На правах рукописи



ПЛАВАН Алексей Игоревич

МЕТОДЫ ОБНАРУЖЕНИЯ АНОМАЛИЙ В СЕТЕВОМ ТРАФИКЕ НА ОСНОВЕ ОПТИМАЛЬНОЙ ЛИНЕЙНОЙ ФИЛЬТРАЦИИ

Специальность 2.3.6. Методы и системы защиты информации, информационная безопасность

Автореферат

диссертации на соискание ученой степени кандидата технических наук

Самара – 2026

Работа выполнена на кафедре информационной безопасности ФГБОУ ВО «Поволжский государственный университет телекоммуникаций и информатики»

Научный руководитель: доктор технических наук, профессор, **Карташевский Вячеслав Григорьевич**

Официальные оппоненты:
Новиков Сергей Николаевич, доктор технических наук, профессор, ФГБОУ ВО «Сибирский государственный университет телекоммуникаций и информатики», и.о. заведующего кафедрой безопасности информации и технологий
Цветов Виктор Петрович, кандидат физико-математических наук, доцент, ФГАОУ ВО «Самарский национальный исследовательский университет имени академика С.П. Королева», доцент кафедры безопасности информационных систем

Ведущая организация: федеральное государственное автономное образовательное учреждение высшего образования «Томский государственный университет систем управления и радиоэлектроники», г. Томск

Защита диссертации состоится 25 декабря 2026 года в 13 ч. 00 мин. на заседании диссертационного совета 24.2.479.07 на базе ФГБОУ ВО «Уфимский университет науки и технологий», по адресу: 450008, г. Уфа, ул. К. Маркса, д. 12.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВО «Уфимский университет науки и технологий» и на сайте *https://uust.ru/*

Автореферат разослан «___» _____ 2026 года.

Ученый секретарь
диссертационного совета,
доктор технических наук  Вульфин Алексей Михайлович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. С ростом объемов глобального сетевого трафика и развитием инструментов искусственного интеллекта количество и масштаб кибератак неуклонно увеличивается, что актуализирует проблему их своевременного и точного обнаружения. Среди каналов распространения вредоносного программного обеспечения (ВПО) компьютерные сети по частоте использования уступают лишь электронной почте, а число зондирующих и DDoS-атак, в том числе с использованием ботнетов, стабильно растет.

Злоумышленник не может получить доступ к трафику, не воздействуя на него, то есть не установив соединение с некоторым сервером и не отправив пакеты данных. Появление нового источника трафика может влиять на состояние сети и значения её статистических характеристик. В отличие от систем обнаружения вторжений (СОВ) на основе злоупотреблений, такие изменения могут быть обнаружены СОВ на основе выявления аномалий, даже при отсутствии известных сигнатур для реализуемой атаки.

Большинство современных систем Network Traffic Analysis (NTA) реализуют обнаружение на основе сигнатур, некоторые применяют простые статистические методы, но чаще сразу переходят к более сложным алгоритмам машинного обучения и нейросетям. Тем не менее, статистические методы, основанные на анализе характеристик временных рядов, остаются актуальными и представляют значительный интерес благодаря высокой точности выявления аномалий при сравнительно более низких вычислительных затратах и лучшей интерпретируемости результатов.

Данные факторы обусловили актуальность исследования, которое посвящено разработке алгоритмов обнаружения аномалий в сетевом трафике на основе оптимальной линейной фильтрации с применением фильтра Колмогорова-Винера и анализа статистических характеристик трафика. Описание разработанных методов и результаты их экспериментального исследования могут быть использованы для повышения безопасности сетей и защиты от кибератак.

Степень разработанности темы исследования. Проблема обеспечения информационной безопасности компьютерных сетей и обнаружения сетевых атак является предметом пристального внимания мирового научного сообщества на протяжении последних десятилетий. Основы построения СОВ были заложены в трудах таких ученых как Anderson J.P., Neumann P.G. и Denning D.E., в работе которой впервые вводится понятие профилей нормального поведения пользователя. Становление и развитие научного направления защиты информации в отечественной науке связано с именами

таких ученых как Котенко И.В., Зегжда П.Д., Герасименко В.А., Кульба В.В., Шелупанов А.А.

Самоподобная природа сетевого трафика, впервые отмеченная в работах Leland W.E., Taqqu M.S., Willinger W., Wilson D.V., поставила под сомнение эффективность классических пуассоновских моделей при анализе телекоммуникационных сетей. Значительный вклад в изучение свойств фрактального сетевого трафика и моделирование его параметров внесли Шелухин О.И., Осин А.В., Потапов А.А., Карташевский И.В., Треногин Н.Г., Соколов Д.Е., Лихтциндер Б.Я., Чабдаров Ш.М., Благов А.В., Привалов А.Ю., Багманов В.Х., Султанов А. Х., Nason G.P., Veitch D., Zukerman M.

Разработке методов обнаружения аномалий на основе интеллектуального анализа данных и машинного обучения посвящены, например, работы Васильева В.И., Вульфина А.М., Картака В.М., Гвоздева В.Е., Гетьмана А.И., Соколова А.Н., Cristophe D. Однако большинство методов на основе нейронных сетей, несмотря на высокую точность, требуют существенных вычислительных затрат, особенно при обучении, что затрудняет их применение в режиме реального времени на магистральных каналах связи.

