ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ АНАЛИЗА СЕТЕВОГО ТРАФИКА

В.В. Денисенко, *канд. техн. наук, доцент*
А.С. Ященко, *магистрант*
Воронежский государственный университет инженерных технологий
(Россия, г. Воронеж)

DOI:10.24412/2500-1000-2023-1-1-19-22

Аннотация. *В статье рассматриваются способы применения искусственного интеллекта для анализа сетевого трафика. Преимущество использования глубокого машинного обучения для анализа сетевого трафика заключается в его способности обрабатывать большие объемы данных и выявлять закономерности. Применение нейронных сетей при анализе безопасности сети включает в себя такие этапы работы, как: обнаружение аномалий, классификация трафика и идентификация атаки. Искусственный интеллект и нейросети в частности способны обучаться, используя данные для развития из сетевого трафика, опираясь на закономерность некоторых выявленных признаков. Процесс глубокого обучения необходимо контролировать, так как если этих данных для обучения будет недостаточно, то в итоге отследить логику действий и причину некоторых выводов может быть проблематично.*

Ключевые слова: *информационная безопасность, искусственный интеллект, машинное обучение, компьютерные сети, нейронные сети, анализ сетевого трафика.*

Искусственный интеллект становится все более важным инструментом для анализа и понимания сетевого трафика. С быстрым ростом подключенных устройств и Интернета объем сетевого трафика резко увеличился, что затрудняет традиционные методы обработки огромного объема данных. ИИ, особенно глубокое обучение, предлагает мощное решение для обработки и анализа этого огромного объема данных в режиме реального времени.

Глубокое обучение, подмножество машинного обучения, включает в себя обучение нейронных сетей на больших объемах данных, чтобы выявлять закономерности и делать прогнозы. Эти нейронные сети можно использовать для различных задач, таких как распознавание изображений и речи, обработка естественного языка и анализ сетевого трафика.

Нейронные сети могут использоваться в информационной безопасности для различных целей, включая обнаружение вторжений, анализ безопасности сети, идентификацию и аутентификацию пользователей, а также классификацию и анализ трафика. Они могут быть обучены обнаруживать аномалии и атаки, которые могут быть сложными для обнаружения с помощью традиционных методов.

Анализ безопасности сети с использованием нейронных сетей может включать в себя несколько этапов, таких как:

- Обнаружение аномалий: Нейронные сети могут быть обучены обнаруживать аномальное поведение в сети, такое как неожиданное сетевое трафик, которое может быть связано с атакой.
- Классификация трафика: Нейронные сети могут использоваться для классификации трафика на безопасный и небезопасный, помогая разделить нормальный трафик от атакующего.
- Идентификация атаки: Нейронные сети могут использоваться для идентификации различных типов атак, например, DDoS-атак, сканирование портов, инъекции SQL [1].

Важно отметить, что классификация трафика сети является необходимым элементом в системах информационной безопасности и позволяет обеспечить безопасность информации в сети, а также улучшить ее производительность. Первые заслуживающие внимания работы появились по научным меркам не так давно в

2016 году. «Deep packet inspection using convolutional neural networks" by A. G. Boteanu et al.» – эта статья описывает, как использовать сверточные нейронные сети для глубокого анализа пакетов и оценивает их эффективность по сравнению с традиционными методами [2]. "Network traffic classification using deep recurrent neural networks" by R. T. Souza et al. – эта статья представляет использование рекуррентных нейронных сетей для классификации сетевого трафика и оценивает их эффективность в сравнении с другими методами [3].

Одним из ключевых преимуществ использования глубокого обучения для анализа сетевого трафика является его способность обрабатывать большие объемы данных и выявлять закономерности, которые людям может быть трудно обнаружить. Это делает его хорошо подходящим для таких задач, как обнаружение вторжений, где он может выявлять аномалии в сетевом трафике, которые могут указывать на попытку кибератаки. Кроме того, глубокое обучение также можно использовать для классификации трафика, т. е. процесса идентификации различных типов сетевого трафика, например электронной почты, просмотра веб-страниц и потокового видео. Это может быть полезно для сетевых администраторов, которым необходимо понять, как используется их сеть, и определить любые потенциальные узкие места или уязвимости в системе безопасности.

Классификация трафика с использованием нейронных сетей может включать в себя следующие шаги:

1. Подготовка данных: необходимо подготовить данные для обучения и тестирования модели. Это может включать в себя преобразование данных сетевого трафика в признаки, которые могут быть использованы для классификации.

2. Обучение модели: используя подготовленные данные, модель нейронной сети может быть обучена классифицировать трафик как безопасный или небезопасный.

3. Тестирование модели: после обучения, модель должна быть протестирована на неизвестных данных, чтобы оценить ее точность и способность классифицировать новый трафик.

4. Использование модели в боевом режиме, после успешного тестирования [4].

К задачам классификации трафика сети можно отнести: определение нежелательного или атакующего трафика для блокировки или отправки в специальный аналитический модуль; определение приоритетного трафика для обеспечения его максимальной доступности; определение трафика, который требует дополнительной аутентификации или авторизации; определение аномального трафика для выявления попыток информационной инцидента или вторжения; оптимизация использования ресурсов сети для более эффективной работы; мониторинг и анализ трафика для поиска уязвимостей или недостатков в конфигурации сети.

Для анализа безопасности по другим направлениям необходимо собрать дополнительные данные для обучения и применить другие виды тестирования.

Рассмотрим подробнее задачу глубокого анализа пакетов с помощью нейросетей. Это метод использования глубоконаучных алгоритмов для анализа и классификации сетевого трафика. Он используется для обнаружения аномалий, вторжений и других нежелательных событий в сети.

