УДК 004.891.3

Разработка алгоритма обнаружения ботнета в корпоративных сетях с использованием частотных характеристик трафика

Золотников И. К. (ИТМО)
Научный руководитель — аспирант факультета безопасности информационных технологий Бучаев А. Я. (ИТМО)

Введение. Ботнеты — компьютерная сеть, состоящая из некоторого количества устройств с доступом в сеть, с запущенным автономным программным обеспечением [4]. Ботом в составе такой сети является сам компьютер с вредоносным ПО, дающим возможность злоумышленнику использовать ресурсы зараженного устройства в своих целях. Ботнеты используются для атак на сервера, подбора паролей на удалённой машине или рассылки спама, хищения данных пользователей [6]. Главные характеристики ботнета: распространенность и децентрализованность. Такие сети имеют зачастую всего один удаленный хост, который лишь посылает команды на заражённые устройства и отследить его довольно сложно. Заражение может произойти в любое время и из любого источника: интернет, съемные носители, почта. После заражения код ботнета распространяется в сети самостоятельно, так как большинство современных ботнетов поддерживает ретрансляцию себя через подконтрольные устройства [5], что позволяет охватить существенно больший объем устройств. Всё это делает ботнеты серьезной угрозой для информационной безопасности организаций, так как наиболее часто именно крупные корпорации подвергаются атакам со стороны киберпреступников, которые занимаются проникновением в систему с целью ее последующей эксплуатации.

Основная часть. Основной задачей борьбы с ботнетами является их своевременное обнаружение и изоляция. Для этого необходимо анализировать присутствующие в сети пакеты в реальном времени с целью обнаружения недекларированной активности. Из-за специфики свойств ботнета задача обнаружения необычного трафика усложняется, так как любой из пораженных хостов может отправлять случайные данные в случайное время по случайному адресу, что делает невозможным простое алгоритмическое решение данной задачи. Развитие и массовое использование продвинутых средств туннелирования и шифрования [2] сделало определение необычных пакетов намного более сложной задачей. Однако, активное развитие алгоритмов машинного обучения позволило создать удобные инструменты для решения задач по анализу огромных объемов данных. Поэтому, было принято решение воспользоваться алгоритмами машинного обучения. Они позволяют достичь высоких точности и скорости при анализе трафика, что важно при борьбе с ботнетом. В нашей работе алгоритм решения состоит из двух частей. Первая часть состоит из сбора пакетов из сети и преобразования их по ряду признаков в набор данных для обучения алгоритма — вектор чисел, содержащий характеристики пакетов. Вторая часть — подать эти данные на вход автокодировщика, после чего решить задачу классификации. Автокодировщик [3] является достаточно удобным инструментом для решения задачи детектирования аномалий. Используемая модель тренируется на сжатие и восстановление нормальных данных, поэтому восстановленные аномальные данные будут существенно отличаться от остальных, что позволит достаточно точно их детектировать. Данное решение актуально тем, что в современных средствах анализа трафика и обнаружения ботнетов алгоритмы машинного обучения пока не так широко используются, как и отсутствуют модули преобразования трафика в конечный вектор чисел. Текущие решения в данной области пользуются заранее подобранным вектором чисел, что делает такие алгоритмы недостаточно эффективными. Данные для анализа трафика собирались из корпоративной сети Университета ИТМО, на что было получено согласие технических специалистов факультета безопасности информационных технологий. Данная сеть характеризуется

высокой разветвленностью, большим количеством устройств в ней и стабильным потоком данных. Для анализа были выбраны следующие характеристики пакета: адреса отправителя и получателя, количество соединений, продолжительность соединения, частота установки соединения, время установки соединения и содержимое пакета. Также интересным параметром является перечень используемых для соединения портов.

Выводы. В ходе работы были проанализированы собранные данные и написана модель автокодировщика, подобрано количество слоев и нейронов для ее оптимальной работы. После этого модель была обучена на собранных данных. В дальнейшем планируется провести тестирование разработанного решения в условиях реальной сети с целью определения его эффективности.

Список использованных источников:

1. Taboada-Crispi A. Anomaly detection in medical image analysis. / A. Taboada-Crispi H. Sahli, D. Hernandez-Pacheco [et al.] // *Handbook of research on advanced techniques in diagnostic imaging and biomedical applications*. — Ukraine: IGI Global, 2009. — pp. 426–446. — URL: https://www.researchgate.net/publication/202974904_Anomaly_Detection_in_Medical_Image_Analysis
2. Булатов А. Е. Аналитический обзор тенденции развития архитектуры сетей передачи данных //T-Comm-Телекоммуникации и Транспорт. – 2011. – №. 7. – С. 31-34. — URL: https://cyberleninka.ru/article/n/analiticheskiy-obzor-tendentsii-razvitiya-arhitektury-setey-peredachi-dannyh
3. Жерон О. Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow ШШ. – 2018. — URL: https://library.bsuir.by/m/12_101945_1_133084.pdf
4. Ковалевский А. И. Ботнет сети и их трафик //Естественные и математические науки в современном мире. – 2014. – №. 25. – С. 35-39. — URL: https://cyberleninka.ru/article/n/botnet-seti-i-ih-trafik
5. Комаров А. А., Назаров А. Н. Функциональные требования к системе обнаружения и противодействия ботнет-атакам на корпоративные сети //Техника средств связи. Серия: Техника телевидения. – 2013. – №. 1. – С. 140-151. — URL: https://elibrary.ru/item.asp?id=20711019
6. Косенко М. Ю., Мельников А. В. Вопросы обеспечения защиты информационных систем от ботнет атак //Вопросы кибербезопасности. – 2016. – №. 4 (17). – С. 20-28. — URL: https://cyberleninka.ru/article/n/voprosy-obespecheniya-zaschity-informatsionnyh-sistem-ot-botnet-atak

Автор: Золотников И. К.
Научный руководитель: Бучаев А. Я.