УДК 004.056
doi:10.31799/1684-8853-2021-1-28-37

Научные статьи
Articles

Анализ моделей глубокого обучения для задач обнаружения сетевых аномалий интернета вещей

Д. А. Гайфулина[a], *младший научный сотрудник, orcid.org/0000-0002-5266-8649*
И. В. Котенко[a], *доктор техн. наук, профессор, orcid.org/0000-0001-6859-7120, ivkote@comsec.spb.ru*
[a]*Санкт-Петербургский федеральный исследовательский центр РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ*

Введение: *актуальность решения задачи выбора моделей глубокого обучения для обнаружения аномалий в сетевом трафике интернета вещей связана с необходимостью анализировать большое число событий безопасности для выявления аномального поведения умных устройств. Мощной технологией анализа таких данных является машинное и, в частности, глубокое обучение.* ***Цель:*** *выработка рекомендаций по выбору моделей глубокого обучения для обнаружения аномалий в сетевом трафике интернета вещей.* ***Результаты:*** *проведен сравнительный анализ моделей глубокого обучения и предоставлены рекомендации по их использованию для обнаружения аномалий в сетевом трафике интернета вещей. В качестве базовых моделей глубокого обучения рассмотрены многослойный персептрон, сверточная нейронная сеть, рекуррентная нейронная сеть, блок долгой краткосрочной памяти, управляемый рекуррентный блок и комбинированная сверточно-рекуррентная нейронная сеть. Дополнительно осуществлен анализ следующих моделей традиционного машинного обучения: наивный байесовский классификатор, метод опорных векторов, логистическая регрессия, метод k-ближайших соседей, бустинг и случайный лес. Показателями эффективности обнаружения аномалий выступали следующие метрики: аккуратность, точность, полнота и F-мера, а также временные затраты на обучение модели. Построенные в процессе эксперимента модели глубокого обучения продемонстрировали более высокие показатели точности обнаружения аномалий в гетерогенном трафике большого объема, характерного для интернета вещей, по сравнению с методами традиционного машинного обучения. Выявлено, что с ростом числа слоев в нейронных сетях возрастает полнота обнаружения аномальных соединений, что улучшает распознавание неизвестных аномалий, но влечет за собой рост ложных срабатываний. Подготовка моделей традиционного машинного обучения в ряде случаев занимает меньшее время. Это связано с тем, что применение методов глубокого обучения требует большего количества ресурсов и вычислительных мощностей.* ***Практическая значимость:*** *полученные в исследовании результаты могут быть использованы для построения систем обнаружения сетевых аномалий в интернете вещей.*

Ключевые слова *— глубокое обучение, глубокие нейронные сети, обнаружение аномалий, интернет вещей, информационная безопасность.*

Для цитирования: Гайфулина Д. А., Котенко И. В. Анализ моделей глубокого обучения для задач обнаружения сетевых аномалий интернета вещей. *Информационно-управляющие системы*, 2021, № 1, с. 28–37. doi:10.31799/1684-8853-2021-1-28-37
For citation: Gaifulina D. A., Kotenko I. V. Analysis of deep learning models for network anomaly detection in Internet of Things. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 28–37 (In Russian). doi:10.31799/1684-8853-2021-1-28-37

Введение

В современном мире технология интернета вещей (Internet of Things — IoT) находит все большее применение в повседневной жизни человека. В общем виде интернет вещей представляет собой сеть распределенных устройств, которые связаны с окружающей средой при помощи датчиков, а также с программным обеспечением и серверами. Такие устройства также называют умными, или интеллектуальными. В то же время с ростом спроса на умные устройства и их доступности совершенствуются способы атак злоумышленников. Разнородность устройств и соединений, а также их ограничения на вычислительные ресурсы усложняют управление системами интернета вещей. В связи с этим обнаружение аномального поведения умных устройств порой сильно затруднено [1].

Для реагирования на угрозы безопасности необходимы инструменты анализа большого числа событий в системах интернета вещей, которые содержатся в сетевом трафике, логах и иных данных, объем которых порой очень велик. Помимо этого, разнородность источников и хранилищ информации приводит к высокой гетерогенности анализируемых данных. Машинное обучение и, в частности, глубокое обучение на данный момент являются мощными технологиями для анализа событий безопасности, обнаружения атак и аномального поведения умных устройств [2]. Ранее авторами был представлен системный анализ современных методов глубокого обучения, применяемых в задачах кибербезопасности [3, 4]. При этом сравнивать между собой различные модели глубоких нейронных сетей в научной литературе достаточно проблематично — в оценке эффективности применения моделей исследователи используют разные наборы данных или отличающиеся подмножества конкретного набора. Научная новизна проводимого исследования со-

стоит в предложенном сравнительном анализе моделей глубокого обучения различных классов и архитектур, основанном на оценке эффективности обнаружения аномалий в сетевом трафике интернета вещей с использованием единого программно-аппаратного обеспечения и одинаковых подмножеств набора данных для обучения и тестирования. Основной задачей проводимого исследования является выработка рекомендаций по выбору моделей глубокого обучения с высокими показателями эффективности обнаружения аномалий в сетевом трафике интернета вещей.

