

Б. Ю. Якобсон^{1✉}

Алгоритм адаптивной фильтрации информационных потоков в корпоративных вычислительных сетях

¹Национальный исследовательский ядерный университет «МИФИ»,
г. Москва, Российская Федерация
e-mail: yazon16300@gmail.com

Аннотация. В работе рассмотрена задача адаптивной фильтрации сетевого трафика в корпоративных вычислительных сетях в условиях роста числа и сложности кибератак. Предложена гибридная двухуровневая архитектура системы обнаружения вторжений, объединяющая модуль сигнатурного анализа, реализованный на языке Go, и модуль поведенческого анализа на языке Python на основе алгоритма Isolation Forest. Для объединения результатов модулей разработан механизм Score Fusion с взвешенным усреднением оценок статистического и машинного анализа. Разработана методика тестирования системы на смешанном трафике с контролируемой долей семи классов атак, приближенная к условиям реальной эксплуатации. Проведено экспериментальное исследование эффективности на виртуализированном стенде из 48 виртуальных машин длительностью 8 часов. Получены следующие показатели: Precision=1,000, TPR=0,942, FPR=0,000, F1-мера=0,970. Проведен сравнительный анализ с учебным решением Simple Python IDS и промышленной системой Snort 3, показавший, что предложенное решение превосходит учебные аналоги и сопоставимо по ключевым метрикам с промышленными системами при существенно меньшем потреблении оперативной памяти.

Ключевые слова: адаптивная фильтрация, обнаружение вторжений, сигнатурный анализ, поведенческий анализ, Isolation Forest, Score Fusion

B. Y. Yakobson^{1✉}

Adaptive traffic filtering algorithm for corporate computer networks

¹National Research Nuclear University MEPhI, Moscow, Russian Federation
e-mail: yazon16300@gmail.com

Abstract. The paper addresses the task of adaptive traffic filtering in corporate computer networks under the growing number and complexity of cyberattacks. A hybrid two-level intrusion detection architecture is proposed, combining a signature analysis module implemented in Go and a behavioral analysis module in Python based on the Isolation Forest algorithm. A Score Fusion mechanism with weighted averaging of statistical and machine learning scores is developed to combine the outputs of the modules. A testing methodology on mixed traffic with a controlled share of seven attack classes is proposed, approximating real-world operating conditions. An experimental study was conducted on a virtualized testbed of 48 virtual machines for 8 hours. The following metrics were obtained: Precision=1.000, TPR=0.942, FPR=0.000, F1-score=0.970. A comparative analysis with the Simple Python IDS educational solution and the industrial Snort 3 system demonstrated that the proposed solution outperforms educational analogues and is comparable in key metrics to industrial systems at significantly lower RAM consumption.

Keywords: adaptive filtering, intrusion detection, signature analysis, behavioral analysis, Isolation Forest, Score Fusion

Введение

Рост масштабов и сложности корпоративных вычислительных сетей (КВС) сопровождается увеличением числа и разнообразия кибератак на инфраструктуру организаций. По данным Positive Technologies, в 2024–2025 гг. доля инцидентов с использованием программ-вымогателей в отдельных кварталах увеличивалась на 13% и более [1]. В этих условиях традиционные подходы к фильтрации трафика, основанные на статических правилах и классическом сигнатурном анализе, не обеспечивают своевременного реагирования на ранее неизвестные угрозы.

Альтернативой являются адаптивные методы, сочетающие сигнатурный и поведенческий анализ с применением машинного обучения [2, 3]. Коммерческие решения класса NGFW реализуют подобные подходы, однако требуют значительных вычислительных ресурсов [4, 5], что ограничивает их применение в организациях малого и среднего размера. В академических работах рассматриваются преимущественно либо изолированный сигнатурный анализ [6–8], либо отдельные методы поведенческого анализа [3, 9]; вопросы совместной обработки результатов статистического и ML-анализа, а также эмпирического подбора параметров такой интеграции проработаны недостаточно.

