ОБЗОР СОВРЕМЕННЫХ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ В ОБЛАСТИ ОБНАРУЖЕНИЯ АНОМАЛИЙ В ИНФОРМАЦИОННЫХ СЕТЯХ

Драгунов А.Д., магистрант гр.467421

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Борботько Т.В. – д-р техн. наук, профессор

Аннотация. Рассмотрены современные методы глубокого обучения в области обнаружения аномалий. Описаны возможности, особенности и принцип работы методов, их слабые места и способы их преодоления, достигнутые на рынке.

Ключевые слова: методы глубокого обучения, нейронные сети, обнаружение аномалий.

Введение. С развитием сетевых технологий и появлением Интернета вещей (IoT), структуры информационных сетей становятся более масштабными и комплексными, из-за чего повышается их уязвимость. Атаки нулевого дня, распределенный отказ в обслуживании (DDoS), программы вымогатели и другие угрозы, с которыми вынуждены сталкиваться современные специалисты информационной безопасности, требуют использовать средства защиты сильно превышающие эффективность классических решений в области обнаружения аномалий. Такими решениями стали методы глубокого обучения, которые позволяют обучать различные модели нейронных сетей для обнаружения отклонений от стандартного потока в трафике.

Основная часть. Глубокое обучение это методология в области искусственного интеллекта, основанная на многослойных нейронных сетях, с помощью которой компьютеры учат обрабатывать данные таким же способом, как и человеческий мозг. Каждый слой сети обрабатывает информацию, придавая ей вес при помощи параметров (весов и смещений), оставляя признаки или отбрасывая их посредством функции активации. Функция активации − нелинейное преобразование, определяющее выходной сигнал нейрона на основе суммы его входов перед передачей его дальше. Методы глубокого обучения разработаны для обработки многомерных данных высокой размерности, что позволяет смягчить сложности моделирования отдельных переменных.

Аномалии могут обладать множеством разнообразных характеристик в низкоразмерных данных и могут оставаться незамеченными в многомерном пространстве, что делает идентификацию нелинейных аномалий высокого порядка сложной задачей. В рамках глубокого обучения существует многоуровневый подход к обучению, который позволяет вести обучение на различных уровнях абстракции, нелинейных преобразованиях и признаках, и дающий модели возможность опознавать такие нелинейные аномалии.

Для обнаружения отклонений, методы глубокого обучения используют архитектуру нейронных сетей, за счет чего они могут извлекать признаки (Feature Extraction). Извлечение признаков – это процесс преобразования необработанных высокоразмерных данных в набор информативных числовых характеристик, что снижает размерность данных, устраняет шум и облегчает работу моделей машинного обучения. Признак можно определить как количественно измеримое свойство в наборе данных, которое влияет на стратегию реализации модели. Определяя признаки, можно определить ключевые закономерности или отклонения. Рассмотренные в рамках работы модели глубокого обучения представлены ниже.

Сверточные нейронные сети (CNN) используют слои свертки для определения признаков и хорошо подходят для анализа пространственных данных, так как изначально создавались для задач по распознаванию изображений. CNN зарекомендовали себя как точный инструмент в области обнаружения аномалий, так как в обучении они способны улавливать локальные закономерности и иерархии в больших массивах данных.

Сеть глубоких убеждений (DBN) – многослойная архитектура, разработанная для опознания абстрактных понятий на более высоких уровнях, в то время как нижние слои позволяют распознавать закономерности. Ключевым компонентом DBN являются ограниченные машины Больцмана (RBM), которые помогают в изучении распределения вероятностей входных данных. Такой подход является стохастическим и использует обратное распространение ошибки для повышения производительности.

Ограниченная машина Больцмана (RBM), являющаяся двухуровневой архитектурой, функционирует путем анализа вероятностного распределения входных данных. Она состоит из видимого и скрытого слоев, при этом разные нейроны в одном и том же слое не могут взаимодействовать, что делает эту модель эффективной в снижении размерности. Наиболее известными вариантами являются бинарные и гауссовские RBM.

Рекуррентные нейронные сети (RNN) разработаны для извлечения признаков из временных рядов и последовательных данных, улавливая временные зависимости и представляя закономерности во времени. Интегрируя RNN с автокодировщиками и долгой краткосрочной памятью (LSTM), увеличивает точность обнаружения аберраций. За счет добавленной памяти RNN способны учитывать прошлые данные, благодаря чему они реагируют на новопоступающие данные, а также данные, обработанные моделью ранее. Этот метод глубокого обучения используется для выявления сложных постоянных угроз (APT), DDoS-атак в сетях Интернета вещей (IoT). Основным отличием RNN от остальных нейронных сетей является наличие узла обратной связи, однако традиционным RNN тяжело с долгосрочными зависимости из-за исчезновений и взрыва градиента, что может привести к недостаточному обучению моделей и неправильной настройке весов. Для борьбы с этой проблемой были разработаны такие передовые архитектуры, как LSTM и управляемые рекуррентные блоки (GRU), которые интегрируют механизмы управления, которые модулируют поток информации, тем самым повышая эффективность улавливания долгосрочных зависимостей, к примеру, за счет того, что позволяет сети сохранять или отбрасывать информацию по мере необходимости, тем самым смягчая трудности, связанные с проблемой исчезающего градиента.

Автокодировщики это нелинейное обобщение метода главных компонент (PCA), представляющие собой нейронные сети прямого распространения без учителя. Модель обучается для восстановления входных данных и аномалий из сжатой формы, уделяя особое внимание снижению размерности. Нормальный трафик обычно восстанавливается с низкой погрешностью, в то время как аномалии восстанавливаются с большими ошибками, тем самым выделяя себя среди остального набора данных.

Генеративно-состязательные сети (GANs) состоят из двух основных компонентов: Генератора (G) и Дискриминатора (D). Функция генератора заключается в создании синтетических данных, тогда как роль дискриминатора – это оценка достоверности этих данных. Генератор получает знания из шума, полученного из набора примеров, в то время как дискриминатор различает искусственные и подлинные образцы. После обучения модели аномалии могут быть выявлены путем сравнения реального сетевого трафика и сгенерированного трафика, тем самым любое существенное отклонение указывает на аномалию. Дискриминатор оценивает эти отклонения, отмечая шаблоны трафика, которые не соответствуют изученному ранее нормальному распределению. Этот подход применяется для обеспечения надежной сетевой безопасности в динамических средах, в частности, в областях, где используется синтез изображений, перевод иностранных языков с помощью технологий искусственного интеллекта и обнаружение аномалий.

Заключение. В ходе работы были рассмотрены современные методы глубокого обучения в области обнаружения аномалий. Описанные методы позволят значительно усилить защиту против атак нулевого дня, распределенного отказа в обслуживании (DDoS) и др. В настоящее время в области обнаружения аномалий сосредоточены на разработке эффективных и точных моделей глубокого обучения, которые бы обеспечивали баланс между улучшенной производительностью и повышенной интерпретируемостью. Ведется активная разработка гибридных систем (ансамблей) где одновременно используются несколько вариантов обучения модели, достигая тем самым еще более точных результатов в обнаружении отклонений.

Список литературы:

1. Kohli, M., Chhabra, I. A comprehensive survey on techniques, challenges, evaluation metrics and applications of deep learning models for anomaly detection. Discov Appl Sci 7, 784 (2025). [Электронный ресурс]. − Режим доступа: https://doi.org/10.1007/s42452-025-07312-7.

2. Как устроенно глубокое обучение нейросетей [Электронный ресурс]. − Режим доступа: https://habr.com/ru/articles/918188/.