Классификация моделей глубокого обучения

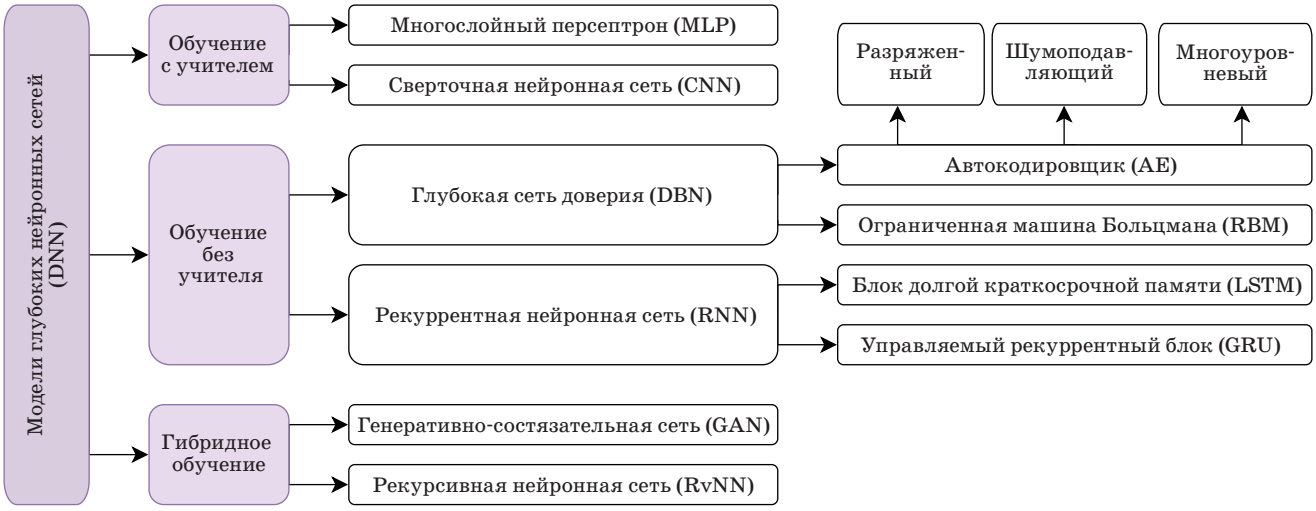
Глубокое обучение является частью семейства методов машинного обучения и основано на применении искусственных нейронных сетей. Глубокая нейронная сеть (Deep Neural Network — DNN) представляет собой нейронную сеть с несколькими слоями между входным и выходным слоями. Целью обучения DNN является нахождение корректного метода математических преобразований для превращения входных данных в выходные, независимо от линейной или нелинейной корреляции. Обучение может проходить как с учителем (supervised learning), так и без (unsupervised learning), а также при сочетании этих двух методов. Классификация наиболее распространенных моделей DNN по способу обучения представлена на рис. 1.

Многослойный персептрон (Multilayer Perceptron — MLP) является классом искусственных нейронных сетей прямого распространения (Feed Forward Neural Network — FFNN). При обнаружении вторжений, как правило, MLP используется для бинарной классификации сетевых соединений (нормальное или аномальное поведение). Взвешенные комбинации выходного слоя представляют собой прогноз, указывающий на принадлежность соединения к определенному классу.

Сверточная нейронная сеть (Convolutional Neural Network — CNN) используется обычно для эффективного распознавания образов, что также позволяет применять их для выявления вторжений. Заголовки пакетов сетевого трафика или последовательности системных вызовов кодируются в двумерную матрицу, которая является входными данными для CNN.

Методы обнаружения аномалий с использованием *глубоких сетей доверия* (Deep Belief Network — DBN), таких как *автокодировщики* (Autoencoder — AE) и *ограниченная машина Больцмана* (RBM), основаны на реконструкции данных, при которой определяется величина расхождения нормальных и аномальных данных. Для этого применяется распространение сигналов ошибки от выходов сети к входам для получения на выходе отклика, наиболее близкого к входному.

Связи между нейронами *рекуррентной нейронной сети* (Recurrent Neural Network — RNN) образуют направленный цикл. RNN может использовать свою внутреннюю память, такую как *блок долгой краткосрочной памяти* (Long Short-Term Memory — LSTM) или *управляемый рекуррентный блок* (Gated Recurrent Units — GRU). Использование RNN позволяет анализировать данные в виде временных рядов: сетевого трафика, последовательностей системных вызовов, журналов событий. Аномалия при этом может



■ ***Рис. 1.*** Классификация основных моделей DNN
■ ***Fig. 1.*** Classification of the main models of DNN

быть распознана как отклонение от предсказанного сетью последующего состояния.

В *генеративно-состязательных сетях* (Generative Adversarial Networks — GAN) используются модели генератора, анализирующего распределение реальных данных, и дискриминатора, оценивающего вероятность того, что входные данные поступают из реальных данных или из генератора. В подходах к обнаружению вторжений GAN применяются для исследования распределения нормальных данных, чтобы распознавать неизвестные аномалии.

Обзор существующих работ

Большая часть исследователей в области безопасности интернета вещей рассматривает методы глубокого обучения в рамках подходов к обнаружению атак и аномального поведения устройств [5, 6]. Основными преимуществами DNN по сравнению с методами традиционного машинного обучения являются высокая производительность и масштабируемость для растущего объема данных, а также возможность автоматически отбирать информативные признаки из необработанных данных.

В статье [7] исследуется потенциал рекуррентных нейронных сетей с LSTM для обнаружения вредоносных программ интернета вещей. Осуществляется сравнение разработанной модели с классификаторами, основанными на традиционных методах машинного обучения: методе опорных векторов (Support Vector Machine — SVM), наивном байесовском классификаторе (Naive Bayes), случайном лесе (Random Forest), бустинге (Adaptive Boosting — AdaBoost) и методе k-ближайших соседей (k-nearest neighbors algorithm — kNN). Анализ демонстрирует, что подход на основе глубокого обучения обеспечивает наилучший возможный результат. Сравнение с другими моделями глубокого обучения не производилось.