Формирование основ теории оптимального оценивания и фильтрации случайных процессов, а также создание соответствующего математического аппарата связаны с именами таких выдающихся ученых, как Wiener N., Kalman R.E., Bucy R.S., Haykin S., Колмогоров А.Н., Левин Б.Р., Смирнов Н.В., Стратонович Р.Л. Попытки применения статистических методов для обнаружения сетевых атак предпринимались в работах ряда современных исследователей, например Фаткиевой Р.Р., Рыбакова С.Ю., Ngong S., Lakhina A., Crovella M., Barford P. В большинстве работ фильтрация применяется для сглаживания временных рядов трафика, на основе прогнозирования которых строятся предположения о наличии аномалий в трафике, однако упрощенные линейные модели не всегда корректно описывают самоподобный характер трафика и не всегда позволяют обнаруживать распределенные по времени аномалии. Таким образом, актуальной научной задачей является разработка методов обнаружения аномалий, сочетающих в себе вычислительную эффективность линейной фильтрации с высокой точностью детектирования для минимизации ошибок первого и второго рода.

Объектом исследования является сетевой трафик в виде последовательности пакетов, порождаемых различными сетевыми устройствами.

Предметом исследования являются аномалии в сетевом трафике, выражающиеся в изменении его статистических характеристик.

Целью диссертационной работы является повышение защищенности телекоммуникационных систем за счет повышения эффективности алгоритмов обнаружения аномалий в сетевом трафике на основе его статистических характеристик.

Для достижения цели исследования поставлены и решены следующие задачи:

1. Провести анализ статистических характеристик реального сетевого трафика, таких как автокорреляционная функция (АКФ) и вероятностное распределение интервалов времени между пакетами, размеров пакетов и других параметров, и установить влияние различных видов атак на статистические характеристики трафика.
2. Провести моделирование реалистичного сетевого трафика, соответствующего нормальной активности пользователей и проведению атаки, изучить свойства объединенного трафика.
3. Разработать методы обнаружения аномалий в сетевом трафике на основе оптимальной линейной фильтрации.
4. Исследовать возможность применения признаков на основе статистических характеристик трафика в методах машинного обучения для повышения точности обнаружения аномалий.
5. Оценить эффективность разработанных методов на смоделированном и реальном трафике по различным метрикам.

Методы исследования включают в себя методы теории вероятностей и математической статистики, анализ временных рядов, в том числе методы корреляционного анализа, проверки статистических гипотез, моделирования случайных процессов, машинного обучения, и различные метрики оценки алгоритмов классификации.

Научная новизна полученных результатов:

1. Разработан алгоритм определения численных индикаторов сетевых атак по статистическим характеристикам трафика, *основанный* на методе «случайного леса», *отличающийся* возможностью определения не только важности признаков, но и их пороговых значений, и *позволяющий* оценивать их статистическую значимость.
2. Разработан метод моделирования случайных процессов, представляющих самоподобный Интернет-трафик, *основанный* на преобразовании распределения Рэлея, *отличающийся* применением копулы Фарли-Гумбеля-Моргенштерна для выражения аналитической связи между АКФ исходного и преобразованного процессов, и *позволяющий* повысить точность воспроизведения заданных статистических характеристик при моделировании легитимного трафика.
3. Разработаны методы обнаружения аномалий в сетевом трафике и реализующий их алгоритм, *основанные* на вычислении минимальной среднеквадратической ошибки фильтрации и сравнении её с эталонным значением для легитимного трафика, *отличающиеся* возможностью учитывать шум наблюдений, и *позволяющие* выявлять сетевые атаки, связанные с изменением статистических характеристик трафика.

4. Предложен механизм формирования признаков для моделей машинного обучения, *основанный* на вычислении минимальной среднеквадратической ошибки фильтрации для различных представлений трафика, *отличающийся* объединением множества отсчетов АКФ в единый скалярный признак ошибки, и *позволяющий* повысить эффективность обнаружения сетевых атак алгоритмами обучения без учителя.

Положения, выносимые на защиту:

1. Алгоритм определения численных индикаторов сетевых атак по статистическим характеристикам трафика, основанный на методе «случайного леса».

2. Метод моделирования случайных процессов, представляющих самоподобный Интернет-трафик, основанный на преобразовании распределения Рэлея и применении копулы Фарли-Гумбеля-Моргенштерна.

3. Методы обнаружения аномалий в сетевом трафике и реализующий их алгоритм, основанные на вычислении минимальной среднеквадратической ошибки фильтрации и сравнении её с эталонным значением для легитимного трафика.

4. Механизм формирования признаков для моделей машинного обучения, предназначенных для обнаружения сетевых атак, основанный на вычислении минимальной среднеквадратической ошибки фильтрации для различных представлений трафика.

Теоретическая значимость исследования заключается в разработке алгоритма определения значимых статистических признаков различных видов сетевых атак, способа моделирования случайных процессов с заданными характеристиками на основе копул, методов обнаружения аномалий в сетевом трафике на основе минимальной среднеквадратической ошибки оптимального фильтра Колмогорова-Винера, которые позволяют эффективно выявлять аномальные последовательности в трафике. Данные результаты вносят важный вклад в совершенствование моделей и методов обеспечения безопасности сетевого взаимодействия, так как их применение может позволить оперативно выявлять потенциальные угрозы и снизить риск инцидентов безопасности.

