МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное
учреждение высшего образования
**«Южно-Уральский государственный университет
(национальный исследовательский университет)»**
Высшая школа электроники и компьютерных наук
Кафедра системного программирования

РАБОТА ПРОВЕРЕНА

Рецензент
Доцент кафедры ИИТиМОИ
ФГБОУ ВО «ЮУрГГПУ», к.п.н.

_____ О.А. Дмитриева

«___»_____ 2024 г.

ДОПУСТИТЬ К ЗАЩИТЕ

Заведующий кафедрой, д.ф.-м.н.,
профессор

_____ Л.Б. Соколинский

«___»_____ 2024 г.

Разработка модели раннего обнаружения DDoS-атак на основе анализа сетевого трафика

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
ЮУрГУ – 09.04.04.2024.308-1469.ВКР

Научный руководитель,
доцент кафедры СП, к.п.н.
_____ О.Н. Иванова

Автор работы,
студент группы КЭ-229
_____ К.Е. Бакунина

Ученый секретарь
(нормоконтролер)
_____ И.Д. Володченко
«___»_____ 2024 г.

Челябинск, 2024 г.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное
учреждение высшего образования
«Южно-Уральский государственный университет
(национальный исследовательский университет)»
Высшая школа электроники и компьютерных наук
Кафедра системного программирования

УТВЕРЖДАЮ
Зав. кафедрой СП
_____ Л.Б. Соколинский
29.01.2024 г.

ЗАДАНИЕ
на выполнение выпускной квалификационной работы магистра
студентке группы КЭ-229
Бакуниной Ксении Егоровне,
обучающейся по направлению
09.04.04 «Программная инженерия»
(магистерская программа «Искусственный интеллект и инженерия данных»)

1. **Тема работы** (утверждена приказом ректора от 22.04.2024 г. № 764-13/12)
Разработка модели раннего обнаружения DDoS-атак на основе анализа сетевого трафика.
2. **Срок сдачи студентом законченной работы**: 20.05.2024 г.
3. **Исходные данные к работе**

3.1. Lopez A. Network Traffic Behavioral Analytics for Detection of DDoS Attacks. / A. Lopez, A. Mohan, S. Nair // SMU Data Science Review, 2019. – 25 p.

3.2. Mishra A. Prediction Approach against DDoS Attack based on Machine Learning Multiclassfier [Электронный ресурс] // arXiv.org, 2017. Дата обновления: 27.04.2022 г. URL: https://arxiv.org/abs/2204.12855 (дата обращения: 01.03.2024 г.).

3.3. Alduailij M. Machine-Learning-Based DDoS Attack Detection Using Mutual Information and Random Forest Feature Importance Method. / M. Alduailij,

Q. Khan, M. Tahir, M. Sardaraz // Cloud Computing and Symmetry: Latest Advances and Prospects, 2022. – 15 p.

4. Перечень подлежащих разработке вопросов

4.1. Провести анализ предметной области и аналогичных решений.

4.2. Провести анализ и выбор набора данных.

4.3. Определить архитектуру модели.

4.4. Обучить нейросетевую модель.

4.5. Оценить результаты работы полученной модели.

5. Дата выдачи задания: 29.01.2024 г.

Научный руководитель,
доцент кафедры СП, к.п.н. О.Н. Иванова

Задание принял к исполнению К.Е. Бакунина

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

В настоящее время фактически все сферы жизни человека зависят от информационных технологий и цифровых инструментов. Именно поэтому современные устройства и сеть Интернет хранят много персональной информации. Помимо личной информации о каждом отдельном человеке также существует корпоративная информация, которая поддерживает работу множества частных и государственных структур.

Вся эта информация может быть украдена посредством кибератаки. Кибератака – это набор действий, совершаемых злоумышленниками, которые пытаются получить несанкционированный доступ, украсть данные или нанести ущерб компьютерам, компьютерным сетям или другим вычислительным системам [1].

Даже одна успешная кибератака может нанести непоправимый ущерб как отдельной личности, так и обществу в целом. Одним из наиболее распространенных типов кибератак считается distributed denial-of-service (DDoS-атака) – распределенная атака типа «отказ в обслуживании».

Актуальность

Актуальность темы обусловлена активным ростом количества совершенных кибератак за последние годы.

Некоторые из основных статистических данных о DDoS-атаках за 2022–2023 годы [2]:

– согласно ежеквартальному отчету компании «Лаборатория Касперского», было зарегистрировано около 57 тысяч DDoS-атак;

– согласно данным компании «Cloudflare», в 2022 году количество DDoS-атак с целью выкупа выросло на 67%.

В настоящее время DDoS-атаки становятся все более технологически сложными, для их реализации используются инструменты искусственного интеллекта (ИИ) и машинного обучения (МО).

Постановка задачи

Целью выпускной квалификационной работы является разработка модели раннего обнаружения DDoS-атак на основе анализа сетевого трафика. Для достижения поставленной цели необходимо решить следующие задачи:

1) провести анализ предметной области и аналогичных решений;
2) провести анализ и выбор набора данных;
3) определить архитектуру разрабатываемой модели;
4) разработать и обучить нейросетевую модель;
5) оценить результаты работы полученной модели.

Структура и объем работы

Работа состоит из введения, четырех глав, заключения, списка литературы и пяти приложений. Объем работы составляет 63 страницы, объем списка литературы – 48 источников.

В первой главе описывается предметная область исследования и анализ аналогичных решений.

