УДК 005.04

АНАЛИЗ СИСТЕМЫ ДИНАМИЧЕСКОГО ОБНАРУЖЕНИЯ ВРЕДОНОСНЫХ ПРОГРАММ НА ОСНОВЕ РЕКУРРЕНТНЫХ И СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Севостьянов Г.Ю. (Национальный Исследовательский Университет ИТМО)

Научный руководитель - кандидат технических наук, доцент Коржук В.М. (Национальный Исследовательский Университет ИТМО)

Цель исследования. Проверка предложенной системы на работоспособность согласно методу оценки точности ROC AUC на статистически значимой выборке образцов безвредного и вредоносного ПО.

В данном докладе описана предлагаемая к анализу система обнаружения вредоносных программ на основе нейронных сетей, детали её построения и функционирования, и ожидаемые результаты работы.

Введение. С постоянным развитием вредоносных программ в сети интернет и практической невозможности эвристического анализа каждого метода их работы существует необходимость в существовании автоматизированных систем их динамического обнаружения, основанных на их чертах.

Нейронные сети, в теории, позволяют построение подобных систем, и один из методов их построения рассматривается в этой работе.

Основная часть. В статье “Malware detection with deep neural network using process behaviour” Тобияма и др. предлагают систему динамического обнаружения вредоносных программ состоящей из рекурентной и сверточной нейронных сетей. Идея состоит в том, чтобы обрабатывать логи запросов программ как натуральный язык, и на основе полученных черт языка делать вывод о принадлежности программы к вредоносным.

Рекуррентная сеть состоит из входного слоя, скрытого слоя, двух слоев долгой краткосрочной памяти и слоя вывода.

Сверточная сеть состоит из входного слоя, двух слоев пуллинга-свертки, полносвязанного слоя и слоя вывода.

Процесс работы системы представлен дальше:
1. Запись логов запросов программ в реальном времени выполняется с использованием системного крюка на уровне пользователя с помощью Cuckoo Sandbox.

2. Рекуррентная сеть строит изображение на основе выделенных характерных черт отдельной программы по полученным логам.

3. Сверточная сеть анализирует изображение характерных черт и классифицирует соответствующую программу.

Выводы. Итогами выполнения данной работы и тестирования написанной системы предполагается один из 3х возможных результатов: система неработоспособна с точностью выполнения ниже ранее представленной; система неработоспособна ввиду её склонности к “запоминанию” вместо анализа; система работоспособна с точностью выполнения соответствующей или превышающей ранее представленную.

Севостьянов Г.Ю. (автор) Подпись

Коржук В.М. (научный руководитель) Подпись