Практическая значимость исследования заключается в возможности применения полученных результатов для разработки перспективных средств защиты информации и мониторинга сетевого трафика. Внедрение разработанных алгоритмов в существующие корпоративные системы безопасности позволяет повысить число выявленных аномалий, обеспечив своевременное обнаружение и предупреждение атак и снижение количества инцидентов информационной безопасности.

Достоверность результатов работы обеспечивается корректным и обоснованным применением методов исследования, проведением вычислительных экспериментов на открытых наборах данных, проверкой

статистической значимости полученных результатов, а также апробацией результатов исследований на международных и всероссийских конференциях и публикацией результатов в рецензируемых научных изданиях.

Связь исследований с научными программами. Часть исследования выполнена в рамках гранта Фонда содействия инновациям УМНИК «Разработка программного комплекса для выявления нарушений информационной безопасности на основе анализа статистических зависимостей в сетевом трафике» и гранта МТУСИ на реализацию научного проекта «Обнаружение аномалий сетевого трафика на основе линейного преобразования трафика по критерию минимума среднеквадратической ошибки» №37/21-к.

Соответствие паспорту специальности. Результаты диссертационной работы соответствуют следующим пунктам паспорта специальности 2.3.6 «Методы и системы защиты информации, информационная безопасность»: **п. 5** «Методы, модели и средства (комплексы средств) противодействия угрозам нарушения информационной безопасности в открытых компьютерных сетях, включая Интернет» (положения 2, 3), **п. 6** «Методы, модели и средства мониторинга, предупреждения, обнаружения и противодействия нарушениям и компьютерным атакам в компьютерных сетях» (положения 1, 3, 4), **п. 15** «Принципы и решения (технические, математические, организационные и др.) по созданию новых и совершенствованию существующих средств защиты информации и обеспечения информационной безопасности» (положения 1, 3, 4).

Апробация работы. Основные результаты и положения диссертационного исследования докладывались и обсуждались на международных и всероссийских научных и научно-технических конференциях. Результаты работы внедрены в ООО «Идентификация» и в учебный процесс кафедры информационной безопасности ФГБОУ ВО «Поволжский государственный университет телекоммуникаций и информатики».

Публикации. По теме диссертационного исследования всего опубликовано 22 работы, отражающих основные полученные результаты, из них: 5 статей в ведущих научных журналах из Перечня рецензируемых научных изданий ВАК при Министерстве науки и высшего образования РФ, 1 научная работа в изданиях, индексируемых ведущими международными реферативными базами Web of Science и Scopus, 6 тезисов международных и 10 всероссийских научных и научно-технических конференций. Во время выполнения научного исследования разработаны 3 программы для ЭВМ, свидетельства о государственной регистрации: №2022614836, №2023664160, №2024611818.

Сведения о личном вкладе автора. Основные результаты диссертационной работы, представляющие научную новизну, а также анализ с соответствующими выводами, выполнены автором самостоятельно. Некоторые из этих результатов подготовлены к публикации в сотрудничестве с соавторами, при этом определяющий вклад вносил автор данной работы.

Структура и объем работы. Работа состоит из введения, четырех глав, заключения, списка сокращений, списка литературы и приложений, содержит 52 рисунка, 11 таблиц, 107 формул, 2 приложения, выполнена на 147 страницах. Список использованной литературы включает 139 источников.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы исследования, определены объект и предмет диссертационного исследования, цель работы и решаемые задачи. Отмечена научная новизна полученных результатов, их теоретическая и практическая значимость, представлены положения, выносимые на защиту.

В первой главе проводится анализ статистических характеристик реального сетевого трафика с целью выявления его отличительных особенностей. Рассматриваются различные методы анализа трафика. Проведено исследование трафика из датасета CIC-IDS2017. Установлено, что длина 99% сетевых потоков не превышает 111 пакетов, что недостаточно для достоверной оценки статистических характеристик. При разбиении трафика по фиксированным временным интервалам (окнам), длительностью 1 с, среднее число пакетов в окне в 17 раз выше среднего размера потока, в связи с чем особую значимость приобретают статистические методы исследования агрегированного трафика без выделения потоков.

Сравниваются различные представления трафика: в виде случайных последовательностей (интервалов времени между поступлениями пакетов и размеров пакетов) и случайных временных рядов (суммарного количества и размера пакетов в байтах за единицу времени). Установлена аналитическая связь между представлениями. Исследуются средние, СКО, и АКФ представлений, производится оценка показателя Херста. Подтверждено наличие устойчивой медленно затухающей АКФ и высоких значений показателя Херста у различных представлений трафика, что свидетельствует о наличии самоподобия.

Установлено влияние различных видов атак на статистические характеристики трафика. Для выделения численных индикаторов аномалий разработан алгоритм на основе метода «случайного леса», включающий в себя агрегацию пороговых условий для каждого признака в узлах деревьев. Полученный в результате набор пар пороговых значений и соответствующего им прироста доли аномальных объектов можно рассматривать как выборку случайной величины, для оценки которой применяется ядерная оценка плотности и находится мода распределения, которая и определяет агрегированный порог. Для каждого признака может быть получено от 0 до 2 наиболее значимых порогов разбиения, которые могут образовывать интервал нормальных или аномальных значений. Для проверки статистической значимости выявленных порогов применяется критерий согласия Колмогорова-

Смирнова. Подробное описание шагов алгоритма представлено в тексте диссертации, указаны его ограничения.