В настоящей работе предлагается гибридная двухуровневая архитектура системы обнаружения вторжений, в которой сигнатурный модуль реализован на языке Go для минимизации накладных расходов, а модуль поведенческого анализа на основе алгоритма Isolation Forest – на языке Python. Результаты модулей объединяются оригинальным механизмом Score Fusion с взвешенным усреднением оценок статистического и машинного анализа, что обеспечивает одновременную устойчивость к известным сигнатурам атак и к ранее неизвестным поведенческим аномалиям. Для оценки эффективности предложенного решения разработана методика тестирования на смешанном трафике с контролируемой долей семи классов атак (SQL-инъекции, XSS, SYN Flood, Port Scan, Host Scan, Path Traversal, Command Injection), приближенная к условиям реальной эксплуатации. Отдельное внимание уделено эмпирическому подбору порога бинарной классификации Score Fusion, обеспечивающего максимум F1-меры при сохранении нулевого уровня ложных срабатываний.

Актуальность исследования обусловлена необходимостью разработки алгоритма адаптивной фильтрации информационных потоков КВС на основе одновременного использования сигнатурных и поведенческих признаков, допускающего интеграцию в существующие средства защиты корпоративного сегмента. Предложенное решение ориентировано на защиту сетей малого и среднего размера, для которых развертывание промышленных систем класса Suricata или Snort 3 затруднено по соображениям стоимости и требуемой квалификации.

Целью исследования является разработка и экспериментальная оценка гибридного алгоритма адаптивной фильтрации информационных потоков в КВС, сочетающего сигнатурный и поведенческий анализ на основе алгоритма Isolation Forest.

Предлагаемый алгоритм

Предложенная система состоит из двух последовательно-параллельных модулей анализа и модуля объединения результатов. Сигнатурный модуль обеспечивает обнаружение известных угроз по базе регулярных выражений и паттернов (SQL-инъекции, XSS, флуд-атаки). Поведенческий модуль осуществляет анализ параметров трафика в скользящем временном окне и обнаружение статистических отклонений методами машинного обучения без учителя. Модуль Score Fusion выполняет взвешенное объединение оценок и принятие итогового решения. Общая схема обработки пакета представлена на рис. 1.