В исследовании [8] авторы предлагают собственную систему обнаружения аномалий для индустриальных систем интернета вещей с использованием автокодировщика и глубокой нейронной сети с прямой связью. Проводится сравнение созданной модели с характеристиками нескольких разработанных методов обнаружения аномалий, в том числе с глубокой сетью доверия [9], рекуррентной сетью [10], DNN [11] и Ensemble-DNN [12]. При этом приведенные модели оценивались на разных подмножествах исходных данных и с использованием разнородного аппаратного и программного обеспечения.

В статье [13] предлагается распределенная облачная среда глубокого обучения для обнаружения и предотвращения фишинговых и ботнет-атак на умные устройства. Разработанная модель RNN-LSTM сравнивается с моделями глубокого обучения, разработанными другими исследователями. Основным недостатком сравнительного анализа в данной работе является то, что рассматриваемые модели DNN оцениваются не на одинаковых наборах данных.

Авторы статьи [14] анализируют несколько методов глубокого обучения для обнаружения DDoS-атак: многослойный персептрон, сверточную нейронную сеть, RNN-LSTM и ансамбль CNN+LSTM. Проведено их сравнение с традиционными методами машинного обучения: методом опорных векторов, байесовским классификатором и случайным лесом. Авторы делают вывод о большей эффективности методов глубокого обучения, в особенности рекуррентных сетей.

В работе [15] также проводится систематическое сравнение CNN и RNN в системах обнаружения вторжений. Оцениваются следующие модели: базовая CNN (Basic CNN), CNN начальной архитектуры (Inception Architecture CNN), RNN-LSTM и управляемый рекуррентный блок. Авторы приходят к выводу, что CNN лучше подходит для бинарной классификации при обнаружении аномалий, а RNN лучше работают при обнаружении сложных атак в задачах мультиклассовой классификации.

Анализ указанных релевантных работ проведен по следующим атрибутам (табл. 1): сравнение моделей глубокого обучения (Г) между собой и с методами традиционного машинного обучения (М), использование метрик аккуратности (A), точности (P), полноты (R), F-меры (F) и временных затрат (T), используемый набор данных и анализируемые модели обучения.

Таким образом, особенностями предлагаемого исследования по сравнению с приведенными релевантными работами являются:

1) проведение эксперимента на едином наборе данных и с использованием одинакового программно-аппаратного обеспечения;

2) расширение сравнительной выборки моделей как глубокого обучения, так и традиционного машинного обучения;

3) введение оценки временных затрат на обучение модели, помимо таких показателей эффективности обнаружения аномалий, как аккуратность, точность, полнота и F-мера.

Выбор моделей глубоких нейронных сетей

В данном разделе проанализированы основные модели глубоких нейронных сетей:

— обучение с учителем — многослойный персептрон (MLP), сверточная нейронная сеть (CNN);

■ ***Таблица 1***. Анализ релевантных работ
■ ***Table 1***. Analysis of relevant works

Авторы, год	Метод		Показатель					Набор данных	Модели	Точность, %
	М	Г	A	P	R	F	T			
Haddad Pajouh, 2018 [7]	+	–	+	–	–	–	–	VirusTotal	RNN-LSTM	98
									SVM	82
									Naive Bayes	90
									Random Forest	92
									Ada Boost	93
									kNN	94
Muna, 2018 [8]	–	+	+	–	–	–	–	NSL-KDD	AE+FFNN	99
									DBN	95
									RNN	73
									DNN	76
									Ensemble-DNN	98
Parra, 2020 [13]	+	+	+	+	+	+	–	Собственный	RNN-LSTM	94
								CSIC 2010	DBN	95
									GRU	97
									SVM	99
Roopak, 2019 [14]	+	+	+	+	+	–	–	CICIDS 2017	MLP	86
									CNN	95
									LSTM	96
									CNN+LSTM	97
									SVM	95
									Naive Bayes	95
									Random Forest	94
Cui, 2018 [15]	–	+	+	+	+	+	–	ISCX2012	Basic CNN	94
									Inception Architecture CNN	95
									RNN-LSTM	93,7
									GRU	94

— обучение без учителя — рекуррентная нейронная сеть (RNN), блок долгой краткосрочной памяти (RNN-LSTM), управляемый рекуррентный блок (RNN-GRU);

— смешанное обучение — сверточно-рекуррентная сеть (CNN+RNN).

Методы глубокого обучения выбраны на основе проведения анализа существующих подходов к выявлению сетевых аномалий. Систематический анализ методов глубокого обучения, используемых в кибербезопасности, продемонстрировал, что данные модели дают хорошие результаты на практике [3, 4]. Анализ показателей обнаружения аномалий в сетевом трафике проводится не только между моделями разных классов, но и между моделями одного класса с различным количеством слоев и нейронов. Это позволяет экспериментально определить зависимость между структурой модели и ее производительностью.

Основные параметры выбранных моделей представлены в табл. 2. Архитектура сети описывается следующим образом: количество слоев и нейронов в каждом из них (h — скрытый слой, p — субдискретизирующий слой, c — сверточный слой, n — полносвязный слой).

Функция активации нейронной сети определяет выходное значение в зависимости от результата взвешенной суммы входов и порогового значения [16]. Для всех моделей в качестве функции активации выходного слоя выбрана сигмои-

■ ***Таблица 2***. Параметры модели глубоких нейронных сетей

■ ***Table 2***. Parameters of the DNN

Обозначение	Архитектура	Функция активации
MLP-1	h(1024)-h(768)	ReLu, sigmoid
MLP-2	h(1024)-h(768)-h(512)	
MLP-3	h(1024)-h(768)-h(512)-h(256)	
MLP-4	h(1024)-h(768)-h(512)-h(256)-h(128)	
CNN-1	2c(64)-1p(2)-1n(128)	ReLu, sigmoid
CNN-2	2c(64)-1p(2)-2c(128)-1p(2)-1n(128)	
RNN-1	h(16)-h(16)-h(16)	Sigmoid
RNN-2	h(32)-h(32)-h(32) -h(32)	
RNN-LSTM-1	h(16)-h(16)-h(16)	
RNN-LSTM-2	h(32)-h(32)-h(32) -h(32)	
RNN-GRU-1	h(16)-h(16)-h(16)	
RNN-GRU-2	h(32)-h(32)-h(32) -h(32)	
CNN+LSTM-1	CNN(2c(64)-1p(2))-LSTM(h(128))	
CNN+LSTM-2	CNN(2c(64)-1p(2)-2c(128)-1p(2)- LSTM(h(128)	