Вторая глава посвящена выбору и предобработке набора данных.

В третьей главе описывается реализация модели, а именно выбор архитектур и подбор гиперпараметров модели.

В четвертой главе представлены вычислительные эксперименты, проведенные для разработанных моделей.

В приложении А содержится таблица результатов анализа аналогичных решений. В приложении Б представлена таблица результатов анализа наборов данных. Приложение В содержит столбчатую диаграмму, отражающую значимость атрибутов выбранного набора данных. В приложении Г представлена сводная таблица метрик, полученных при тестировании разрабатываемой модели. Приложение Д содержит характеристики выбранного предобработанного набора данных.

1. АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ И АНАЛОГИЧНЫХ РЕШЕНИЙ

1.1. Анализ предметной области

Распределенная атака типа «отказ в обслуживании» (DDoS-атака) – это злонамеренная попытка нарушить нормальный трафик целевого сервера, службы или сети путем переполнения цели или окружающей ее инфраструктуры потоком интернет-трафика [3].

Согласно отчету Radware Global Application and Network Security Report 2016-173 и 2017-184 [4], типичная тактика, техника и процедура DDoS-атаки заключается во взломе и проникновении в данные жертвы и требовании выкупа в обмен на отказ от публичного обнародования похищенных данных.

Обнаружение DDoS-атак – это процесс отделения распределенных атак типа «отказ в обслуживании» от обычного сетевого трафика с целью эффективного подавления атак [5].

Принцип работы DDoS-атак [6]

Реализация распределенной атаки типа «отказ в обслуживании» основана на применении сетей компьютеров, подключенных к Интернету.

Сети состоят из компьютеров и других устройств, которые были заражены вредоносным программным обеспечением (ПО), что позволяет злоумышленнику удаленно управлять ими. Эти отдельные устройства называются ботами, а группа ботов называется ботнетом.

В результате атаки сервер или сеть жертвы становятся целью ботнета, каждый бот отправляет запросы на IP-адрес цели, что приводит к перегрузке, которая вызывает отказ в обслуживании обычного трафика.

DDoS-атака с точки зрения сетевого подключения

DDoS-атака происходит посредством сетевого подключения, которое описывается с помощью модели OSI (Open Systems Interconnection) – это

концептуальная модель, созданная Международной организацией по стандартизации, которая позволяет различным системам связи взаимодействовать с использованием стандартных протоколов [7].

Данная модель основана на концепции разделения системы связи на семь абстрактных уровней.

Многоуровневость архитектуры сетевого подключения говорит о том, что осуществление DDoS-атак может быть произведено на различных уровнях модели OSI.

Типы DDoS-атак

Выделяют несколько типов DDoS-атак в соответствии с различными уровнями сетевого подключения [8]:

– объемный;
– протокольный;
– прикладной.

Объемные DDoS-атаки

Объемные DDoS-атаки нацелены на сетевой уровень (уровень 3) модели OSI, фокусируясь на чрезмерной пропускной способности сети и инфраструктуре с большим объемом вредоносного трафика.

Концепция объемной атаки – отправить на сайт как можно больше трафика, чтобы перегрузить пропускную способность сервера. Обычно производятся с использованием методов усиления, например, таких как DNS-усиление.

Злоумышленники также создают объемные атаки, используя ботнеты, состоящие из устройств Интернета вещей. Таким устройствам обычно не хватает базовых средств защиты, но поскольку они подключены к Интернету и могут выполнять код, их можно легко взломать.

К объемным DDoS-атакам относятся: SYN-флуд, ICMP-флуд, UDP-флуд, атаки с усилением DNS и с усилением NTP.

DDoS-атака сетевого протокола

Данный тип атак нацелен на транспортный (уровень 4) и прикладной (уровень 7) уровни модели OSI.

Обычно эти атаки используют сценарии, в которых сервер получает пакет или запрос от компьютера и ожидает дальнейшего взаимодействия. Сервер выделяет память и ресурсы для поддержания состояния сеанса и канала связи, тем самым ненамеренно замедляя или останавливая обмен данными и истощая ресурсы.

К DDoS-атакам сетевого протокола относятся: атаки типа «человек посередине» (MitM), подмена DNS, атаки SYN Flood, атаки смурфов, взлом TCP/IP, IP-спуфинг, отравление ARP.

DDoS-атака прикладного уровня

Атаки на уровне приложений (уровень 7 модели OSI) направлены на кражу данных, отключение сетей, нарушение бизнес-процессов и вымогательство средств у компаний.

Данные атаки легко запустить, но трудно обнаружить и предотвратить, поскольку злоумышленники обычно отправляют запросы как законные пользователи.

Также DDoS-атака прикладного уровня может представлять собой многовекторную атаку, в которой используется комбинация объемных и протокольных атак для повышения вероятности отключения службы.

К DDoS-атакам прикладного уровня относятся: SQL-инъекция, межсайтовый скриптинг (XSS) и удаленное включение файлов (RFI).

Основные аспекты о типах DDoS-атак представлены в таблице 1.