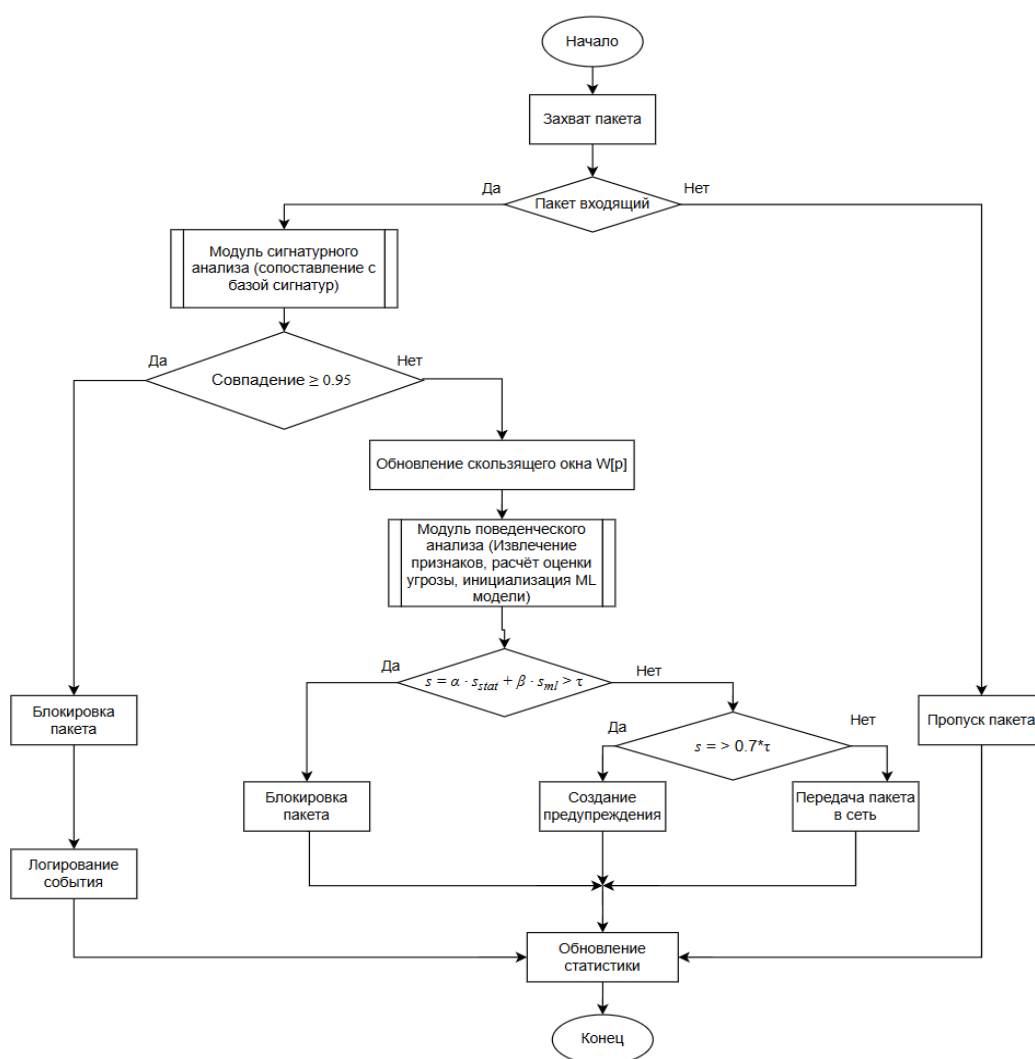


Рис. 1. Обобщенная схема алгоритма адаптивной фильтрации

В качестве модели машинного обучения выбран алгоритм Isolation Forest, обеспечивающий оценку аномальности за время $O(n \cdot \log n)$ без размеченной обучающей выборки [10]. Оценка аномальности вычисляется как

$$s(x, n) = 2^{-E(h(x))/c(n)}, \quad (1)$$

где $E(h(x))$ – математическое ожидание длины пути до изоляции наблюдения x в ансамбле случайных деревьев; $c(n)$ – нормализующий коэффициент для выборки объема n . Значения $s(x, n)$, близкие к 1, соответствуют аномалиям, близкие к 0 – нормальным наблюдениям.

Для каждого IP-адреса источника в скользящем временном окне длительностью 60 секунд извлекается вектор из 12 поведенческих признаков, описывающих активность узла: количество пакетов и байт в секунду, средний размер пакета, число уникальных портов и IP-адресов назначения, энтропия распределения портов, доля SYN-пакетов, среднее время между пакетами, энтропия содержимого и др.

Для объединения результатов необходимо, чтобы обе оценки находились в единой шкале. Сигнатурный модуль не возвращает бинарный ответ – вместо этого формируется вещественная оценка $s_{stat} \in [0, 1]$, вычисляемая как максимальный вес срабатывания сигнатуры с учетом степени совпадения (длины совпавшего паттерна, специфичности регулярного выражения) и дополнительных статистических признаков (частота срабатывания за последние 60 секунд, число уникальных правил, сработавших на данный источник). Модель машинного обучения по формуле (1) возвращает оценку $s_{ml} \in [0, 1]$ в той же шкале. Итоговая оценка угрозы формируется как взвешенное среднее двух нормализованных оценок:

$$s = \alpha \cdot s_{stat} + \beta \cdot s_{ml}, \quad \alpha + \beta = 1, \quad (2)$$

где α и β – весовые коэффициенты. В работе используются значения $\alpha = 0,4$, $\beta = 0,6$, полученные предварительной настройкой на обучающей подвыборке; больший вес модели машинного обучения обусловлен ее способностью выявлять распределенные атаки, слабо представленные в сигнатурной базе. Приведение обеих оценок к одному вещественному диапазону $[0, 1]$ обеспечивает корректность их взвешенного сложения. Решение о блокировке пакета принимается при $s \geq \tau$, где τ – порог, подбираемый эмпирически.

