УДК 004.056.53

ВОЗМОЖНОСТИ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ DDOS АТАК

Шаронова Е. И., магистрант гр. 267241

Белорусский государственный университет информатики и радиоэлектроники

г. Минск, Республика Беларусь

Петров С. Н.. – канд. техн. наук

Аннотация. Показаны возможности методов машинного обучения для обнаружения признаков атак типа DDoS.

Ключевые слова. Машинное обучение, ботнет, DDoS-атака, нейронные сети, сетевой поток, датасет, библиотеки машинного обучения.

DDos-атаки нацелены на сайты и онлайн-сервисы. Смысл такой атаки заключается в том, чтобы заблокировать сеть или сервер чрезмерным трафиком. Эффективность достигается за счет использования нескольких скомпрометированных систем в качестве источников атакующего трафика. DDoS-атаки делятся на различные подкатегории в зависимости от уровня сетевого подключения, которое они пытаются атаковать (согласно модели OSI). Выделяют следующие категории: SYN Flood, UDP Flood, MSSQL, LDAP, Portmap и т.п.

Использование нейронных сетей для детектирования DDoS-атак (атак распределенного отказа в обслуживании) является одним из подходов в области кибербезопасности, который может помочь в выявлении подозрительной активности, связанной с DDoS-атаками. Ниже приведены основные шаги, связанные с использованием нейронных сетей для детектирования DDoS-атак [1].

Подготовка данных: Для тренировки нейронной сети необходимо иметь подготовленный набор данных, содержащий размеченные примеры DDoS-атак и нормальной сетевой активности. Этот набор данных будет использоваться для обучения нейронной сети, чтобы она могла "научиться" распознавать характерные признаки DDoS-атак [2].

Выбор архитектуры нейронной сети: В зависимости от конкретной задачи и объема данных можно выбрать подходящую архитектуру нейронной сети. Например, сверточные нейронные сети (Convolutional Neural Networks, CNN) могут быть эффективны для анализа сетевого трафика, а рекуррентные нейронные сети (Recurrent Neural Networks, RNN) могут быть полезны для анализа последовательностей событий.

Обучение нейронной сети: Набор данных с размеченными примерами DDoS-атак и нормальной сетевой активности используется для обучения нейронной сети. Нейронная сеть "изучает" характерные признаки DDoS-атак на основе предоставленных данных.

Тестирование и настройка нейронной сети: После обучения нейронной сети необходимо протестировать ее на отдельном наборе данных, который не был использован в процессе обучения. Это поможет оценить точность и эффективность нейронной сети и внести дополнительные настройки при необходимости [3].

Интеграция в систему обнаружения атак: После успешного тестирования и настройки нейронной сети ее можно интегрировать в систему обнаружения атак, такую как система SIEM (Security Information and Event Management), для автоматического обнаружения DDoS-атак в реальном времени [4].

Рассмотрим более подробно некоторые пункты этого списка.

Для обучения нейронных сетей на обнаружение DDoS-атак необходимо использовать подготовленные датасеты, содержащие размеченные примеры DDoS-атак и нормальной сетевой активности. Ниже приведены некоторые из популярных датасетов, которые можно использовать для этой цели:

DARPA Intrusion Detection Data Sets: Датасеты, предоставленные DARPA (Defense Advanced Research Projects Agency), включают большой объем данных с размеченными примерами различных типов атак, включая DDoS-атаки. Эти датасеты широко используются в исследованиях в области кибербезопасности и могут быть полезны для обучения нейронных сетей на обнаружение DDoS-атак.

CICDDoS2019: Датасет CICDDoS2019, разработанный в Канадском институте компьютерных исследований (Canadian Institute for Cybersecurity), содержит данные о сетевом трафике, сгенерированном DDoS-атаками разных типов, а также нормальной сетевой активности.

UNSW-NB15: Датасет UNSW-NB15, разработанный в Университете Нового Южного Уэльса (University of New South Wales), включает данные о сетевом трафике с размеченными примерами атак, включая DDoS-атаки.

А также UNSW_NB15 [5], Kitsune Network Attack, Network Intrusion Detection [6].

Это лишь некоторые из возможных датасетов, которые можно использовать для обучения нейронных сетей на обнаружение DDoS-атак. Важно выбирать подходящий датасет в зависимости от конкретных требований и целей исследования.

Выбор архитектуры нейронной сети для обнаружения DDoS-атак зависит от множества факторов, таких как доступность данных, требования к производительности, предполагаемые типы DDoS-атак и другие особенности конкретной задачи. Вот несколько распространенных архитектур нейронных сетей, которые могут быть использованы для обнаружения DDoS-атак:

Сверточные нейронные сети (Convolutional Neural Networks, CNNs) широко применяются в обработке изображений, но также могут быть использованы для обработки сетевого трафика, включая обнаружение DDoS-атак. Они могут автоматически извлекать важные признаки из входных данных, таких как пакеты или сетевые потоки, и использовать их для классификации на атаки или нормальную активность.

