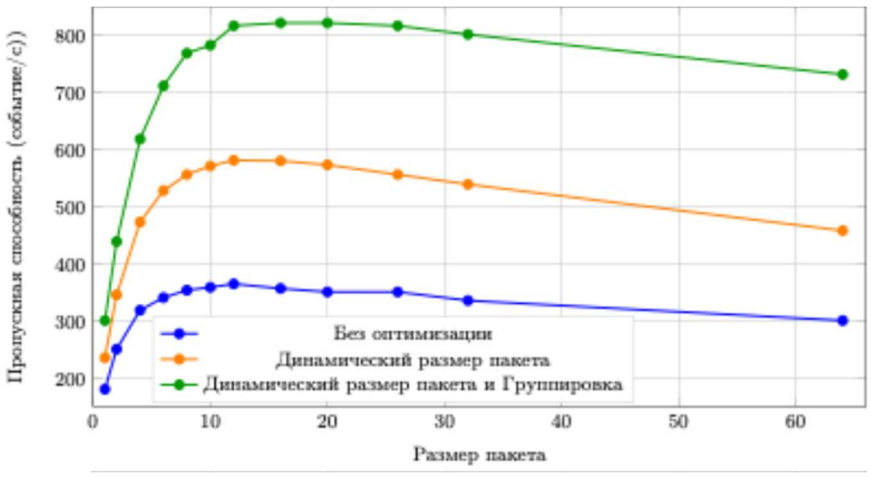


Рисунок 11 – Зависимость пропускной способности от размера пакета

Пакетная обработка снижает накладные расходы запуска модели, динамическое выравнивание уменьшает объём вычислений, а группировка событий по длине сокращает неоднородность пакетов и повышает производительность. Совместное применение механизмов увеличивает пропускную способность в 2,23 раза относительно базового варианта и в 1,4 раза относительно стандартной пакетной обработки. Оптимальный размер пакета составил 12 токенизированных событий: после этого значения производительность стабилизируется и постепенно снижается. В традиционной схеме пакеты формируются без учёта длины последовательностей, поэтому объём вычислений определяется максимальной длиной события в пакете, что ограничивает эффективность использования ресурсов.

Основные результаты и выводы по работе

В рамках проведенного диссертационного исследования разработаны интеллектуальные методы, модели и программные средства формирования модулей обнаружения и классификации несанкционированных вторжений в операционные системы семейства Linux. Разработанные интеллектуальные методы позволяют не только расширить класс задач, решаемых с использованием языковых моделей, но и уменьшить время обработки данных без снижения уровня точности. Основными результатами диссертационного исследования являются:

1. Разработан метод выявления потенциально компрометируемых аутентификационных данных, позволяющий уточнить вероятность их компрометации, что обеспечивает возможность его применения в подсистемах

контроля защищенности, аудита безопасности и управления доступом.

2. Разработан интеллектуальный метод анализа событий операционных систем семейства Linux на основе языковых моделей, позволяющий уменьшить время их обработки в среднем в 5 раз, что обеспечивает возможность его применения при создании программного обеспечения для центров мониторинга информационной безопасности.

3. Разработан метод параллельной потоковой обработки событий операционных систем семейства Linux, позволяющий уменьшить время обработки в среднем в 2 раза и повысить пропускную способность интеллектуальных подсистем мониторинга информационной безопасности на 123% без снижения точности классификации, что обеспечивает возможность его применения при создании программного обеспечения для центров мониторинга информационной безопасности.

Публикации по теме диссертации

1. Rusanov M. A. Research of machine learning methods for detecting network attacks / M. A. Lapina, N. V. Podruchny, M. A. Rusanov, M. G. Babenko // Труды Института системного программирования РАН. – 2025. – Т. 37. – №. 4-2. – С. 147-174. (0,42/1,69 п.л.)
2. Rusanov M. A. Detection of SQL injection attacks through the network logs using machine learning methods / M. A. Lapina, N. R. Kapshuk, M. A. Rusanov, E. F. Timofeeva // Труды института системного программирования РАН. – 2025. – Т. 37. – №. 5. – С. 81-92. (0,19/0,75 п.л.)
3. Rusanov M. Identification of exploited unreliable account passwords in the information infrastructure using machine learning methods / M. Rusanov, M. Babenko, M. Lapina, M. Sajid // Big Data and Cognitive Computing. – 2024. – Т. 8. – №. 11. – С. 159. (0,31/1,25 п.л.)
4. Rusanov M. Optimization of Machine Learning Algorithms with Distillation and Quantization for Early Detection of Attacks in Resource-Constrained Systems / M. Rusanov, M. Babenko, M. Lapina // Big Data and Cognitive Computing. – 2025. – Т. 9. – №. 12. – С. 303. (0,41/1,25 п.л.)

5. Русанов М.А. Криптографическая система на основе хаотической системы Лоренца / М.А. Русанов, С.В. Данилкин, П.И. Карасев, А.А.Л. Алмали // Приборы и системы. Управление, контроль, диагностика. 2023, – №. 3. – С. 48-53. (0,1/0,38 п.л.)
6. Русанов М.А. Эффективные модели машинного обучения для раннего выявления угроз / М.А. Русанов, М.Г. Бабенко // Актуальные проблемы информационной безопасности: материалы XIII Всероссийской научно-практической конференции 18 декабря 2025 г. 2026, – С. 126-131. (0,19/0,375 п.л.)
7. Свидетельство о государственной регистрации программы для ЭВМ №2025696118 Российская Федерация. Программа обнаружения вредоносной активности в системных событиях и приложениях на базе нейронной сети: № 2025694833: заявл. 01.12.2025: опубл. 16.12.2025 / М. А. Русанов, М. А. Лапина, М. Г. Бабенко; заявитель Федеральное государственное автономное образовательное учреждение высшего образования «Северо-Кавказский федеральный университет».
8. Свидетельство о государственной регистрации программы для ЭВМ №2025695537 Российская Федерация. Программа идентификации тактик вредоносной активности в системных событиях на основе нейронной сети: № 2025694747 : заявл. 01.12.2025: опубл. 11.12.2025 / М. А. Русанов, М. А. Лапина, М. Г. Бабенко; заявитель Федеральное государственное автономное образовательное учреждение высшего образования «Северо-Кавказский федеральный университет».