Для оценки качества классификации использовались стандартные метрики бинарной классификации: Precision (доля истинных срабатываний среди всех предсказанных как атакующие), Recall/TPR (доля обнаруженных атакующих источников среди всех реальных), FPR (доля ошибочно заблокированных легитимных источников) и F1-мера как гармоническое среднее Precision и Recall.

Результаты

Экспериментальное исследование предложенного алгоритма проведено на виртуализированном тестовом стенде, воспроизводящем типовую структуру КВС средней организации. Стенд включал 48 виртуальных машин в четырех VLAN-сегментах: 30 рабочих станций, 8 серверов приложений, 4 публичных сервиса в DMZ и 6 атакующих узлов под управлением Kali Linux. Центральный узел

защиты на Ubuntu Server 24.04 размещен на границе серверного сегмента в режиме прозрачного моста. Общая продолжительность эксперимента составила 8 часов непрерывной работы стенда, за которые было обработано порядка 4,82 млн пакетов, включающих 80,7% фоновое трафика и 19,3% атакующего. Обучение модели Isolation Forest выполнено на предварительно собранной выборке из 102 340 векторов поведенческих признаков, полученной на изолированной конфигурации стенда при отсутствии атакующей активности; среднее значение F1-меры на пятикратной стратифицированной кросс-валидации составило 0,91 со стандартным отклонением 0,02, что свидетельствует об отсутствии переобучения. Эксперимент проводился трижды с недельными интервалами для контроля воспроизводимости.

Таблица 1. Сравнение разработанного решения с Simple Python IDS и Snort 3

Параметр	Simple IDS	Разработанное решение	Snort 3
Архитектура	Монолит	Гибрид (2 модуля)	Модульная
Сигнатурный анализ	Да	Да	Да
ML-анализ	Нет	Да (Isolation Forest)	Нет
Пропускная способность, Мбит/с	60	620	940
TPR	0,70	0,942	0,951
FPR	0,12	0,000	0,014
Precision	0,65	1,000	0,968
F1-мера	0,67	0,970	0,960
Потребление RAM, МБ	120	248	512
Загрузка CPU-ядра, %	68	41	58

По результатам тестирования при оптимальном значении порога Score Fusion $\tau = 0,55$ система корректно классифицировала 278 атакующих источников из 295, при этом из 312 наблюдаемых источников фоновое трафика ни один не был ошибочно помечен как угроза. Итоговые значения метрик качества составили: Precision=1,00, TPR (Recall)=0,942, FPR=0, F1-мера=0,970. При консервативном значении порога $\tau = 0,60$, рекомендуемом для развертываний с повышенными требованиями к устойчивости, F1-мера снижается до 0,89 за счет увеличения доли пропусков атак с пограничными оценками угрозы. Для сравнения эффективности разработанного решения с существующими на том же стенде независимо развернуты учебный проект Simple Python IDS и промышленная система Snort 3 (версия 3.1.77) с набором правил Emerging Threats Open. Все три системы последовательно тестировались на одном и том же сгенерированном трафике. Результаты сравнения приведены в табл. 1.

Обсуждение

В отличие от учебного проекта Simple Python IDS, предложенный алгоритм использует двухуровневую схему принятия решения: быстрый сигнатурный этап блокирует трафик с явными признаками известных атак без накопления контекста, а поведенческий этап на базе Isolation Forest выявляет аномалии у источников с недостаточно выраженными сигнатурами. Такая схема позволяет достичь полного отсутствия ложных срабатываний ($FPR=0$) при сохранении высокой полноты обнаружения, что особенно важно при практической эксплуатации, где каждое ложное срабатывание требует ручного разбора оператором и увеличивает нагрузку на подразделение мониторинга.

Сравнение с промышленной системой Snort 3 показывает, что разработанное решение достигает сопоставимых значений F1-меры (0,970 против 0,960) и превосходит Snort 3 по Precision за счет поведенческого фильтра, отсеивающего нетипичные, но легитимные всплески активности. В то же время Snort 3 сохраняет преимущество по пропускной способности (940 против 620 Мбит/с) благодаря многолетней оптимизации движка и низкоуровневой реализации. Таким образом, предложенное решение не претендует на замену промышленных систем в высоконагруженных сегментах, однако оказывается конкурентоспособным в сетях малого и среднего размера при вдвое меньшем потреблении оперативной памяти.

