УДК: 004.492.3

Мухаметшина Р.А.
Уфимский университет науки и технологий, Уфа

Научный руководитель:
Сенцова А.Ю.
Уфимский университет науки и технологий, Уфа

ОПТИМИЗАЦИЯ СИГНАТУРНОГО АНАЛИЗА С ИСПОЛЬЗОВАНИЕМ БИТОВЫХ МАСОК: СРАВНИТЕЛЬНОЕ ИССЛЕДОВАНИЕ

Аннотация. В данной статье представлен комплексный анализ метода оптимизации сигнатурного анализа с использованием битовых масок. Исследование включает теоретическое обоснование использования метода, детальное описание экспериментальной реализации, сравнительный анализ производительности и практические рекомендации по внедрению.

Ключевые слова: сигнатурный анализ, битовые маски, оптимизация производительности, IDS, обнаружение атак, Python.

При работе с современными системами обнаружения вторжений (IDS) специалисты по информационной безопасности сталкиваются с фундаментальной проблемой обработки экспоненциально растущего объема сетевого трафика при необходимости сохранения высокой точности обнаружения угроз. В современных системах защиты основным инструментом выявления атак остаются сигнатурные методы анализа сетевого трафика. Их эффективность в обнаружении известных угроз подтверждена многолетней практикой применения, однако она напрямую зависит от актуальности и полноты используемой базы сигнатурных правил [1]. Эта проблема особенно актуальна для промышленных систем, где требования к скорости передачи информации зависят от спецификаций протокольных сред каждого уровня [2].

Традиционные методы сигнатурного анализа, основанные на последовательном поиске, демонстрируют временную сложность O(n×m), что становится критичным при обработке гигабитных потоков данных. В данной статье предлагается метод оптимизации сигнатурного анализа с использованием битовых масок, который позволит значительно ускорить процесс обнаружения известных угроз без потери точности.

Метод битовых масок реализует эффективный механизм предварительной фильтрации трафика за счет использования компактных бинарных шаблонов, где каждый бит маски соответствует определенному признаку в анализируемом пакете. Как следует из принципов работы битовых масок, значение каждого бита отвечает за наличие соответствующего элемента в

подмножестве [3], что позволяет быстро идентифицировать потенциально опасные пакеты путем битовых операций.

Актуальность исследования подтверждается тем, что большинство коммерческих и открытых IDS до сих пор используют устаревшие алгоритмы поиска сигнатур. В отличие от статистических методов, анализирующих аномалии поведения, сигнатурный анализ остается основным подходом в коммерческих продуктах благодаря своей простоте и эффективности, но нуждается в модернизации алгоритмов поиска [4].

Ключевые преимущества метода битовых масок перед классическим линейным поиском:

1. Снижение вычислительной нагрузки на 30–50 % за счет использования быстрых битовых операций вместо последовательного перебора.
2. Режим реального времени обработки благодаря эффективному представлению данных (использование целочисленных типов для хранения масок).
3. Гибкость адаптации к изменениям в сигнатурных базах, так как перебор всех масок размера n можно удобно осуществлять путем перебора всех чисел от 0 до 22^n – 1.

Для проведения экспериментов был разработан скрипт generate_data.py, создающий:

Файл сигнатур (signatures.txt) – 1000 случайных строк длиной 20 символов, имитирующих сигнатуры атак.

Наборы пакетов (packets_*.txt) – файлы с данными, включающие легитимный и вредоносный трафик в различных пропорциях.

Для воспроизводимости результатов использовался фиксированный сид (random.seed(42)).

Были разработаны два подхода:

Базовый алгоритм (simple_detection.py) – последовательный поиск каждой сигнатуры в пакетах.

Оптимизированный алгоритм (optimized_detection.py) – предварительная фильтрация пакетов с помощью битовых масок (длиной 8 бит), что сокращает количество полных проверок.

Для сравнения алгоритмов использовались:

Время выполнения (сек) – среднее значение по 4 запускам и загрузка CPU (%) – измерена с помощью psutil.