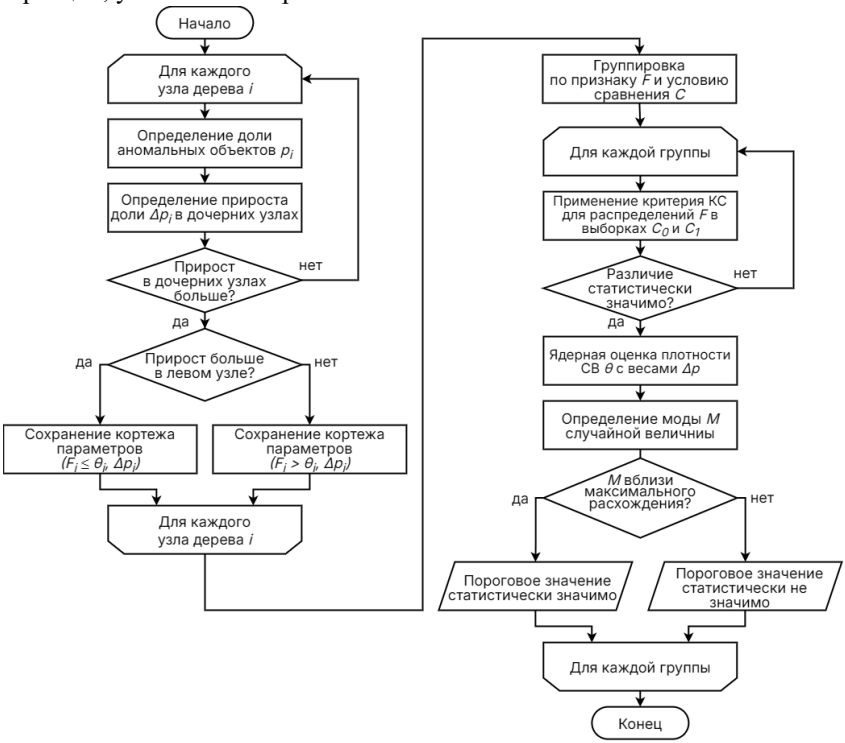


Рисунок 1 – Алгоритм выделения численных порогов

Получены значения численных индикаторов атак из данных CIC-IDS2017. Несмотря на то, что эти значения специфичны для трафика сети, в которой производилось моделирование, выделенные закономерности совпадают с интуитивным пониманием особенностей атак, а разработанный подход является достаточно общим. Полученные индикаторы для отдельных видов атак показывают, что их можно обнаружить по изменению АКФ последовательности.

Во второй главе рассмотрены задачи и методы моделирования трафика в телекоммуникационных сетях, в том числе в виде случайного процесса с заданными статистическими характеристиками. Рассмотрены методы моделирования самоподобного интернет-трафика, такие как модель ON/OFF источников и модель Poisson-Pareto Burst Process, которая позволяет генерировать в симуляторе ns-3 самоподобный трафик с заданным значением показателя Херста. Рассмотрены различные варианты моделирования DoS-атак и их основные признаки. Исследованы способы объединения записей трафика, представляющие легитимный трафик и трафик атаки различной интенсивности.

На основе известных методов разработан новый способ моделирования случайных процессов с одновременно заданным законом распределения и АКФ, позволяющий выразить взаимосвязь между АКФ исходного и преобразованного процессов через параметр *копулы* Фарли-Гумбеля-Моргенштерна (ФГМ). Разработанный способ состоит в двухэтапном преобразовании исходных нормальных процессов: сначала в процесс с экспоненциальным распределением или распределением Рэлея, а затем в процесс с распределением Вейбулла. Данное распределение обладает «тяжелым хвостом» при параметре формы $p < 1$ и часто используется для моделирования сетевого трафика.

Копула ФГМ используется для установления аналитической связи между коэффициентами корреляции исходного и преобразованного процессов через параметр θ, который для распределения Рэлея равен r/π, а для Вейбулла:

$$\theta = r \cdot \frac{\Gamma\left(1+\frac{2}{p}\right) - \Gamma^2\left(1+\frac{1}{p}\right)}{\Gamma^2\left(1+\frac{1}{p}\right)\left[1-\left(1-2^{-\left(1+\frac{1}{p}\right)}\right)\cdot 2^{\left(1-\frac{1}{p}\right)}\right]} = r \cdot k(p). \qquad (1)$$

Проведено моделирование случайных процессов с распределением Вейбулла и двумя видами АКФ: экспоненциальной и треугольной. Так как копула ФГМ позволяет моделировать только слабую корреляционную связь $|r| < 1/3$, также рассматриваются отмасштабированные версии АКФ, где значения для всех лагов, кроме первого, делятся на 3. Сравниваются известные методы моделирования и методы с использованием разработанного способа в различных вариантах. Способ на основе копулы позволяет снизить евклидово расстояние между заданной и эмпирической АКФ в среднем на 55%, и уменьшить значение критерия согласия Колмогорова-Смирнова на 34% по сравнению с известными численными методами.