Отдельного внимания заслуживает вопрос применимости алгоритма к шифрованному трафику, доля которого в современных КВС превышает половину. Из 12 используемых поведенческих признаков большинство (количество пакетов и байт в секунду, число уникальных портов и IP-адресов назначения, энтропия распределения портов, доля SYN-пакетов, среднее время между пакетами) извлекаются из заголовков сетевого и транспортного уровней и сохраняют информативность независимо от шифрования полезной нагрузки. Ограниченно применимой остается лишь энтропия содержимого пакетов, теряющая дискриминирующую способность на уже зашифрованных потоках. Это позволяет ожидать сохранения основных показателей качества обнаружения для шифрованного трафика, что составляет предмет отдельной экспериментальной проверки. Дополнительным направлением дальнейших исследований является анализ многоэтапных АРТ-сценариев с длинным временным горизонтом, выходящим за пределы текущего окна наблюдения в 60 секунд.

Заключение

В ходе исследования разработан гибридный алгоритм адаптивной фильтрации информационных потоков в корпоративных вычислительных сетях, объединяющий сигнатурный и поведенческий анализ через механизм Score Fusion с взвешенным усреднением. Предложенная методика тестирования на смешанном трафике с контролируемой долей семи классов атак и масштабный экспериментальный стенд из 48 виртуальных машин позволили получить показатели $Precision=1,000$, $TPR=0,942$, $FPR=0,000$ и $F1\text{-мера}=0,970$, сопоставимые с про-

мышленной системой Snort 3 при вдвое меньшем потреблении оперативной памяти.

В дальнейшем планируется исследование устойчивости алгоритма к шифрованному трафику и многоэтапным АРТ-атакам, реализация автоматической подстройки порога Score Fusion на основе онлайн-оценки распределения оценок угрозы, а также интеграция с промышленными SIEM-системами через стандартизованные интерфейсы.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Актуальные киберугрозы: III квартал 2025 года [Электронный ресурс]. – URL: <https://ar2025.ptsecurity.com/> (дата обращения: 15.05.2026).
2. Басыня Е. А., Равтович Ю. К. Разработка модуля системы обнаружения и предотвращения вторжений // Современные материалы, техника и технологии. – 2016. – № 2 (5). – С. 26–29.
3. Программная реализация интеллектуально-адаптивного управления // Вестник цифровых технологий. – 2021. – № 3. – С. 56–60.
4. Межсетевые экраны нового поколения: обзор технологий // Вестник информационных технологий. – 2020. – № 5 (3). – С. 34–40.
5. Подключение к Usergate DCFW [Электронный ресурс]. – URL: https://docs.usergate.com/podklyuchenie-k-usergate-dcfw_1601.html (дата обращения: 15.05.2026).
6. Алексеев А. В. Анализ систем обнаружения вторжений // Экономика и социум. – 2017. – № 6 (37). – С. 686–688.
7. Анчишкин А. П. Сравнительный анализ современных систем обнаружения вторжений // Colloquium-journal. – 2019. – № 26 (50). – С. 25–26.
8. Белова А. Л., Бородавкин Д. А. Сравнительный анализ систем обнаружения вторжений // Актуальные проблемы авиации и космонавтики. – 2016. – Т. 1. – С. 742–743.
9. Адаптивный метод фильтрации аномалий // Технологии информационной безопасности. – 2021. – № 2. – С. 45–49.
10. Liu F. T., Ting K. M., Zhou Z.-H. Isolation Forest // Proceedings of the 8th IEEE International Conference on Data Mining (ICDM '08). – Pisa, Italy : IEEE Computer Society, 2008. – P. 413–422. – DOI: 10.1109/ICDM.2008.17.

© Б. Ю. Якобсон, 2026