да (sigmoid), и в ряде моделей она дополнена линейным выпрямителем (Rectified linear unit — ReLU) на скрытых слоях. Также применяется метод контроля емкости дропаут (dropout), позволяющий предотвратить переобучение нейронной сети [17].

Эксперименты

Оценка производительности любых систем обнаружения аномалий для интернета вещей требует наличия исходных данных, включающих в себя набор сетевых признаков, таких как признаки на основе номеров портов источника и назначения, полезной нагрузки (payload-based), поведения (behaviour based) и потока данных (flow-based).

В качестве экспериментальных данных для анализа моделей DNN в задачах обнаружения сетевых аномалий интернета вещей был выбран открытый набор данных UNSW-NB15 [18, 19], содержащий 2 540 044 записей — векторов признаков сетевых соединений TCP/IP и соответствующих им меток классов. В этом наборе данных сетевые пакеты включают как информацию о реальной нормальной активности сети, так и девять типов атак: фаззеры (Fuzzers), анализаторы (Analysis), бэкдоры (Backdoors), отказ в обслуживании (DoS), эксплоиты (Exploits), обобщенные (Generic), разведка (Reconnaissance), шелл-код (Shellcode) и черви (Worms). Данные UNSW-NB15 для обучения и тестирования систем обнаружения вторжений содержат 47 признаков, таких как IP-адреса, номера портов, байты транзакции и др. [20], и две метки класса — категорию атаки и метку аномальности соединения. Первые 35 признаков представляют собой интегрированную информацию из пакетов данных, а остальные определяются для сценариев подключения.

Обнаружение аномалий представляет собой процесс идентификации отклонений от нормального профиля системы. Таким образом, для обнаружения аномалий в сетевом трафике UNSW-NB15 используется бинарная классификация, и в качестве метки класса используется критерий аномальности соединения, где 0 соответствует нормальному профилю, а 1 — аномалии.

Анализ моделей DNN для задач обнаружения сетевых аномалий интернета вещей состоит из описанных ниже этапов.

Предобработка данных **(1)** заключается в преобразовании входного набора данных: 47 признаков сетевых соединений и метки класса — в форму, подаваемую на вход анализируемым моделям. К признакам номинального типа, таким как IP-адреса, название протокола и сервиса передачи данных, применяется горячее кодирование (one-hot encoding) — метод представления категориальных переменных в виде двоичных векторов. Далее производится нормализация значений всех признаков к диапазону [0...1]. Нормализация данных осуществляется, так как дисбаланс между значениями признаков может вызвать неустойчивость работы модели, ухудшить результаты обучения и замедлить процесс моделирования. В качестве данных для обучения моделей выбирается 80% исходного набора данных (1 547 081 запись), а для тестирования моделей — 20% (386 771 запись). Важной особенностью данного этапа исследований является отсутствие высокой сбалансированности нормального и аномального класса сетевых соединений, что наиболее близко к реальным условиям при возникновении аномалий в сетевом трафике. Так, в данном случае отношение аномальных данных к нормальным составляет 1:4. Обучающая и тестовая выборка являются однородными.

Обучение моделей **(2)** осуществляется на одинаковом тренировочном наборе данных, а обнаружение аномалий **(3)** — на одинаковом тестовом наборе данных. Для обучения и валидации моде-

лей глубокого обучения использовались следующие гиперпараметры: размер пакта (batch size) — 64, алгоритм оптимизации — adam, функция потерь (loss function) — binary cross-entropy.

Разработанные модели глубоких нейронных сетей и традиционного машинного обучения были реализованы с использованием Python 3.6, Tensorflow 2.1, Scikit-learn 0.23.2, Numpy 1.19.2, Pandas 1.1.3 и Scipy 1.5.2. Все эксперименты проводились на Acer Swift SF315-52G с процессором Intel Core i5 с тактовой частотой 1,8 ГГц, ОЗУ 8 ГБ и операционной системой Windows 10.

Оценка эффективности обнаружения аномалий **(4)** заключается в вычислении следующих метрик: аккуратности (A), точности (P), полноты (R), F-меры (F) и временных затрат на обучение (T). Аккуратность характеризует долю экземпляров сетевых соединений, по которым модель приняла правильное решение о принадлежности к нормальному или аномальному классу. Точность характеризует долю верно классифицированных экземпляров сетевых соединений относительно всех экземпляров сетевого трафика. Полнота характеризует долю найденных моделью экземпляров сетевых соединений, принадлежащих нормальному или аномальному классу относительно всех экземпляров сетевого трафика. F-мера представляет собой гармоническое среднее между точностью и полнотой.

