Выпускная квалификационная работа на степень бакалавра

Автоматизированная классификация функционального назначения программного обеспечения по исходным кодам с использованием нечетких хэш-функций

Задверняк Яна Анатольевна

Научный руководитель: к.ф.-м.н, доцент Нестеренко В.А.

Направление подготовки: 02.03.02 - Фундаментальная информатика и информационные технологии

Цели и задачи

Целью работы является исследование возможности использования нейронных сетей для классификации исходных кодов вредоносных программ.

Задачи

- анализ существующих решений в сфере классификации программ с использованием нейросетей
- сбор обучающей выборки из исходных кодов вредоносных программ разных классов
- обучение нейронной сети на различных метриках абстрактных синтаксических деревьев
- анализ результатов обучения

Предметная область

Сейчас широко распространены вредоносные скрипты на JavaScript.
Написанный на JavaScript исходный код легко обфусцируется, поэтому его сложно обнаруживать.
Существует два основных метода обнаружения:

- Статический - неустойчив к обфускации
- Поведенческий - требует значительных ресурсов и много времени

Предлагаемый метод обнаружения

В данной работе будем рассматривать метод классификации программ по их АСД, поскольку они не меняются от обфускации.
Проверим гипотезу, что разные классы вредоносных программ имеют общие черты в своих АСД.
Это позволит обнаруживать ранее неизвестные вредоносные программы.

Использование нейронных сетей

Нейронные сети очень популярны для решения задач NLP, в том числе для решения задач, связанных с анализом исходного текста программ:

- Самопризнанный технический долг
- Поиск потенциальных уязвимостей
- Поиск кода с уязвимостями
- Определение JavaScript-атак
- Генерация фрагментов кода

Таким образом, нейросети широко применяются для анализа и классификации исходных текстов программ. Используем их для классификации программ по их АСТ.

Экспериментальные данные

Использована обучающая выборка из исходных текстов программ двух классов:

- Вредоносные криптомайнеры
- Кейлоггеры, отслеживающие ввод пользователя и отправляющие их на сервер злоумышленника

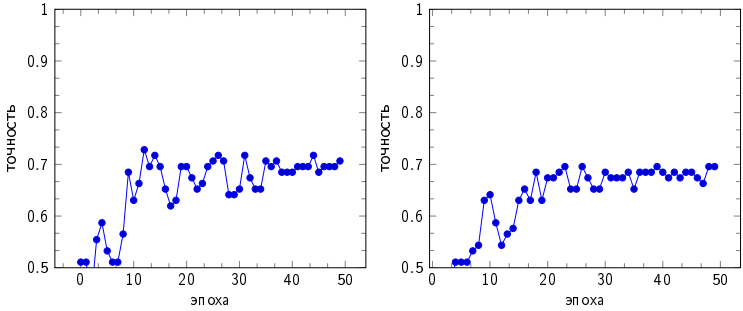
Парсинг и обучение модели

Для примера в эксперименте использована библиотека TensorFlow.
Исходные тексты обучающей выборки обрабатывались в несколько этапов:

- Парсинг в AST
- Выделение метрик из AST
- Нормализация данных методом min-max

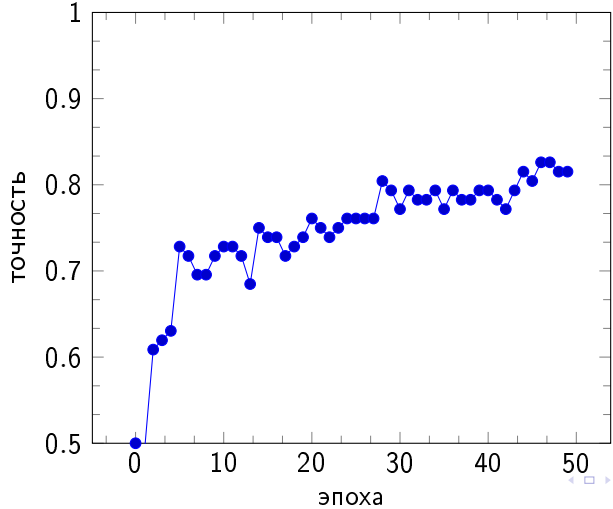
Парсинг и обучение модели

Графики обучения моделей для метрик количества узлов Identifier, CallExpression, ExpressionStatement, AssignmentExpression (слева) и общего количества узлов с количеством узлов Identifier, VariableDeclaration, CallExpression (справа)



Парсинг и обучение модели

График обучения модели для метрик количества узлов VariableDeclaration, CallExpression, Literal, AssignmentExpression



Приложение

Создано веб-приложение, демонстрирующее использование обученной модели на сервере для произвольного JavaScript-файла, отправленного с клиента.

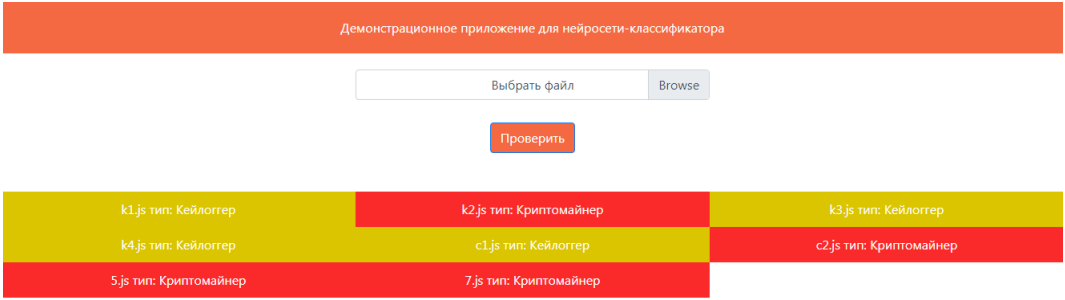


Рис.: Пример работы нейросети в демонстрационном приложении

Результаты

В итоге можно сделать вывод, что абстрактные синтаксические деревья могут использоваться для классификации вредоносных программ по их назначению. Получены следующие результаты:

- проанализированы существующие решения в сфере классификации программ с использованием нейросетей
- на основании собранной выборки обучена модель
- создано приложение, демонстрирующее использование обученной модели

Код, используемый при эксперименте и код демонстрационного приложения размещены по ссылке https://github.com/DasIgnis/AstNeuralNetworkClassify