Рекуррентные нейронные сети (Recurrent Neural Networks, RNNs) подходят для анализа последовательных данных, таких как сетевой трафик, и могут быть использованы для обнаружения DDoS-атак, учитывая последовательную природу сетевого трафика. Например, LSTM (Long Short-Term Memory) и GRU (Gated Recurrent Unit) – это популярные типы RNN, которые могут быть использованы для обработки временных рядов, включая сетевой трафик.

Автоэнкодеры (Autoencoders) это нейронные сети, которые могут быть использованы для извлечения важных признаков из входных данных и их реконструкции. Они могут быть использованы для обнаружения аномалий, включая DDoS-атаки, путем обучения на нормальной активности и выявления отклонений от нее.

Сочетание различных архитектур: В ряде случаев, комбинирование нескольких типов нейронных сетей может привести к лучшим результатам. Например, сочетание CNN и RNN может учитывать как пространственные, так и временные признаки сетевого трафика.

Также используются глубокие ансамбли (Deep Ensembles): Это архитектуры, состоящие из нескольких нейронных сетей, которые обучаются вместе. Они могут использоваться для повышения точности и устойчивости модели на обнаружение атак.

Важно проводить эксперименты с различными архитектурами и настраивать их параметры для конкретной задачи обнаружения DDoS-атак, чтобы достичь наилучших результатов.

Обнаружение IoT-трафика, используемого для DDoS-атак, является важной задачей в области кибербезопасности. Для решения этой задачи можно использовать различные методы, включая машинное обучение и анализ трафика. Один из подходов заключается в использовании алгоритмов машинного обучения, таких как классификация на основе дерева решений или нейронные сети, для обнаружения аномалий в трафике. Для этого необходимо собрать достаточное количество данных о трафике IoT-устройств и обучить модель на этом наборе данных. Другой подход заключается в анализе самого трафика с использованием методов, таких как анализ частоты и длительности пакетов, анализ наборов данных, таких как размер пакета, время задержки и количество пакетов в секунду, и другие методы. Эти методы могут быть использованы для определения аномалий в трафике IoT-устройств. Также можно использовать комбинацию методов машинного обучения и анализа трафика для обнаружения DDoS-атак, использующих IoT-устройства.

Для обнаружения IoT DDoS-трафика можно использовать различные датасеты, содержащие записи трафика, собранные в реальных условиях. Некоторые из популярных датасетов для обнаружения IoT DDoS-трафика включают:

IoT-23 – датасет, разработанный на основе DDoS-атак на устройства Интернета вещей (IoT) в реальных сетях. Он содержит записи трафика с 23 различных устройств IoT, включая камеры видеонаблюдения, медиаплееры, маршрутизаторы, принтеры и другие, и включает различные типы DDoS-атак [7].

Android Mischief Dataset – набор данных сетевого трафика с мобильных телефонов, зараженных троянами удаленного доступа Android [8].

Для обучения моделей распознавать DDoS атаки можно использовать различные языки программирования и фреймворки, в зависимости от предпочтений и требований проекта. Некоторые из популярных языков программирования и фреймворков для обучения моделей на обнаружение DDoS атак включают:

Python – является одним из наиболее популярных языков программирования для машинного обучения и имеет множество библиотек и фреймворков, таких как TensorFlow, Keras, PyTorch, scikit-learn, которые предоставляют мощные инструменты для создания и обучения моделей машинного обучения для обнаружения DDoS атак.

R – язык программирования и окружение для статистической обработки данных, который также может быть использован для обучения моделей машинного обучения. R предоставляет богатый набор библиотек, таких как caret, randomForest, xgboost и другие, которые могут быть использованы для создания моделей обнаружения DDoS атак.

TensorFlow – открытая библиотека машинного обучения, разработанная Google, которая предоставляет мощные инструменты для создания и обучения моделей машинного обучения, включая обнаружение DDoS атак. TensorFlow имеет широкий выбор предварительно обученных моделей и также позволяет создавать собственные модели с использованием глубоких нейронных сетей.

Scikit-learn – библиотека машинного обучения для языка программирования Python, которая предоставляет множество алгоритмов машинного обучения, таких как Decision Trees, Random Forest, SVM, и другие, которые могут быть использованы для создания моделей обнаружения DDoS атак.

Keras – высокоуровневый фреймворк машинного обучения для Python, который предоставляет простой и интуитивно понятный интерфейс для созд

PyTorch – фреймворк машинного обучения, разработанный Facebook, и также имеет множество инструментов и библиотек для обучения моделей на данных о DDoS атаках.