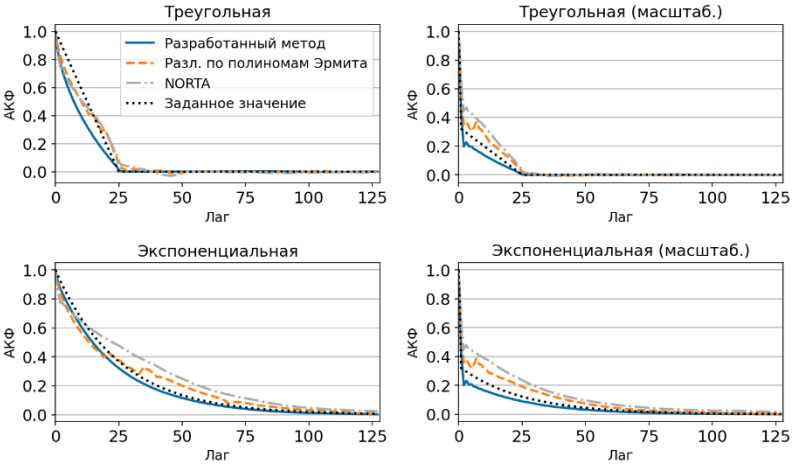


Рисунок 2 – АКФ смоделированного трафика

В третьей главе проведен обзор существующих методов обнаружения сетевых аномалий в компьютерных сетях. Описаны основные методы, применяющиеся в СОВ. Разработаны новые методы на основе минимальной среднеквадратической ошибки фильтра Винера.

Фильтр Винера применяется для получения оптимальной оценки смеси сигнала с шумом: $x(t) = s(t) + n(t)$. Минимальная среднеквадратическая ошибка фильтрации (оценивания), которая неизбежно возникает при взаимном наложении энергетических спектров сигнала и шума, равна при этом:

$$\varepsilon_{\min}^2 = \sigma_s^2 - \mathbf{r}_{\mathbf{xs}}^{\mathbf{T}} \mathbf{R}_x^{-1} \mathbf{r}_{\mathbf{xs}}. \quad (2)$$

NMSE представляет собой нормализованное относительно дисперсии полезного сигнала значение среднеквадратической ошибки и определяется как:

$$\epsilon = \frac{\varepsilon_{\min}^2}{\sigma_s^2} = 1 - \frac{\sigma_y^2}{\sigma_s^2}. \quad (3)$$

Для оптимального фильтра Винера $0 \le \epsilon \le 1$. Можно установить «эталонное» значение NMSE для смеси случайного процесса, представляющего легитимный трафик, и шума наблюдений с заданной мощностью. Сравнивая для наблюдаемого трафика текущее значение среднеквадратической ошибки с эталонным, по величине отклонения можно судить о наличии в трафике аддитивной аномальной составляющей. Если на вход фильтра поступает сигнал вместе с аномалией: $x'(t) = s(t) + n(t) + a(t)$, то отклонение от эталонного значения равняется дисперсии аномальной составляющей после фильтрации.

Если рассматривать шум в смеси сигнала на входе фильтра, как аномалию: $x(t) = s(t) + a(t)$, тогда минимальное значение ошибки $\varepsilon_{\min}^2$ (MMSE) и будет являться признаком аномалии, так как в случае отсутствия шума минимальная ошибка равна нулю.

Рассмотрены варианты выбора пороговых значений на основе ROC-кривых бинарного классификатора или собственных чисел матрицы корреляции легитимного трафика. Через отношение Рэлея можно выразить пороговые значения, соответствующие проекции АКФ на оси в направлении минимальной и максимальной дисперсии полезного сигнала.

$$\varepsilon_{\min}^2 \in \left[\sigma_s^2 - \frac{\|\mathbf{r}_{xs}\|^2}{\lambda_{\min}}, \sigma_s^2 - \frac{\|\mathbf{r}_{xs}\|^2}{\lambda_{\max}}\right]. \quad (4)$$

Разработан алгоритм обнаружения аномалий, который включает в себя шаги с применением методов оптимальной линейной фильтрации. Поскольку фильтр Винера применим только к стационарным процессам, алгоритм включает в себя фазу обучения, которая необходима для выделения типичных сценариев сетевой активности $U = \{U_1, U_2, \dots, U_n\}$, описывающих трафик в определенное время или для определенных приложений, соответствующих стационарным периодам. Затем производится синтез фильтра Винера и определение минимальной ошибки фильтрации для каждого сценария. На этапе

применения алгоритма сначала выполняется классификация текущего трафика. Если ни один из сценариев не определяется как подходящий, или характеристики трафика подходят под сценарий, который в данный момент не является допустимым, это может также служить признаком аномалии.