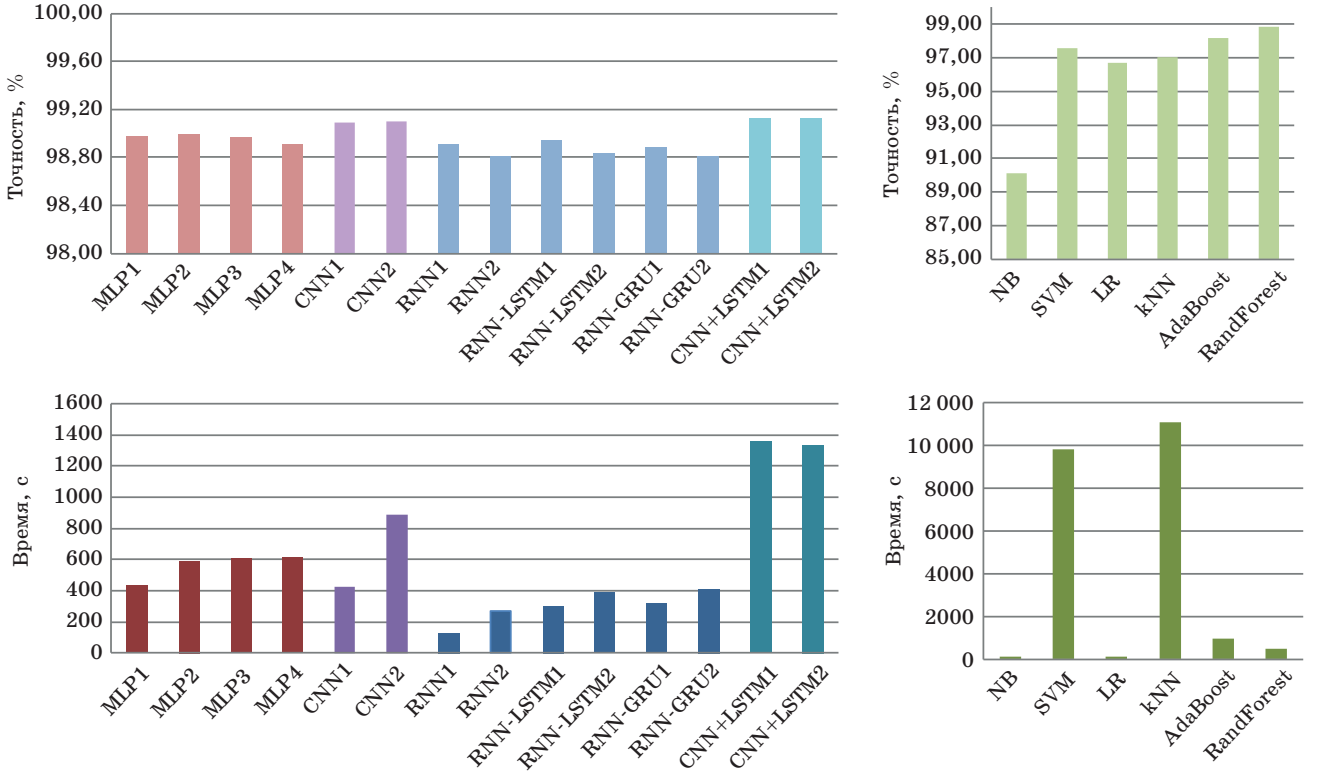
Результаты экспериментов по анализу моделей глубокого обучения представлены в табл. 3. Данные приводятся для первой эпохи обучения.

■ ***Таблица 3***. Анализ моделей глубокого обучения
■ ***Table 3***. Analysis of deep learning models

Модель	Оценка				
	A, %	P, %	R, %	F, %	T, с
MLP-1	98,976	93,365	98,465	95,847	430,946
MLP-2	98,996	93,612	98,384	95,939	595,026
MLP-3	98,965	93,198	**98,552**	95,800	606,213
MLP-4	98,913	92,665	**98,669**	95,573	617,150
CNN-1	**99,095**	**95,049**	97,741	**96,376**	423,574
CNN-2	**99,096**	94,713	98,086	**96,370**	892,186
RNN-1	98,912	93,621	97,688	95,612	**128,943**
RNN-2	98,810	92,199	98,295	95,149	**273,471**
RNN-LSTM-1	**98,938**	93,952	97,569	95,726	**301,561**
RNN-LSTM-2	98,840	92,629	98,106	95,289	388,294
RNN-GRU-1	98,883	93,212	97,867	95,483	317,043
RNN-GRU-2	98,808	92,104	**98,381**	95,139	412,727
CNN+LSTM-1	**99,126**	**95,932**	97,127	**96,526**	1355,819
CNN+LSTM-2	**99,124**	**95,972**	97,078	**96,522**	1334,130

■ ***Таблица 4***. Анализ моделей традиционного машинного обучения
■ ***Table 4***. Analysis of traditional machine learning models

Модель	Оценка				
	A, %	P, %	R, %	F, %	T, с
Naive Bayes	90,12	94,22	90,12	91,18	**130,36**
SVM	97,56	97,58	97,58	97,58	9786,67
Logistic Regression	96,73	96,77	96,73	96,74	**131,26**
kNN	97,07	97,07	97,07	97,07	11074,83
AdaBoost	98,2	98,2	98,2	98,2	940,17
Random Forest	**98,87**	**98,85**	**98,85**	**98,45**	468,05



■ ***Рис. 2.*** Сравнение точности и времени обучения моделей
■ ***Fig. 2.*** Comparison of accuracy and time of model learning

Дополнительно на тех же данных проанализированы следующие модели традиционного машинного обучения: наивный байесовский классификатор, метод опорных векторов, логистическая регрессия (Logistic Regression), метод k-ближайших соседей, бустинг и случайный лес. Результаты представлены в табл. 4.

Сравнение точности рассмотренных моделей и времени обучения для обнаружения аномалий продемонстрировано на рис. 2.

Анализ результатов экспериментов

Результаты проведенных экспериментов позволяют сделать вывод, что большинство моделей глубоких нейронных сетей обладает высокой точностью обнаружения аномалий в гетерогенном трафике большого объема для применения их на практике. Среди моделей традиционного машинного обучения сходной высокой точностью обладают ансамбли классификаторов, такие как AdaBoost (A = 98,2 %) и случайный лес (A = 98,87 %).