Таблица 1 – Типы DDoS-атак

Тип	Метрика	Уровнь	Особенности	Примеры
Объемный	Гбит/с, бит/с	Сетевой (уровень 3)	Большой объем трафика, использование ботов и ботнетов IoT	ICMP-флуд, UDP-флуд NTP амплификация

Окончание таблицы 1

Тип	Метрика	Уровень	Особенности	Примеры
Протокольный	Пакетов в секунду (PPS)	Транспортный (уровень 4) и прикладной (уровень 7)	Истощение памяти и ресурсов сервера жертвы	MitM, SYN-флуд, подмена DNS
Прикладной	Запросов в секунду (RPS)	Прикладной (уровень 7)	Несложная реализация, трудно предотвратить и смягчить	XSS, SQL-инъекция, RFI

Сетевое подключение

Данные, передаваемые по сети, делятся на пакеты – единицы данных на сетевом уровне в модели OSI, которые повторно объединяются устройствами назначения. Разделение данных на пакеты позволяет сети управлять различными полосами пропускания, маршрутами и несколькими подключенными устройствами, которые обмениваются данными и получают пакеты независимо друг от друга. Это упрощает повторную передачу потерянных или прерванных фрагментов данных.

Большинство сетевых пакетов разделены на три части, которые содержат определенный набор атрибутов (рисунок 1) [9].

Рисунок 1 – Строение сетевого пакета

Признаки обнаружения DDoS-атаки

Объективными признаками совершения DDoS-атаки являются замедление или недоступность служб.

Также можно рассматривать другие более специфические признаки [10], основанные на строении сетевого пакета:

1) подозрительные объемы трафика, исходящие с одного IP-адреса или диапазона IP-адресов;

2) поток трафика от пользователей с одним и тем же поведенческим профилем: тип устройства, геолокация, версия веб-браузера;

3) большое количество запросов к одной странице или конечной точке (endpoint);

4) неестественные модели трафика;

5) другие специфические признаки.

В процессе детектирования проводится анализ аномалий сетевого трафика – ведется поиск отклонений от контрольных характеристик трафика в штатных условиях работы сети.

Классификация подходов обнаружения DDoS-атаки

В общем случае классификация систем обнаружения DDoS-атак аналогична классификации систем обнаружения вторжений.

Обнаружение атак на основе сигнатур

Детектирование на основе сигнатур используется для известных типов атак. Для обнаружения атаки не требуется какое-либо описание типичных действий при ней, однако для этих видов атак необходима база данных с известными сигнатурами атак.

Чтобы протестировать эффективность системы обнаружения вторжений на основе сигнатур, необходимо создать базу данных из сотен или даже тысяч сигнатур, каждая из которых требует свой отдельный раздел. Процесс сопоставления каждого пакета с этой базой данных может потребовать значительных ресурсов и занять всю доступную пропускную способность.

Обнаружение атак на основе аномалий

Методы обнаружения вторжений, основанные на противоречивости (аномалиях), распознают необычную активность и создают предупреждения аномалий в действиях системы или действиях приложений.

Наибольшие трудности в использовании методов обнаружения на основе аномалий заключаются в определении типичного поведения системы,

выборе предела для срабатывания предупреждения и предотвращении ложных предупреждений.

Существуют базовые встроенные решения обнаружения DDoS-атак сетевых устройств, к ним относятся: балансировщики нагрузки, брандмауэры или системы предотвращения вторжений. Данные решения являются базовыми и универсальными, что делает их неустойчивыми для большинства новых типов атак.

На сегодняшний день основным способом оперативного обнаружения DDoS-атак являются специализированные методы, направленные исключительно на определения вредоносного трафика DDoS-атак.

Обнаружение DDoS-атаки с использованием ИИ и МО

В настоящее время злоумышленники стали использовать технологии ИИ и МО для совершения кибер-атак. В связи с этим наиболее логичным решением будет использование этих технологии для раннего обнаружения.

Для определения незаконного трафика необходимо изучить характеристики подлинных пакетов, генерируемых настоящими приложениями, сравнить их с поддельными пакетами, генерируемыми атакующими инструментами.

Результаты сравнения будут выступать в качестве паттернов для отделения трафика, которые в дальнейшем будут представлены в качестве входных переменных для обучения искусственной нейронной сети (ИНС).

Для обучения используются шаблоны заголовков пакетов, которые включают адреса источников, идентификаторы и порядковые номера в сочетании с номерами портов источника и назначения.

1.2. Современные подходы обнаружения DDoS-атак

В настоящее время существует ряд работ, решающих задачи обнаружения и классификации DDoS-атак [11–18], направленных на разработку модели ИНС на основе следующих архитектур: ANN, Random Forest (RF),

Logistic Regression (LG), K-Nearest Neighbors (KNN), Naive Bayes (NB), Multi-Layer Perceptron (MLP), Support vector machine (SVM).

Для дальнейшего анализа необходимо сформулировать критерии оценки исследований в рамках предметной области:

1) цель исследования – обнаружение, классификация DDoS-трафика;

2) входные данные – данные, используемые для обучения модели нейронных сетей;

3) метод классификации – метод, используемый для обнаружения и классификации DDoS-атак;

4) результат классификации – оценка результатов обнаружения и классификации DDoS-атак.

Perakovic, D. Model for Detection and Classification of DDoS-Traffic Based on Artificial Neural Network [11]

Цель исследования

Целью данного исследования является разработка модели системы на основе ИНС для обнаружения DDoS-трафика и его классификации с целью повышения точности обнаружения определенных классов DDoS-трафика и применения соответствующих методов защиты.

Гипотеза этого исследования заключается в том, что с извлеченными параметрами собранного трафика и внедрением ИНС можно с высокой точностью классифицировать DDoS-трафик на новом наборе данных.