Рисунок 3 – Диаграмма фазы обучения

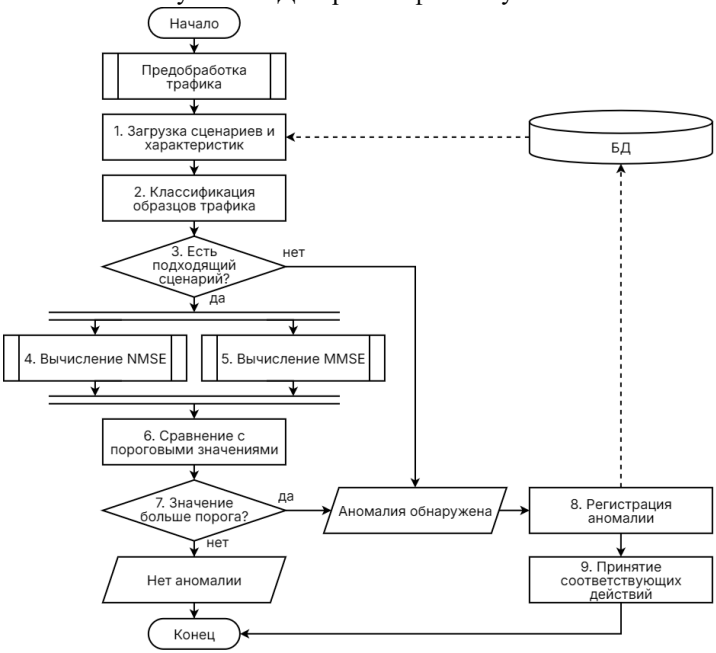


Рисунок 4 – Диаграмма обнаружения аномалий

Главным ограничением данного алгоритма является то, что он основан на сравнении АКФ, поэтому может обнаруживать только те аномалии, которые влияют на статистические характеристики сетевого трафика, а также то, что при изменении структуры сети или характера её использования необходимо обновление моделей нормального трафика.

В четвертой главе представлены результаты моделирования обнаружения аномалий для различных сценариев. Исследовано влияние формы АКФ на эффективность обнаружения аномалий. Установлено, что при любой форме АКФ полезного трафика и аномалии, метод позволяет получить статистически значимые результаты отклонения NMSE от эталонного значения.

Исследовано влияние мощности шума наблюдений, используемого на этапе обучения, на результаты обнаружения аномалий. Установлено, что обнаружение достоверно производится при ОСШ от 0,01 до 100.

Исследована эффективность алгоритма при применении к реальному сетевому трафику из датасетов MAWI и CIC-IDS2017. Установлено, что использование среднеквадратической ошибки фильтрации в качестве бинарного классификатора позволяет получить значения AUC (площади под кривой) до 0,8 и 0,95 при рассмотрении отклонения и значения ошибки соответственно, для брутфорс атак, сканирования портов, DoS-атак и атак внедрения.

Было проведено сравнение с классическими методами, применяющимися для обнаружения аномалий на основе статистических характеристик, для среднего: выход скользящего среднего за интервал в 3σ, CUSUM, для АКФ: значение упомянутой ранее метрики CMD, для распределения: статистика критерия согласия Колмогорова-Смирнова и расстояние Вассертштейна. Метод на основе отклонения ошибки дает наилучшие результаты при применении к последовательности межпакетных интервалов, позволяя обнаруживать брутфорс, сканирование портов и атаки внедрения. Метод на основе значения ошибки позволяет получить наибольшие значения AUC, превышающие значения для других методов, при обнаружении атак внедрения и сканирования портов.

Предложено использование значений NMSE, получаемых с использованием разработанных алгоритмов, в качестве дополнительных признаков для методов машинного обучения. Рассмотрены различные методы обучения без учителя: гауссовские смеси распределений (GMM), ядерная оценка плотности (KDE), изоляционный лес (IF), анализ главных компонент (PCA) и одноклассовый метод опорных векторов (OCSVM). Рассматриваются потоки двух узлов, один из которых считается «нормальным», а другой «аномальным».

Обучение моделей производится на выборке, содержащий признаки только нормальных потоков, в тестовой содержится одинаковое количество нормальных и аномальных потоков. Установлено, что добавление отклонения и значения среднеквадратической ошибки к вектору признаков для различных

представлений трафика в среднем приводит к увеличению метрики AUC, максимальный прирост составляет 4,91%. Для вектора признаков на основе статистических характеристик добавление новых признаков позволяет увеличить AUC в среднем от 6,39% до 14,06%, в зависимости от модели.

Таблица 1 – Прирост AUC для статистических характеристик

	Модель					
Значение	OCSVM	PCA	GMM	IF	KDE	Среднее
AUC	0,46	0,57	0,55	0,67	0,68	0,59
ΔNMSE, %	4,51	11,85	5,31	7,41	2,98	6,39
ΔMMSE, %	30,52	1,24	19,58	-0,77	2,40	9,09
ΔBOTH, %	31,76	12,64	22,22	7,81	2,88	14,06

В заключении представлены основные выводы и результаты, полученные в ходе проведенного исследования.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

В диссертационной работе решена научная задача по повышению защищенности телекоммуникационных систем посредством разработки алгоритмов обнаружения аномалий, которые могут быть вызваны атаками злоумышленника. Основные выводы и результаты диссертационной работы:

1. Проведен анализ статистических характеристик реального сетевого трафика, сгенерированного в локальной сети и взятого из открытых наборов данных. Разработан алгоритм определения численных индикаторов атак и установлено, что воздействие сетевых атак на трафик приводит к изменению АКФ и вероятностного распределения, из чего следует возможность их обнаружения на основе рассмотрения данных характеристик.
2. Проведено моделирование самоподобного сетевого трафика и DoS-атак, исследованы различные варианты объединения трафика и его свойства. Разработан метод моделирования случайных процессов с распределением Вейбулла, представляющих легитимный трафик, который позволяет улучшить точность воспроизведения заданных статистических характеристик по сравнению с известными численными методами.
3. Разработаны методы обнаружения аномалий на основе оптимального линейного фильтра Колмогорова-Винера и реализующие их алгоритмы, которые позволяют определить аномалии на основе несовпадения характеристик наблюдаемого в данный момент трафика с легитимным.
4. Разработан механизм формирования признаков для моделей машинного обучения, предназначенных для обнаружения сетевых атак, на основе минимального значения среднеквадратической ошибки фильтрации и ее отклонения от эталонного значения.