В глубоком анализе пакетов с помощью нейросетей, данные сетевого трафика представляются в виде набора признаков, которые могут быть использованы для обучения нейронной сети. Например, признаки могут включать в себя информацию о протоколе, источнике и назначении, длине пакета и т.д. После обучения нейронная сеть может использоваться для классификации нового трафика как безопасный или небезопасный.

Глубокий анализ пакетов с помощью нейросетей может быть более точным и эффективным методом для обнаружения аномалий и атак в сети, чем традиционные методы, такие как сигнатурный анализ или правила базирующиеся на совпадениях. Так как нейронные сети могут автоматически извлекать признаки из данных и обучаться на основе новых данных, они могут быть более адаптивными и способными обнаруживать новые и неизвестные виды атак [5].

Однако также следует отметить, что глубокий анализ пакетов с помощью нейросетей также имеет некоторые ограничения. Например, он может столкнуться с проблемой переобучения, если недостаточно данных для обучения, и может быть сложно интерпретировать результаты и понять, как нейронная сеть пришла к своим предсказаниям. Также требуется более высокий уровень знаний и опыта в области машинного обучения и нейронных сетей для разработки и использования такой системы [6].

В целом, глубокий анализ пакетов с помощью нейросетей может быть полезным инструментом для обеспечения безопасности сети, но также требует внимания и особых усилий для его разработки и использования.

Стоит отметить, что использование моделей глубокого обучения для безопасности сети является активной и развивающейся областью исследования, и есть другие проекты и компании, которые работают над подобными решениями и были реализованы в реальных сценариях. Однако стоит отметить, что безопасность и эффективность этих систем может варьироваться, и перед тем, как внедрять их в производственную среду, необходимо тщательно оценить любое решение по безопасности.

Можно уже встретить довольно хорошо работающие и широко известные открытые проекты:

- "DeepPacket" – это открытый исходный код нейронной сети, разработанной для глубокого анализа пакетов сетевого трафика. Этот проект доступен на GitHub: https://github.com/sdnds-tw/DeepPacket

- "DeepFlow" – это открытый исходный код системы глубокого обучения для классификации сетевого трафика, разработанной в Университете Калифорнии в Беркли. Этот проект доступен на GitHub: https://github.com/caesar0301/deepflow

- "DeepIntrusion" – это открытый исходный код нейронной сети, разработанной для обнаружения вторжений в сети с использованием глубокого обучения. Этот проект доступен на GitHub: https://github.com/deepintrusion/deepintrusion

Можно предположить, что упомянутые выше проекты или подобные им используются в реальных системах безопасности сети, но также возможно, что они находятся в стадии исследования и разработки и еще не были внедрены в системы в производстве. Кроме того, стоит отметить, что использование модели в реальной системе требует более серьезного рассмотрения, чем сама модель, таких как производительность и масштабируемость модели, интеграция системы и т.д.

В заключение, глубокое обучение – это мощный инструмент для обработки и анализа сетевого трафика. Он хорошо подходит для таких задач, как обнаружение вторжений и классификация трафика, может обрабатывать большие объемы данных и адаптироваться к меняющимся условиям. Однако есть и такие проблемы, как отсутствие размеченных данных и интерпретируемость моделей глубокого обучения. Несмотря на эти проблемы, исследования в этой области продолжаются, и вполне вероятно, что глубокое обучение продолжит играть важную роль в анализе и понимании сетевого трафика в будущем.

Библиографический список

1. Скрыпников, А.В. Использование методов машинного обучения при решении задач информационной безопасности / А.В. Скрыпников, В.В. Денисенко, И.А. Саранов // Вестник Воронежского института ФСИН России. – 2020. – №4. – С. 69-73. – EDN MYVNUV.

2. Ботяну А.Г. Глубокая проверка пакетов с использованием сверточных нейронных сетей / Ботяну А.Г., Лотфоллахи Мохаммад и др. // Мягкие вычисления. – 2020. – С. 1-14.

3. Соуза Р.Т. Классификация сетевого трафика с использованием глубоких рекуррентных нейронных сетей / Р.Т. Соуза, Франческо Палмиери, Джанни Д'Анджело // Журнал сетевых и компьютерных приложений, том 173, 1 января 2021 г., 102890.

4. Дайнотти А. Проблемы и будущие направления в классификации трафика / А. Дайнотти, А. Пескейп, К.С. Клаффи // IEEE Network. – 2012. – Vol. 26, №1. – С. 35-40.

5. Кузнецова, А.В. Искусственный интеллект и информационная безопасность общества: монография / А.В. Кузнецова, С.И. Самыгин, М.В. Радионов; ред. П.С. Самыгин. – М.: Русайнс, 2016. – 117 с.

6. Маршалко, Г. Игры искусственного разума: безопасность систем машинного обучения // Информационная безопасность. – 2018. – №4. – С. 6-7.

APPLICATION OF ARTIFICIAL INTELLIGENCE FOR NETWORK TRAFFIC ANALYSIS

V.V. Denisenko, *Candidate of Technical Sciences, Associate Professor*
A.S. Yaschenko, *Graduate Student*
Voronezh State University of Engineering Technologies
(Russia, Voronezh)

Abstract. *The article discusses ways to use artificial intelligence to analyze network traffic. The advantage of using deep machine learning to analyze network traffic is its ability to process large amounts of data and discover patterns. The use of neural networks in the analysis of network security includes such stages of work as: anomaly detection, traffic classification and attack identification. Artificial intelligence and neural networks in particular are able to learn using data for development from network traffic, based on the pattern of some identified features. The deep learning process needs to be controlled, because if this data is not enough for training, then in the end it can be problematic to track the logic of actions and the reason for some conclusions.*

Keywords: *information security, artificial intelligence, machine learning, computer networks, neural networks, network traffic analysis.*