Набор данных

Поскольку целью данного исследования является обнаружение и классификация DDoS-трафика, были выделены три класса DDoS-трафика: UDP, DNS, CharGen – и один класс легитимного сетевого трафика. Данные, используемые в этом исследовании, были собраны из онлайн источников. Были задействованы четыре общедоступных набора данных, из которых был создан набор данных из 4 986 записей сетевого трафика. Набор входных данных представляет собой предварительно созданную матрицу, содержа-

щую выборку из 4986 экземпляров со значениями пяти выбранных атрибутов [5x4986] и матрицу [4x4986], которая содержит значения 0 или 1 в зависимости от классификации конкретного класса трафика.

Параметрами, используемыми для классификации, являются время прибытия пакета, исходный IP-адрес, целевой IP-адрес, используемый протокол и длина пакета.

С целью использования данных для классификации DDoS-атак была проведена нормализация и классификация данных для получения значений всех выявленных параметров во взаимном соотношении.

Модель

Представленная в работе архитектура соответствует многослойному восприятию (MLP) – типу ИНС, у которого входные сигналы представлены набором входных данных скрытого слоя, выходного слоя и входного слоя. Скрытый слой имеет 50 нейронов, которые по сравнению с другими комбинациями показали себя лучше всего.

Чистая сумма весов представляет собой входные данные для расчета передаточной функции f(net). Передаточная функция является сигмоидной или логистической функцией. Результат сигмоидной передаточной функции в скрытом слое представляет собой ввод в выходной слой. Внутри выходного слоя использовалась передаточная функция softmax.

Результат вывода представляет один из четырех определенных классов трафика DDoS-атаки.

Результаты исследования

В данном исследовании было проведено моделирование разработанной модели ИНС с различным количеством нейронов в скрытом слое (30, 35, 40, 45, 50 и 55). Наилучшие результаты в обнаружении нелегитимного трафика и его классификации были получены при 50 нейронах в скрытом слое, точность классификации составила 95,6%.

Saied, A. Artificial Neural Networks in the Detection of Known and Unknown DDoS-Attacks: Proof-of-Concept [12]

Цель исследования

Целью исследования является обнаружение известных и неизвестных DDoS-атак с последующим созданием защитного механизма, который предотвращает попадание поддельных пакетов к жертве, но пропускает настоящие пакеты. Также авторы данного исследования ставят перед собой несколько второстепенных целей:

1) обучение, развертывание и тестирование решения в реалистичной физической среде;

2) снижение силы атаки до того, как она достигнет жертвы;

3) оценивание используемого подхода с помощью старых и современных наборов данных в сравнении с соответствующими работами на основе точности, чувствительности, специфичности.

Набор данных

Выбор шаблонов для входных данных основан на идеи создания новых сетевых инфраструктур в корпоративной и изолированной среде с различными типами DDoS-атак, запущенных на разных уровнях.

Наборы данных были организованы и структурированы для размещения подлинных и атакующих шаблонов в квалифицированном формате, который принимает симулятор Java Neural Network Simulator (JNNS) для дальнейшей нормализации.

Модель

Механизм обнаружения основан на архитектуре ANN. Для обучения использовались шаблоны заголовков пакетов, которые включают адреса источников, идентификаторы и порядковые номера в сочетании с номерами портов источника и назначения. Авторы изучили характерные особенности подлинных пакетов, генерируемых настоящими приложениями, сравнили их с поддельными пакетами, генерируемыми атакующими инструментами, и представили их в качестве входных переменных для обучения ИНС.

Типичная ИНС состоит из входного, скрытого и выходного слоев, где шаблоны поступают в алгоритм обучения через входные узлы. Входные значения представляют собой характерные шаблоны, которые отделяют атаки от подлинного трафика. Авторами были выбраны три топологические структуры ИНС, каждая из которых имеет три слоя.

Выбор соответствующего алгоритма обучения или количества скрытых узлов и функции активации был основан на первоначальных экспериментах, в которых sigmoid и обратное распространение дали наиболее точные результаты. Сравнение проводилось между QuickProp, Backpropagation, Backprop Weight Decay, Backprop thru time, а в качестве функций использовались sigmoid, Elliott, SoftMax, BAM. Эксперименты показали, что Backpropagation в сочетании с сигмоидной функцией активации и выбранной топологической структурой может обеспечить точность до 98,6%.

Интеграция всех ИНС в одно приложение, а не в отдельные экземпляры, может привести к отсутствию доступности в случае технического сбоя системы. Если один экземпляр не работает (например, экземпляр, обнаруживающий ICMP-атаку), другие будут по-прежнему доступны для обнаружения атак, связанных с UDP и TCP. Введение экземпляров ANN для каждого протокола обеспечивает лучшее обслуживание и дополнительный контроль для обучения и анализа алгоритма.

Результаты исследования

В ходе экспериментов были запущены известные и неизвестные DDoS-атаки, каждая из которых состояла из 80–90 зомби-машин, формирующих в общей сложности 60 раундов TCP, UDP и ICMP DDoS-атак на целевой объект, наряду с 60 раундами обычного трафика.

Точность обнаружения данного решения составила 98%, где 50% – обнаружение известных атак, а 48% – обнаружение неизвестных DDoS-атак. Дополнительно была оценена точность обнаружения при тестировании против DDoS-атак с низкой и высокой скоростью: 97,4% и 98% соответственно.

Li, Y. DDoS-attack detection method based on feature extraction of deep belief network [13]

Цель исследования

Целью данного исследования является обнаружения DDoS-атак на основе извлечения признаков глубокой сети убеждений (deep belief network).

