УДК 004

РАЗРАБОТКА СИСТЕМЫ АНАЛИЗА ИСХОДНОГО КОДА ВРЕДОНОСНЫХ ANDROID-ПРИЛОЖЕНИЙ НА ОСНОВЕ СВЕРТОЧНОЙ СЕТИ С ГРАФАМИ

Донг Суан Тхань (Национальный исследовательский университет ИТМО)
Чан Дык Мань (Национальный исследовательский университет ИТМО)
Нгуен Хоанг Хиеп (Национальный исследовательский университет ИТМО)
Научный руководитель– доцент, кандидат технических наук Левко И.В. (Национальный исследовательский университет ИТМО)

Аннотация. Android является наиболее широко используемой мобильной операционной системой с 73,0% доли рынка смартфонов в ноябре 2021 года. Однако открытость операционной системы Android не только обеспечивает удобство для пользователей, но и ведет к угрозе атаки со стороны большого количества вредоносных приложений.

Введение. Традиционные решения для обнаружения вредоносных приложений на основе сигнатур, которые основаны на анализе и сравнении сигнатур атак вредоносного ПО со списком предварительно идентифицированных сигнатур. Однако этот метод не может эффективно обнаруживать неизвестные варианты вредоносных программ, такие как вредоносные программы нулевого дня.

Основная часть. Граф вызовов функций Android (FCG) состоят из набора функций и их межпроцедурных вызовов, которые можно извлечь из файлов пакета приложений Android (APK) и представить в виде графика. Соответствующие структуры графов можно использовать для обнаружения вредоносных программ Android на основе изучения представления графа. Мы пытаемся извлечь наиболее полезные атрибуты и внедрить алгоритм классификации графа для выявления вредоносных приложений. Метод использует новейшие алгоритмы на графах, такие как Graph Convolution Network, Graph Attention Network, GraphSage,..

Выводы. На основе экспериментальных результатов мы достигли точности 96%, а шкала F1 достигла 0,95 на наборе данных из 10 000 вредоносных приложений. Исследования по классификации 5 самых популярных типов вредоносных программ дают хорошие результаты наряду с более низким временем выполнения алгоритма, чем у других методов использования глубокого обучения (из-за сложности алгоритма на нижнем графике). В будущем мы хотим разработать более эффективный алгоритм графа, основанный на самоконтроле, который доказал свою эффективность в других областях глубокого обучения.

Донг Суан Тхань (автор)

Левко И.В. (научный руководитель)