5. Проведена оценка эффективности разработанных методов и алгоритмов на смоделированном и реальном трафике в различных условиях.

Все поставленные задачи были решены. Разработанные методы позволили повысить эффективность обнаружения сетевых атак по сравнению с существующими подходами. Добавление в вектор признаков для моделей машинного обучения без учителя новых параметров, основанных на вычислении MSE, позволило увеличить точность классификации аномалий, за счет чего была достигнута цель исследования.

Перспективы дальнейшей разработки темы включают в себя исследование характеристик статистической взаимосвязи отсчетов трафика выше второго порядка, называемых *кумулянтами*. Для алгоритма определения численных индикаторов атак целесообразно учитывать порядок порогов в решающих деревьях. Для моделирования случайных последовательностей помимо ФГМ могут быть рассмотрены другие виды копул, которые позволяют выразить связь через коэффициент корреляции. Разработанные методы могут быть расширены для обнаружения не только наличия аномалии, но и определение её типа, для чего могут использоваться, например, несколько фильтров для каждого известного типа аномалии. Также, текущий подход рассматривает аномалии только в виде аддитивного процесса, поэтому перспективной задачей является исследование других моделей аномалий.

СПИСОК ОСНОВНЫХ ПУБЛИКАЦИЙ ПО ТЕМЕ ИССЛЕДОВАНИЯ

Публикации в рецензируемых научных изданиях из перечня ВАК:

1. *Поздняк И.С.* Выявление DoS-атак с помощью анализа статистических характеристик трафика / И.С. Поздняк, А.И. Плаван // Инфокоммуникационные технологии. – 2021. – Т. 19. – № 1. – С. 73-80

2. *Плаван А.И.* Среднеквадратическая ошибка фильтрации как критерий обнаружения аномалий сетевого трафика / А.И. Плаван, В.Г. Карташевский // Вестник Российского нового университета. Серия: Сложные системы модели, анализ и управление. – 2023. – № 1. – С. 94-101

3. *Плаван А.И.* Метод выявления аномалий в сетевом трафике на основе фильтра Колмогорова-Винера / А.И. Плаван // Вестник УрФО. Безопасность в информационной сфере. – 2024. – № 1 (51). – С. 53-59

4. *Плаван А.И.* Алгоритм определения численных индикаторов аномалий трафика на основе модели случайного леса / А.И. Плаван // Инфокоммуникационные технологии. – 2026. – Т. 24. – № 1 (93). – С. 26-32

5. *Плаван А.И.* Применение копулы для моделирования случайных процессов с распределением Вейбулла и заданной автокорреляционной функцией / А.И. Плаван, В.Г. Карташевский // Известия Института инженерной физики. Серия: Информатика, вычислительная техника и управление. – 2026. – № 3 (81). – С. 47-52

Публикации в изданиях, индексируемых в Scopus и Web of Science:

6. *Plavan A.I.* Statistical analysis of aggregated network traffic model from multiple on/off sources / A.I. Plavan, V.G. Kartashevsky, I.S. Pozdnyak // Proc. SPIE 12564, 2nd International Conference on Computer Applications for Management and Sustainable Development of Production and Industry (CMSD-II-2022) / ред. A. Gibadullin, S. Sadullozoda. – Dushanbe, Tajikistan: SPIE, 21–23 декабря 2022 г. – Т. 12564. – С. 118-124

Другие публикации по теме диссертации:

7. *Поздняк И.С.* Исследование трафика на наличие атак / И.С. Поздняк, В.Г. Карташевский, А.И. Плаван // Проблемы техники и технологий телекоммуникаций ПТиТТ-2020: Материалы XXII Международной научно-технической конференции. – Самара: ПГУТИ, 17–20 ноября 2020 г. – С. 206-207

8. *Карташевский В.Г.* Моделирование и анализ DDoS-атаки / В.Г. Карташевский, И.С. Поздняк, А.И. Плаван // Материалы XXVIII Российской научно-технической конференции профессорско-преподавательского состава, научных сотрудников и аспирантов университета с приглашением ведущих ученых и специалистов родственных вузов и организаций. – Самара: ПГУТИ, 5–8 апреля 2021 г. – С. 63-64

9. *Плаван А.И.* Выявление аномалий в сетевом трафике по критерию среднеквадратической ошибки фильтрации линейного преобразования / А.И. Плаван, И.С. Поздняк // Фундаментальные и прикладные аспекты компьютерных технологий и информационной безопасности: материалы VII Всероссийской научно-технической конференции. – Таганрог: Южный федеральный университет, 5–11 апреля 2021 г. – С. 70-74

10. *Плаван А.И.* Сравнительный анализ статистических характеристик DDoS-атак и нормального трафика / А.И. Плаван, В.Г. Карташевский, И.С. Поздняк // Актуальные проблемы науки и техники: Материалы I Международной научно-технической конференции. – Сарапул: Издательство УИР ИжГТУ имени М. Т. Калашникова, 20–22 мая 2021 г. – С. 551-555

11. *Карташевский В.Г.* Исследование статистических характеристик вредоносного трафика / В.Г. Карташевский, А.И. Плаван, И.С. Поздняк // Проблемы техники и технологий телекоммуникаций ПТиТТ-2021: Материалы XXIII Международной научно-технической конференции. – Самара: ПГУТИ, 23–26 ноября 2021 г. – С. 165-166

