УДК 004.056.5

Разработка интеллектуальных методов для обнаружения потенциальных угроз и уязвимостей компонентов распределенного приложения

Семыкин В. (ИТМО)
Научный руководитель – кандидат физико-математических наук, доцент Иванов С.Е. (ИТМО)

Введение. Современные распределенные приложения широко применяются в различных отраслях, включая облачные вычисления, финансовые технологии и критически важные информационные системы. Усложнение их архитектуры и рост числа атак на базы данных и сетевые сервисы требуют разработки эффективных методов обнаружения угроз. Традиционные системы безопасности часто не справляются с выявлением новых типов атак, таких как SQL-инъекции и эксплойты уязвимостей. Поэтому применение машинного обучения и анализа данных становится актуальным направлением для автоматизированного выявления и классификации угроз [1, 2, 3].

Основная часть. Разработка системы автоматического обнаружения угроз основывается на применении интеллектуальных методов анализа SQL-запросов. В исследовании решены следующие задачи:

1. Создание и обработка датасета. Для обучения моделей машинного обучения сгенерирован синтетический набор данных, содержащий безопасные и вредоносные SQL-запросы. Данные были векторизованы с помощью метода CountVectorizer и Word2Vec.
2. Выбор и обучение модели. Для классификации SQL-запросов использована логистическая регрессия, которая продемонстрировала точность 97,92% на тестовых данных. Оценка модели проводилась с использованием F1-меры, точности и полноты.
3. Разработка архитектуры системы. Разработан прототип, состоящий из трех ключевых компонентов:
   - Генератор данных, создающий SQL-запросы для обучения модели.
   - Анализатор SQL-запросов, использующий обученную модель для выявления потенциальных атак.
   - Модуль управления данными в разработанной архитектуре, обеспечивает логирование и интеграцию с брокерами сообщений (Kafka/RabbitMQ).
4. Экспериментальная проверка. Проведено тестирование системы в контролируемых условиях, продемонстрировавшее высокую эффективность детектирования угроз.

Используемый микросервисный подход с контейнеризацией (Docker, Kubernetes) позволяет системе быть гибкой и масштабируемой.

Выводы. Разработан и протестирован прототип системы для автоматического выявления уязвимостей в SQL-запросах с применением машинного обучения. Полученные результаты подтверждают эффективность предложенной модели, однако дальнейшие исследования будут направлены на интеграцию системы в реальные облачные инфраструктуры, улучшение алгоритмов обработки аномалий и разработку механизмов динамической адаптации к новым видам атак.

Список использованных источников:

1. Бирюков А. А. Информационная безопасность: защита и нападение. 3-е изд., перераб. и доп. М.: ДМК Пресс, 2023. 440 с.
2. Басс Л., Клементс П., Казман Р. Архитектура программного обеспечения: основы,

анализ, проектирование. 3-е изд., перераб. и доп. М.: Лори, 2021. 624 с.
3. Фаулер М. Архитектура корпоративных программных приложений. СПб.: Символ-Плюс, 2020. 432 с.