Модель

Для исследования обнаружения DDoS-атак в данной статье предлагается метод обнаружения DDoS-атак на основе извлечения признаков сети глубоких убеждений и модели Long short-term memory (LSTM).

Сначала этот метод извлекает особенности IP-пакетов с помощью глубокой сети убеждений, затем создает модель прогнозирования трафика LSTM и, наконец, идентифицирует DDoS-атаки на основе созданной LSTM модели. Модель может точно предсказать тенденцию нормального сетевого трафика, идентифицировать аномалии, вызванные DDoS-атаками, и применяться для решения других методов обнаружения DDoS-атак в будущем.

Результаты вычислительных экспериментов и описание используемого в работе датасета приведены не были.

Zeinalpour, A. Addressing the Effectiveness of DDoS-Attack Detection Methods Based on the Clustering Method Using an Ensemble Method [14]

Цель исследования

Целью данного исследования было определить, является ли использование методов фильтра и обертки, помещаемых перед комбинированными алгоритмами кластеризации с использованием метода классификатора Vote, эффективным для снижения частоты ложных срабатываний.

Набор данных

Для данного исследования был использован набор данных сетевого трафика CIC-DOS2017, включающего в себя данные о DDoS-атаках и обычном трафике.

Чтобы подготовить этот набор данных, авторы исследования применили метод шести процедур очистки данных:

1) ручное удаление атрибутов шести характеристик: идентификатор потока, IP-адрес источника, IP-адрес назначения, порт источника, порт назначения и метка времени;

2) очистка числовых данных с помощью процедуры NumericCleaner для дальнейшей нормализации;

3) нормализация данных с использованием процедуры min-max для установки значений атрибутов между 0 и 1;

4) метод EM Imputation для замены отсутствующего значения атрибута Flow Bytes среди четырех экземпляров данных;

5) коррекция несбалансированных экземпляров данных DDoS-атаки и легитимный трафик с помощью процедуры Spreadsubsample;

6) рандомизация данных о сетевом трафике.

Модель

В данном исследовании авторы использовали концепцию группы этапов A-B-A-BC:

– на этапе A использовался весь набор данных CIC-DOS2017 для оценки комбинации алгоритмов кластеризации с использованием метода Vote;

– на этапе B применили метод фильтра для отбора признаков для оценки комбинации алгоритмов кластеризации;

– на этапе BC добавили метод обертки после метода фильтра при выборе признаков для оценки.

На рисунке 2 представлена архитектура проектируемой модели для классификации трафика.