Среди моделей глубокого обучения с учителем лучшую точность обнаружения демонстрирует сверточная нейронная сеть (A = 99,1 %). При этом с увеличением числа слоев время обучения существенно возрастает, в отличие от показателя точности, изменяющегося не так сильно. Многослойный персептрон обладает наибольшей полнотой обнаружения аномалий (R = 98,67 %). Это значит, что данная модель распознает большее количество экземпляров аномальных соединений, что позволяет избежать их пропусков. С увеличением числа слоев точность многослойного персептрона ухудшается, а время, затраченное на обучение, возрастает (см. рис. 2). В данном эксперименте предпочтительной архитектурой многослойного персептрона является модель MLP-2 с точностью обнаружения аномалий A = 99 %.

Для моделей глубокого обучения без учителя, представленных рекуррентными сетями, лучшие результаты показывает блок долгой краткосрочной памяти (A = 98,938 %). Обучение данного вида моделей занимает наименьшее количество времени, следовательно, и меньшее количество вычислительных ресурсов, что нередко является существенным параметром для устройств интернета вещей. С увеличением числа слоев в архитектуре рекуррентных нейронных сетей возрастает полнота обнаружения аномальных соединений, но снижается точность, что связано с накоплением ошибок обучения. Таким образом, в модели обнаружения аномалий, настроенной на низкий коэффициент ложных срабатываний, точность будет

представлять собой более значимую характеристику, тогда как модель с высокой полнотой классификации предпочтительней для распознавания ранее неизвестных типов аномалий.

Наивысшей точностью обнаружения аномалий среди представленных моделей DNN обладает комбинированная сверточно-рекуррентная нейронная сеть (A = 99,13 %). При этом время обучения данной сети является самым продолжительным.

При сравнении между собой базовых моделей DNN разных классов можно установить, что различие в точности обнаружения аномалий не является весьма значительным — не более 1 %. Более разнящейся характеристикой является время обучения сети, которое также возрастает соответственно увеличению числа слоев. Стоит отметить, что подготовка моделей традиционного машинного обучения по большей части занимает меньшее количество времени, за исключением моделей опорных векторов и k-ближайших соседей. Это связано с тем, что применение методов глубокого обучения требует большего количества вычислительных мощностей.

Представлены рекомендации (табл. 5) для выбора наиболее предпочтительной модели глубокого обучения исходя из временных затрат на обучение и приоритета в обнаружении аномалий в сетевом трафике интернета вещей.

Для систем, работающих в режиме реального времени и часто обновляемых, скорость моделирования является значимой характеристикой и должна быть минимизирована. В то время как для некоторых систем, обучаемых офлайн, время моделирования может быть увеличено для более тщательной настройки и повышения эффективности функционирования.

■ ***Таблица 5***. Рекомендации по использованию моделей глубокого обучения

■ ***Table 5***. Recommendations for using deep learning models

Приоритет в обнаружении аномалий	Временные затраты	Рекомендация
Низкий коэффициент ложных срабатываний	Не имеют значения	Комбинированная нейронная сеть
	Минимизированы	Рекуррентная нейронная сеть, в частности блок долгой краткосрочной памяти
Распознавание неизвестных типов аномалий	Не имеют значения	Многослойный персептрон с большим количеством слоев
	Минимизированы	Управляемый рекуррентный блок с большим количеством слоев

Заключение

В данном исследовании представлен анализ базовых моделей глубокого обучения для задач обнаружения аномалий в сетевом трафике интернета вещей. Экспериментальная оценка моделей глубокого обучения проводилась с использованием единого программно-аппаратного обеспечения и одинаковых подмножеств набора данных UNSW-NB 15 для обучения и тестирования. В качестве базовых моделей глубоких нейронных сетей рассмотрены многослойный персептрон, сверточная нейронная сеть, рекуррентная нейронная сеть, блок долгой краткосрочной памяти, управляемый рекуррентный блок и комбинированная сверточно-рекуррентная нейронная сеть.

Построенные модели глубокого обучения продемонстрировали высокие показатели точности обнаружения аномалий — от 98,8 %. В работе представлены рекомендации для выбора наиболее предпочтительной модели глубокого обучения исходя из временных затрат на обучение модели и приоритета в обнаружении аномалий в сетевом трафике интернета вещей. При настройке модели обнаружения аномалий на низкий коэффициент ложных срабатываний точность будет представлять собой значимую характеристику. С увеличением числа слоев в архитектуре DNN возрастает полнота обнаружения аномальных соединений, что в свою очередь предпочтительней для распознавания ранее неизвестных типов аномалий. Увеличение числа слоев в модели обнаружения аномалий требует мощных вычислительных ресурсов центральных компонентов интернета вещей.

В дальнейшем предполагается продолжить анализ характеристик моделей DNN, применяемых в задачах кибербезопасности. Одним из направлений будущих работ является исследование влияния структуры сетевого трафика на показатели эффективности использования моделей глубокого обучения. На основании полученных результатов планируется разработать подход к выявлению и корреляции событий безопасности на базе методов глубокого обучения.

Финансовая поддержка

Работа выполнена при частичной финансовой поддержке проекта РФФИ 18-29-22034 мк и бюджетной темы 0073-2019-0002.