Машинное обучение используется для обнаружения DDoS атак во многих продуктах, включая:

– SIEM системы (Security Information and Event Management)
– Системы обнаружения вторжений (Intrusion Detection Systems)
– Системы защиты от DDoS атак (DDoS Protection Systems)
– Брандмауэры и системы безопасности сети
– Кластеры веб-серверов
– CDN (Content Delivery Network)

В этих продуктах используются различные методы машинного обучения, включая классические алгоритмы, такие как машины опорных векторов (SVM) и наивные байесовские классификаторы, а также более сложные алгоритмы, такие как глубокие нейронные сети и алгоритмы кластеризации данных.

Системы обнаружения вторжений (Intrusion Detection Systems, IDS) могут использовать машинное обучение для обнаружения DDoS атак, анализируя сетевой трафик или лог-файлы на наличие аномального поведения или характеристик, связанных с DDoS атаками. Системы облачной безопасности (Cloud Security) могут использовать машинное обучение для обнаружения DDoS атак, анализируя сетевой трафик и активность в облачной среде на предмет аномалий и признаков DDoS атак. Системы анализа журналов и лог-файлов (Log Analysis Systems) могут использовать машинное обучение для обнаружения DDoS атак, анализируя лог-файлы на наличие аномальных событий, включая признаки DDoS атак. Специализированные продукты для обнаружения DDoS атак разработанные специально для обнаружения DDoS атак с использованием машинного обучения, такие как Radware DefensePro, Arbor Networks APS, F5 Networks Silverline.

Сделаем краткое обобщение изложенных выше фактов.

Машинное обучение широко используется для обнаружения DDoS атак в различных продуктах и системах. Наиболее распространенными языками программирования и фреймворками для обучения моделей являются Python, TensorFlow, Keras, PyTorch и scikit-learn. В качестве датасетов используются как синтетические данные, так и данные с реальных атак. Для обнаружения IoT-сетей, используемых для DDoS атак, могут быть использованы датасеты, содержащие данные трафика с устройств IoT. Архитектуры нейронных сетей для обнаружения DDoS атак могут быть различными, включая сверточные нейронные сети, рекуррентные нейронные сети и комбинации из них. В качестве входных данных для моделей могут использоваться параметры сетевого трафика, такие как размер пакетов, частота запросов, распределение IP-адресов и протоколов. Продукты, использующие машинное обучение для обнаружения DDoS атак, включают в себя коммерческие решения от компаний, а также открытые проекты.

В целом, машинное обучение дает новые возможности для обнаружения DDoS атак, позволяя автоматически анализировать и выявлять необычное поведение в сети, что может быть связано с атакой. Однако, необходимо помнить, что машинное обучение не является универсальным решением для всех сценариев обнаружения DDoS атак и может требовать дополнительной настройки и адаптации для каждого конкретного случая.

Список использованных источников:

1. DDoS-атаки и электронная коммерция: современные подходы к защите [Электронный ресурс] / 1-С Битрикс / / Сайт Хабрахабр. - Режим доступа: https://habrahabr.ru/ company/bitrix/blog/267947

2. Шелухин О.И., Ванюшина А.В., Габисова М.Е. Фильтрация нежелательных приложений интернет трафика с использованием алгоритма классификации Random Forest. Вопросы кибербезопасности, № 2 (26), 2018 г., стр. 44-51

3. Artifcial intelligence and machine learning) / Niklas Kühl, Max Schemmer, Marc Goutier, Gerhard Satzger.] // Electronic Marketsи, https://link.springer.com/article/10.1007/s12525-022-00598-0 .

4. Котов, В. Д. Современное состояние проблемы обнаружения сетевых вторжений / В.Д. Котов, В.И. Васильев / / Вестник УГАТУ. - 2012. - Т. 16, № 3. - С. 198-204.

5. UNSW_NB15 – Kaggle [Электронный ресурс] – Режим доступа: https://www.kaggle.com/mrwellsdavid/unsw-nb15/. – Дата доступа: 02.04.2023

6. Network Intrusion Detection – Kaggle [Электронный ресурс] – Режим доступа: https://www.kaggle.com/sampadab17/network-intrusion-detection/. – Дата доступа: 02.04.2023

7. Aposemat IoT-23 – Stratosphereips [Электронный ресурс] – Режим доступа: https://www.stratosphereips.org/datasets-iot23. – Дата доступа: 02.04.2023

8. Android-mischief-dataset – Stratosphereips [Электронный ресурс] – Режим доступа: https://www.stratosphereips.org/android-mischief-dataset. – Дата доступа: 02.04.2023

UDC 004.056.53

MACHINE LEARNING CAPABILITIES FOR DETECTING DDOS ATTACKS

Sharonava E.I.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Petrov S.N. – PhD in Technical Sciences

Annotation. The possibilities of machine learning methods for detecting signs of DDoS attacks are shown.

Keywords. Machine learning, botnet, DDoS attack, neural networks, network flow, dataset, machine learning libraries.