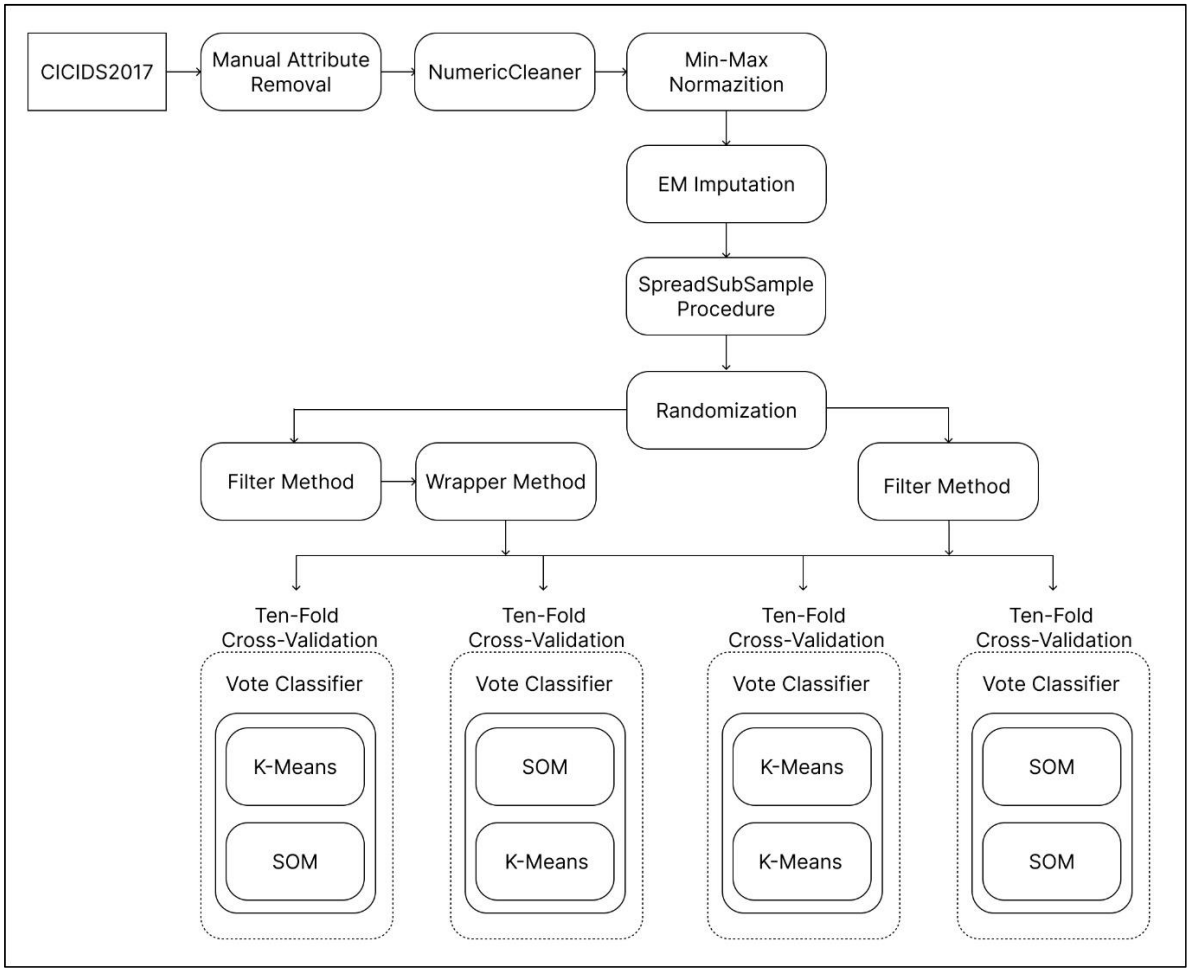

Рисунок 2 – Архитектура проектируемой модели
для классификации трафика

Результаты исследования

В данном исследовании авторам удалось достичь самого низкого показателя ложных срабатываний в 0,012% для методов обнаружения DDoS-атак в двух экземплярах сигнатур сетевого трафика путем включения хи-квадрат и J48 до комбинированных алгоритмов кластеризации с использованием K-means и SOMs.

David, J. DDoS-Attack Detection using Fast Entropy Approach on Flow-Based Network Traffic [15]

Цель исследования

Целью данного исследования является обнаружение DDoS-атак с использованием подхода быстрой энтропии в потоке.

Набор данных

Эксперименты исследования проводились на подмножестве набора данных CAIDA. Предлагаемый метод основан на анализе потока, для которого требуется только информация о заголовке пакета.

При агрегации потоков информация заголовка была агрегирована за определенный интервал времени, который принадлежит пяти одинаковым кортежам (IP-адрес источника, IP-адрес назначения, порт источника, порт назначения, номер протокола).

В исследовании вместо анализа набора данных на основе захваченной информации заголовка пакета для быстрого вычисления энтропии используется счетчик потока каждого соединения в наборе данных. Счетчик потоков – это свойство, показывающее серьезность атаки переполнения. В данном случае это было сделано для увеличения скорости анализа.

Модель

Предлагаемый метод обнаружения DDoS-атак основан на трех целях:

1) агрегация потоков для обнаружения атак на основе потоков;

2) быстрое вычисление энтропии для обнаружения DDoS-атаки с меньшим временем вычисления;

3) адаптивный пороговый алгоритм для повышения точности обнаружения.

Детектирование DDoS-атак характеризуется тем, что есть ли разница между быстрой энтропией и средним пороговым значением. Этот адаптивный пороговый алгоритм повышает точность обнаружения, а быстрое вычисление энтропии сокращает время вычислений по сравнению с обычным вычислением энтропии.

Алгоритм обнаружения показан на рисунке 3.