Литература

1. **Alrawais A., Alhothaily A., Hu C., Cheng X.** Fog computing for the internet of things: Security and privacy issues. *IEEE Internet Computing*, 2017, no. 21(2), pp. 34–42. doi:10.1109/MIC.2017.37
2. **Браницкий А. А., Котенко И. В.** Обнаружение сетевых атак на основе комплексирования нейронных, иммунных и нейро-нечетких классификаторов. *Информационно-управляющие системы*, 2015, № 4, с. 69–77. doi:10.15217/issn1684-8853.2015.4.69
3. **Гайфулина Д. А., Котенко И. В.** Применение методов глубокого обучения для решения задач кибербезопасности. Ч. 1. *Вопросы кибербезопасности*, 2020, № 3(37), с. 76–86. doi:10.21681/2311-3456-2020-03-76-86
4. **Гайфулина Д. А., Котенко И. В.** Применение методов глубокого обучения для решения задач кибербезопасности. Ч. 2. *Вопросы кибербезопасности*, 2020, № 4(38), с. 11–21. doi:10.21681/2311-3456-2020-04-11-21
5. **Al-Garadi M. A., Mohamed A., Al-Ali A., Du X., Guizani M.** A Survey of machine and deep learning methods for internet of things (IoT) security. *IEEE Communications Surveys & Tutorials*, 2020, vol. 22, no. 3, pp. 1646–1685. doi:10.1109/COMST.2020.2988293
6. **Левшун Д. С., Гайфулина Д. А., Чечулин А. А., Котенко И. В.** Проблемные вопросы информационной безопасности киберфизических систем. *Информатика и автоматизация*, 2020, т. 19, № 5, с. 1050–1088. doi:10.15622/ia.2020.19.5.6
7. **HaddadPajouh H., Dehghantanha A., Khayami R., Choo K. K. R.** A Deep recurrent neural network based approach for internet of things malware threat hunting. *Future Generation Computer Systems*, 2018, vol. 85, pp. 88–96. doi:10.1016/j.future.2018.03.007
8. **Muna Al H., Moustafa N., Sitnikova E.** Identification of malicious activities in industrial internet of things based on deep learning models. *Journal of Information Security and Applications*, 2018, vol. 41, pp. 1–11. doi:10.1016/j.jisa.2018.05.002
9. **Alom M. Z., Bontupalli V., Taha T. M.** Intrusion detection using deep belief networks. *2015 National Aerospace and Electronics Conference (NAECON)*, Dayton, 2015, pp. 339–344. doi:10.1109/NAECON.2015.7443094
10. **Yin C., Zhu Y., Fei J., He X.** A Deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access*, 2017, vol. 5, pp. 21954–21961. doi:10.1109/ACCESS.2017.2762418
11. **Tang T. A., Mhamdi L., McLernon D., Zaidi S., Ghogho M.** Deep learning approach for network intrusion detection in software defined networking. *2016 International Conference on Wireless Networks and Mobile Communications (WINCOM)*, 2016, pp. 258–263. doi:10.1109/WINCOM.2016.7777224
12. **Ludwig S. A.** Intrusion detection of multiple attack classes using a deep neural net ensemble. *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, Honolulu, 2017, pp. 1–7. doi:10.1109/SSCI.2017.8280825.
13. **Parra G. D. L. T., Rad P., Choo K. K. R., Beebe N.** Detecting internet of things attacks using distributed deep learning. *Journal of Network and Computer Applications*, 2020, vol. 163, pp. 102662. doi:10.1016/j.jnca.2020.102662
14. **Roopak M., Tian G. Y., Chambers J.** Deep learning models for cyber security in IoT networks. *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, USA, 2019, pp. 0452–0457. doi:10.1109/CCWC.2019.8666588
15. **Cui J., Long J., Min E., Liu Q., Li Q.** Comparative study of CNN and RNN for deep learning based intrusion detection system. *International Conference on Cloud Computing and Security*, Springer, Cham, 2018, pp. 159–170. doi:10.1007/978-3-030-00018-9_15
16. **Nwankpa C., Ijomah W., Gachagan A., Marshall S.** Activation functions: Comparison of trends in practice and research for deep learning. *ArXiv preprint arXiv:1811.03378*, 2018. 20 p.
17. **Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdino R.** Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 2014, vol. 15, no. 1, pp. 1929–1958.
18. *UNSW-NB15 Dataset*. https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/ (дата обращения: 27.10.2020).
19. **Moustafa N., Slay J.** UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). *2015 Military Communications and Information Systems Conference (MilCIS)*, Canberra, ACT, 2015, pp. 1–6. doi:10.1109/MilCIS.2015.7348942
20. **Moustafa N., Turnbull B., Choo K. R.** An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. *IEEE Internet of Things Journal*, 2019, vol. 6, no. 3, pp. 4815–4830. doi:10.1109/JIOT.2018.2871719

UDC 004.056
doi:10.31799/1684-8853-2021-1-28-37

Analysis of deep learning models for network anomaly detection in Internet of Things

D. A. Gaifulina[a], Junior Researcher, orcid.org/0000-0002-5266-8649
I. V. Kotenko[a], Dr. Sc., Tech, Professor, orcid.org/0000-0001-6859-7120, ivkote@comsec.spb.ru
[a]St. Petersburg Federal Research Center of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation

Introduction: The article discusses the problem of choosing deep learning models for detecting anomalies in Internet of Things (IoT) network traffic. This problem is associated with the necessity to analyze a large number of security events in order to identify the abnormal behavior of smart devices. A powerful technology for analyzing such data is machine learning and, in particular, deep learning. **Purpose:** Development of recommendations for the selection of deep learning models for anomaly detection in IoT network traffic. **Results:** The main results of the research are comparative analysis of deep learning models, and recommendations on the use of deep learning models for anomaly detection in IoT network traffic. Multilayer perceptron, convolutional neural network, recurrent neural network, long short-term memory, gated recurrent units, and combined convolutional-recurrent neural network were considered the basic deep learning models. Additionally, the authors analyzed the following traditional machine learning models: naive Bayesian classifier, support vector machines, logistic regression, k-nearest neighbors, boosting, and random forest. The following metrics were used as indicators of anomaly detection efficiency: accuracy, precision, recall, and F-measure, as well as the time spent on training the model. The constructed models demonstrated a higher accuracy rate for anomaly detection in large heterogeneous traffic typical for IoT, as compared to conventional machine learning methods. The authors found that with an increase in the number of neural network layers, the completeness of detecting anomalous connections rises. This has a positive effect on the recognition of unknown anomalies, but increases the number of false positives. In some cases, preparing traditional machine learning models takes less time. This is due to the fact that the application of deep learning methods requires more resources and computing power. **Practical relevance:** The results obtained can be used to build systems for network anomaly detection in Internet of Things traffic.