Эксперименты проводились в виртуальной среде со следующей конфигурацией:

Хост-система:

Процессор: Intel Core i7-9700 (8 ядер @ 3.0 GHz)

Память: 32 GB DDR4;

Хранилище: NVMe SSD 1 TB;

Виртуальная машина (Oracle VirtualBox 7.1.4):

Гостевая ОС: Ubuntu 25.04 (Plucky Puffin).

Особое внимание уделялось мониторингу ресурсов во время тестов.

Использование памяти не превышало 60 % от доступной, температура CPU поддерживалась в допустимых пределах.

Ниже представлены результаты по времени и по загрузки процессора при проведении экспериментов с базовым алгоритмом и с предложенным в статье оптимизированным алгоритмом, который использует метод битовых масок, где БА – базовый алгоритм, ОА – оптимизированный алгоритм.

Таблица 1

Результаты проведенных экспериментов

Сценарий	Пакеты	Атаки	БА (время, с)	БА (CPU, %)	ОА (время, с)	ОА (CPU, %)
1	10000	7650	0,5295	0,505	0,2432	0,23
2	786	460	0,04885	0,73	0,019575	0,35
3	200	20	0,013975	1,19	0,005075	0,705
4	16	1	0,001175	0,7225	0,000475	2,5625

Были сделаны следующие выводы:

Оптимизированный алгоритм в 2–2,5 раза быстрее базового при больших объемах данных.

Загрузка CPU снижается на 30–50 % для большинства сценариев.

В качестве заключения можно отметить следующее: применение битовых масок позволило значительно ускорить сигнатурный анализ без потери точности. Однако исследование выявило несколько направлений для дальнейшей работы:

Оптимизация длины масок – подбор оптимального размера для баланса между скоростью и точностью.

Динамические маски – адаптация масок под статистику трафика с помощью машинного обучения.

Гибридные методы – комбинация с другими алгоритмами (например, хешированием) для дальнейшего ускорения.

Результаты могут быть интегрированы в легковесные IDS для IoT-устройств или серверов с ограниченными ресурсами, где критична эффективность обработки трафика.

Дополнительные исследования: Тестирование на реальных сетевых данных (например, дампы трафика с атаками DDoS).

Сравнение с другими методами оптимизации (например, префиксными деревьями).

Таким образом, использование битовых масок демонстрирует высокий потенциал для оптимизации сигнатурного анализа в современных системах кибербезопасности.

Список использованных источников:

1. Гетьман А.И. Сравнение системы обнаружения вторжений на основе машинного обучения с сигнатурными средствами защиты информации / А.И. Гетьман, М. Н. Горюнов, А.Г. Мацкевич, Д.А. Рыболовлев // Труды ИСП РАН. 2022. Т. 34, № 5. С. 111–126.
2. Подтопельный В.В. Особенности формирования сигнатурных последовательностей для обнаружения сетевых атак в АСУТП / В.В. Подтопельный // Modern Science. 2020. № 12-3. С. 303–307.
3. Переборы с помощью битовых масок // Средняя школа № 9. URL: https://sch9.ru/news/
4. Актуальные вопросы выявления сетевых атак // InfoSecPortal.ru. URL: https://infosecportal.ru/stati/aktualnye-voprosy-vyyavleniya-setevyh-atak/

Mukhametshina R.A.
Ufa University of Science and Technology, Ufa

Scientific supervisor:
Sentsova A.U.
Ufa University of Science and Technology, Ufa

OPTIMIZATION OF SIGNATURE ANALYSIS USING BITMASKS: A COMPARATIVE STUDY

Abstract. This paper presents a comprehensive analysis of a method for optimizing signature analysis using bitmasks. The study includes a theoretical justification of the method, a detailed description of its experimental implementation, a comparative performance analysis, and practical recommendations for deployment.

Keywords: signature analysis, bitmasks, performance optimization, IDS, attack detection, Python.