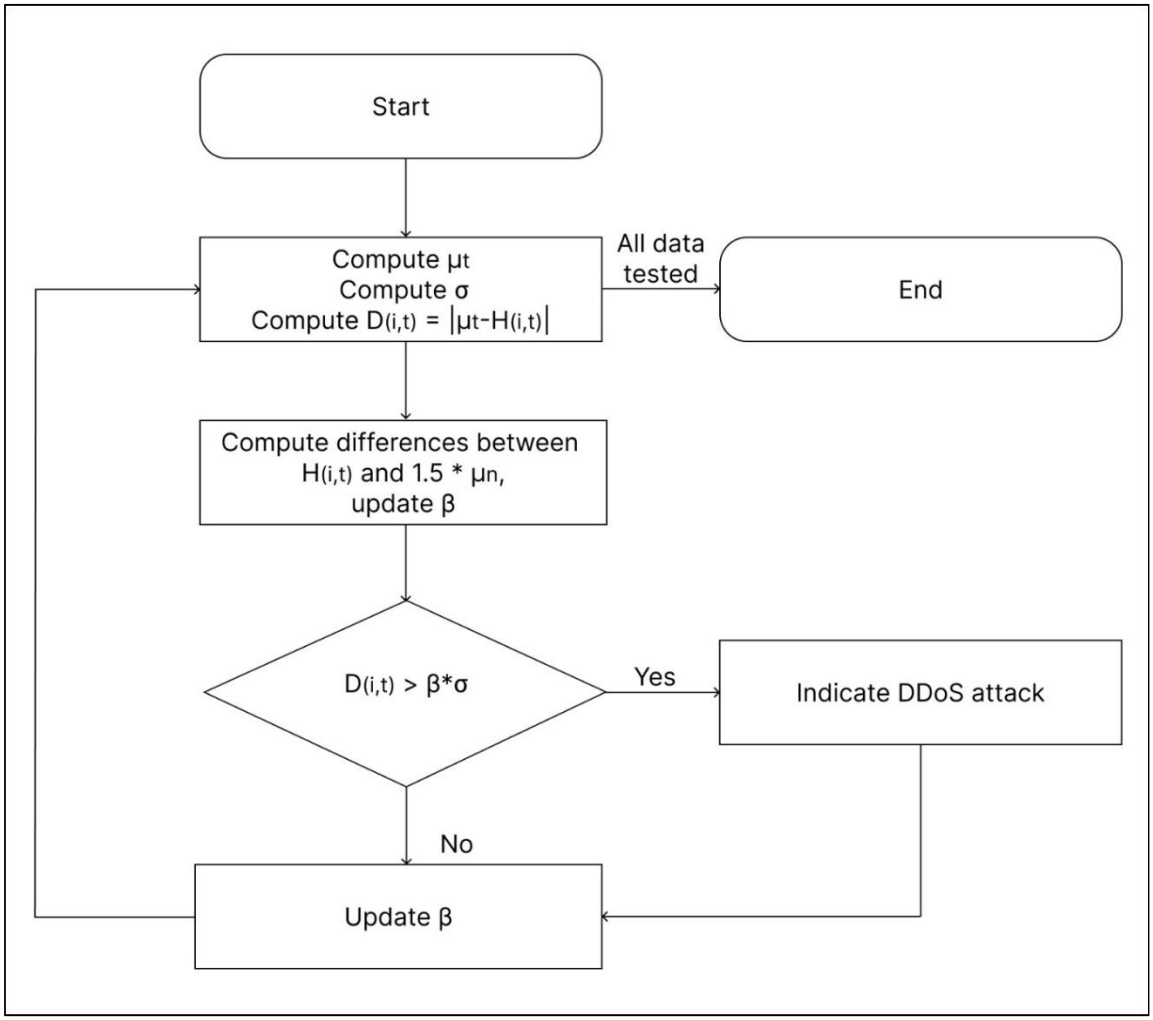

Рисунок 3 – Алгоритм обнаружения DDoS-атак

Lopez, A. Network Traffic Behavioral Analytics for Detection of DDoS-Attacks [16]

Цель исследования

Цель данного исследования – построить общую статистическую модель после сравнительного анализа различных алгоритмов машинного обучения, которая может быть использована для обнаружения сетевых аномалий. Анализатор потока сетевого трафика может быть использован для сбора информации о потоке.

Набор данных

Для данного исследования был использован набор данных CIC-DOS2017. Разведочный анализ данных (EDA) был выполнен для пони-

мания данных, захваченных в потоке трафика, а характеристики были сгенерированы анализатором сетевого трафика CICFlowMeter, в результате этого было получено 85 признаков с помощью методов рекурсивного устранения признаков (RFE).

Модель

В данном проекте авторы отдельно рассматривали несколько методов классификации данных и запускали несколько основных моделей классификации. Были использованы следующие алгоритмы: MLR, KNN, RF, NB, MLP и Dense Neural Networks.

Результаты исследования

В результате проведения исследования была получена сводная таблица (таблица 2), которая иллюстрирует результаты accuracy, precision и training time при различных методах классификации.

Таблица 2 – Результаты классификации для рассматриваемых методов

Model	Overal Accuracy	Overall Precision	Training Time (sec)
MLR	0,88	0,89	364
KNN	0,95	0,66	63
RF	0,99	0,99	88
NB	0,26	0,66	188
MLP	0,77	0,77	116
DNN	0,97	0,98	523

Mishra, A. Prediction Approach against DDoS-Attack based on Machine Learning Multiclassfier [17]

Цель исследования

Цель авторов – обеспечить лучшее понимание комплексной структуры для характеристик уязвимостей в информационных системах, таких как категория уязвимости, к которой относится конкретная уязвимость, потенциальные угрозы и предупреждающие признаки приближения к ней для получения помощи в устранении проблемы.

Набор данных

CICDoS-2019 – это набор данных, предоставленный CIC, который предназначен для предоставления достоверной информации для обучения алгоритмов обнаружения DDoS-атак.

Набор данных содержит достаточно информации для обучения и проверки модели, поскольку он включает в себя в общей сложности 5 775 786 строк информации в файлах с расширением CSV. Всего в наборе данных CICDoS имеется 88 атрибутов.

В данном исследовании авторы использовали классификатор Extra Tree Classifier для выбора лучших 20 признаков. В таблице 3 представлены выбранные признаки и их описание.

Таблица 3 – 20 лучших признаков и их описание из классификатора Extra Tree Classifier

Признак	Описание
Timestamp	Время сеанса
Min Packet Length	Минимальная длина пакета
Source Port	Исходный порт, используемый сеансом
Protocol	Протокол, используемый сеансом
Packet Length Mean	Средняя длина пакета
Avg Fwd Segment Size	Средний размер, наблюдаемый в прямом направлении
Fwd Packet Length Min	Минимальный размер пакета в прямом направлении
Average Packet Size	Средний размер пакета
Fwd Packet Length Max	Максимальный размер пакета в прямом направлении
Max Packet Length	Максимальная длина пакета
Fwd Packet Length Mean	Средний размер пакета в прямом направлении
Flow Id	Внутренний числовой идентификатор, применяемый к каждому потоку
Total Length of Fwd Packets	Общий размер пакета в прямом направлении
Init_Win_bytes_forward	Общее количество байт, отправленных в начальном окне в прямом направлении
Down/Up Ratio	Коэффициент загрузки и выгрузки
Destination Port	Порт назначения, используемый сеансом
Source Ip	IP-адрес источника исходного сеанса
Fwd Header Length	Длина заголовка в прямом направлении
Min_seg_size_forward	Минимальный размер сегмента в прямом направлении

Модель

Для сравнения были выбраны четыре различные модели контролируемого обучения. Для выбора модели используется несколько критериев, которые включают наличие как параметрических, так и непараметрических моделей, использование алгоритмов из нескольких категорий, а также применение моделей, которые широко использовались в предыдущих исследованиях и публикациях.

Были использованы методы Random Forest, Decision Tree, Naive Bayes и Support Vector Machine.