12. *Карташевский В.Г.* Статистический анализ агрегированной модели сетевого трафика / В.Г. Карташевский, А.И. Плаван, И.С. Поздняк // Материалы XXIX Российской научно-технической конференции профессорско-преподавательского состава, научных сотрудников и аспирантов университета с приглашением ведущих ученых и специалистов родственных вузов и организаций. – Самара: ПГУТИ, 22–25 марта 2022 г. – С. 58-59

13. *Плаван А.И.* Исследование влияния скорости on/off источника на характеристики моделируемого интернет-трафика / А.И. Плаван // Доклады 77-й всероссийской конференции «Радиоэлектронные устройства и системы для

инфокоммуникационных технологий» посвящённой дню радио (РЭУС-2022). – М.: РНТОРЭС имени А.С. Попова, 8–10 июня 2022 г. – С. 102-105

14. *Плаван А.И.* Обнаружение аномалий сетевого трафика на основе фильтрации линейного преобразования трафика по критерию минимума среднеквадратической ошибки / А.И. Плаван // Сборник трудов II Всероссийской научной школы-семинара «Современные тенденции развития методов и технологий защиты информации». Москва, МТУСИ. – М.: ООО «Брис-М», 18–21 октября 2022 г. – С. 189-198

15. *Плаван А.И.* Моделирование самоподобного трафика с использованием модели Poisson-Pareto-Burst-Process в симуляторе ns-3 / А.И. Плаван // Материалы XXVII Международной научно-технической конференции «Современные средства связи». – Минск, Республика Беларусь: Белорусская государственная академия связи, 27–28 октября 2022 г. – Т. 1. – С. 197-199

16. *Плаван А.И.* Определение ошибки фильтрации трафика с целью обнаружения атаки / А.И. Плаван, В.Г. Карташевский, И.С. Поздняк // Материалы XXIV Международной научно-технической конференции и материалы XX Международной научно-технической конференции (ПТиТТ-2022). В 2-х томах. – Уфа: УУНиТ, 23–25 ноября 2022 г. – Т. 1. – С. 217-219

17. *Плаван А.И.* Обнаружение аномалий сетевого трафика на основе рассмотрения автокорреляционной функции в скользящем окне / А.И. Плаван // Сборник материалов XIII межведомственной научной конференции «Актуальные направления развития систем обеспечения безопасности объектов государственной охраны и охраняемых объектов, специальной связи для нужд органов государственной власти и специального информационного обеспечения государственных органов». – Орёл: Академия ФСО России, 10 февраля 2023 г. – С. 2

18. *Плаван А.И.* Среднеквадратическая ошибка фильтрации в качестве метрики расстояния для корреляционных матриц / А.И. Плаван // Актуальные проблемы радиоэлектроники и телекоммуникаций: материалы Всероссийской научно-технической конференции. – Самара: ООО «Артель», 25–28 апреля 2023 г. – С. 169-172

19. *Плаван А.И.* Исследование корреляционных свойств сетевого трафика при воздействии DoS-атаки / А.И. Плаван // Информационные технологии обеспечения комплексной безопасности в цифровом обществе: сборник материалов VI Всероссийской молодежной научно-практической конференции с международным участием. – Уфа: РИЦ УУНиТ, 19–20 мая 2023 г. – С. 105-108

20. *Плаван А.И.* Сравнительный анализ метрик для определения схожести корреляционных свойств трафика / А.И. Плаван, В.Г. Карташевский // Сборник докладов и тезисов XIII Всероссийской научно-практической конференции «Проблемы передачи информации в инфокоммуникационных системах». – Волгоград: Волгоградский государственный университет, 26 мая 2023 г. – С. 27-30

21. *Плаван А.И.* Использование среднеквадратической ошибки фильтрации для обнаружения кратковременных аномалий сетевого трафика / А.И. Плаван, В.Г. Карташевский // Проблемы техники и технологий телекоммуникаций ПТиТТ-2023: Материалы XXV Международной научно-технической конференции. – Казань: КНИТУ-КАИ, 22–24 ноября 2023 г. – Т. 1. – С. 502-504

22. *Плаван А.И.* Исследование метода обнаружения аномалий в сетевом трафике на основе оптимального фильтра Колмогорова-Винера / А.И. Плаван // Актуальные проблемы радиоэлектроники и телекоммуникаций: материалы Всероссийской конференции. – Самара: ООО «Артель», 23–26 апреля 2024 г. – С. 228-231

Свидетельства о государственной регистрации программ для ЭВМ:

23. Свидетельство о государственной регистрации программы для ЭВМ №2022614836. Программный комплекс «PcapStat» : №2022613421 : заявл. 09.03.2022 : опубл. 28.03.2022, Бюл. № 4 / А.И. Плаван. – 1 с.

24. Свидетельство о государственной регистрации программы для ЭВМ №2023664160. PcapConvert : №2023663148 : заявл. 21.06.2023 : опубл. 03.07.2023, Бюл. № 7 / А.И. Плаван. – 1 с.

25. Свидетельство о государственной регистрации программы для ЭВМ №2024611818. PcapReplay : №2024611026 : заявл. 24.01.2024 : опубл. 24.01.2024, Бюл. № 2 / А.И. Плаван. – 1 с.

Диссертант  А.И. Плаван