Keywords — deep learning, deep neural networks, anomaly detection, Internet of Things, information security.

For citation: Gaifulina D. A., Kotenko I. V. Analysis of deep learning models for network anomaly detection in Internet of Things. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 1, pp. 28–37 (In Russian). doi:10.31799/1684-8853-2021-1-28-37

References

1. Alrawais A., Alhothaily A., Hu C., Cheng X. Fog computing for the internet of things: Security and privacy issues. *IEEE Internet Computing*, 2017, no. 21(2), pp. 34–42. doi:10.1109/MIC.2017.37
2. Branitskiy A. A., Kotenko I. V. Network attack detection based on combination of neural, immune and neuro-fuzzy classifiers. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2015, no. 4, pp. 69–77 (In Russian). doi:10.15217/issn1684-8853.2015.4.69
3. Gaifulina D. A., Kotenko I. V. Application of deep learning methods in cybersecurity tasks. Part 1. *Voprosy kiberbezopasnosti*, 2020, no. 3(37), pp. 76–86 (In Russian). doi:10.21681/2311-3456-2020-03-76-86
4. Gaifulina D. A., Kotenko I. V. Application of deep learning methods in cybersecurity tasks. Part 2. *Voprosy kiberbezopasnosti*, 2020, no. 4(38), pp. 11–21 (In Russian). doi:10.21681/2311-3456-2020-04-11-21
5. Al-Garadi M. A., Mohamed A., Al-Ali A., Du X., Guizani M. A Survey of machine and deep learning methods for internet of things (IoT) security. *IEEE Communications Surveys & Tutorials*, 2020, vol. 22, no. 3, pp. 1646–1685. doi:10.1109/COMST.2020.2988293.
6. Levshun D., Gaifulina D., Chechulin A., Kotenko I. Problematic issues of information security of cyber-physical systems. *Informatics and Automation*, 2020, vol. 19, no. 5, pp. 1050–1088. doi:10.15622/ia.2020.19.5.6
7. HaddadPajouh H., Dehghantanha A., Khayami R., Choo K. K. R. A Deep recurrent neural network based approach for internet of things malware threat hunting. *Future Generation Computer Systems*, 2018, vol. 85, pp. 88–96. doi:10.1016/j.future.2018.03.007
8. Muna Al H., Moustafa N., Sitnikova E. Identification of malicious activities in industrial internet of things based on deep learning models. *Journal of Information Security and Applications*, 2018, vol. 41, pp. 1–11. doi:10.1016/j.jisa.2018.05.002
9. Alom M. Z., Bontupalli V., Taha T. M. Intrusion detection using deep belief networks. *2015 National Aerospace and Electronics Conference (NAECON)*, Dayton, 2015, pp. 339–344. doi:10.1109/NAECON.2015.7443094
10. Yin C., Zhu Y., Fei J., He X. A Deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access*, 2017, vol. 5, pp. 21954–21961. doi: 10.1109/ACCESS.2017.2762418
11. Tang T. A., Mhamdi L., McLernon D., Zaidi S., Ghogho M. Deep learning approach for network intrusion detection in software defined networking. *2016 International Conference on Wireless Networks and Mobile Communications (WINCOM)*, 2016, pp. 258–263. doi:10.1109/WINCOM.2016.7777224
12. Ludwig S. A. Intrusion detection of multiple attack classes using a deep neural net ensemble. *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, Honolulu, 2017, pp. 1–7. doi:10.1109/SSCI.2017.8280825
13. Parra G. D. L. T., Rad P., Choo K. K. R., Beebe N. Detecting internet of things attacks using distributed deep learning. *Journal of Network and Computer Applications*, 2020, vol. 163, pp. 102662. doi:10.1016/j.jnca.2020.102662
14. Roopak M., Tian G. Y., Chambers J. Deep learning models for cyber security in IoT networks. *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, USA, 2019, pp. 0452–0457. doi:10.1109/CCWC.2019.8666588
15. Cui J., Long J., Min E., Liu Q., Li Q. Comparative study of CNN and RNN for deep learning based intrusion detection system. *International Conference on Cloud Computing and Security*, Springer, Cham, 2018, pp. 159–170. doi:10.1007/978-3-030-00018-9_15
16. Nwankpa C., Ijomah W., Gachagan A., Marshall S. Activation functions: Comparison of trends in practice and research for deep learning. *ArXiv preprint arXiv:1811.03378*, 2018. 20 p.
17. Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdino R. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 2014, vol. 15, no. 1, pp. 1929–1958.
18. *UNSW-NB15 Dataset*. Available at: https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/ (accessed 27 October 2020).
19. Moustafa N., Slay J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). *2015 Military Communications and Information Systems Conference (MilCIS)*, Canberra, ACT, 2015, pp. 1–6. doi:10.1109/MilCIS.2015.7348942
20. Moustafa N., Turnbull B., Choo K. R. An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things. *IEEE Internet of Things Journal*, 2019, vol. 6, no. 3, pp. 4815–4830. doi:10.1109/JIOT.2018.2871719