Результаты исследования

В результате исследования была создана сводная таблица 4 с показателями классификации: Recall (1), Precision (2) и F1-Score (3) – для методов классификации.

Таблица 4 – Результаты классификации для методов классификации RF, DT, NB, SVM

	Random Forest			**Decision Tree**			**Naive Bayes**			**SVM**		
Показатель	1	2	3	1	2	3	1	2	3	1	2	3
BENIGN	100	100	100	40	50	50	77	100	87	100	100	100
DDoS_ LDAP	100	100	100	97	99	98	98	99	99	99	99	99
DDoS_ MSSQL	100	100	100	100	100	100	100	100	100	100	100	100

Alduailij, M. Machine-Learning-Based DDoS-Attack Detection Using Mutual Information and Random Forest Feature Importance Method [18]

Цель исследования

Цель исследования – разработка метод обнаружения DDoS-атак, используя различные методы отбора признаков и машинного обучения.

Набор данных

В данной работе были использованы наборы данных CIC-DOS2017 и CIC-DDoS2019, взятые с соответствующих официальных веб-сайтов.

Из набора данных CIC-DOS2017 эксперименты проводились с файлом журнала сетевого трафика за вечер пятого дня, содержащим 225 711 экземпляров и 79 признаков, включая метку класса.

Из набора данных CIC-DDoS2019 был выбран файл DrDoS_NTP, содержащий 1 209 961 экземпляров и 84 входных признаков. Атрибут класса представляет собой бинарную метку класса.

Для указанных наборов данных была применена предобработка: преобразование категориальной метки класса в дискретную форму (0,1) с применением кодирования метки, где 0 – это доброкачественный класс, а 1 – DDoS-атака.

Для выбора наиболее релевантных признаков из имеющихся были использованы методы RFFI и MI, с целью нахождении лучшего метода отбора признаков из методов на основе фильтров и встроенных методов.

Ниже перечислены основные цели отбора признаков:

1) улучшение эффективности обобщения по сравнению с моделью со всеми характеристиками;

2) обеспечение более надежного обобщения и более быструю реакцию на невидимые данные;

3) получение более простого понимания процесса генерации данных.

Подход, основанный на отборе признаков, используется как шаг предварительной обработки, в регрессии и классификации.

Модель

В предлагаемой работе используется логистическая регрессия, она применяется к выбранным признакам для обнаружения DDoS-атак.

Также в данной работе используется классификация KNN с параметром K равным 2 и метрикой расстояния Минковского.

Градиентное усиление – один из самых популярных алгоритмов прогнозирования в машинном обучении. Для регулирования эволюции дерева решений алгоритма используются различные специальные параметры, которые управляют размером дерева и величиной веса (таблица 5).

Таблица 5 – Параметры, используемые для обучения GB

Параметр	Значение
learning rate	0,5
Max depth	4
Max	2
Min samples split	2
Random state	0
N-estimators	19
Min samples leaf	1

Также в данной работе был использован алгоритм классификации случайный лес, параметры которого приведены в таблице 6.

Таблица 6 – Параметры, используемые для обучения RF

Параметр	Значение
Bootstrap	True
Criterion	Gini
Min samples split	2
N estimators	30
Random state	0
Max features	Auto
Min samples leaf	1

Также рассматривался подход WVE – это репрезентативный подход для объединения прогнозов в парной классификации, в которой классификаторы не считаются равными. На оценочном множестве D каждому классификатору присваивается весовой коэффициент, который обычно равен его точности классификации. В данной работе в качестве базовых методов используются KNN, RF и DT, которые предсказывают DDoS-атаку путем объединения результатов с WVE.

Результаты исследования

Общая точность предсказания случайного леса с 16 признаками составляет 0,99993, а с 19 признаками – 0,999977, что лучше других методов. Сделан вывод, что RF, GB, WVE, KNN и LR достигают хороших результатов, используя MI и RFFI в качестве методов отбора признаков. Также ав-

торы предлагают в будущем для обнаружения DDoS-атак и других атак использовать методы выбора признаков, например, последовательный выбор признаков.

В качестве результата анализа аналогичных решений была сформирована сводная таблица с характеристиками работ в соответствии с критериями, определенными перед анализом, представленная в форме таблицы 1 приложения А.

Исходя из данных, полученных при анализе, можно сделать вывод, что наиболее результативным и используемым методом классификации является Random Forest с показателями классификации на уровне 99%. Также, исходя из этого, можно выделить LSTM и Gradient Boosting, результаты которых не были представлены в анализе.

1.3. Анализ работ по раннему обнаружению DDoS-атак

Поскольку цель работы заключается в разработке модели раннего обнаружения DDoS-атак, необходимо рассмотреть ряд работ, фокусирующихся на раннем обнаружении.

Сформулируем пункты, по которым будем рассматривать работы:

1) параметр t – время до обнаружения DDoS-атаки;

2) методы – методы сокращения времени до обнаружения DDoS-атаки;

3) результаты обнаружения – оценка результатов обнаружения DDoS-атаки при различных значениях параметра t.

Xylogiannopoulos, K. Early DDoS Detection Based on Data Mining Techniques [19]

Цель

Целью данной статьи является разработка инновационного метода обнаружения DDoS-атак, который сочетает в себе аномалии и обнаружение закономерностей.

К данным, полученным в сети, применяется разработанная авторами методика интеллектуального анализа данных, позволяющая идентифицировать все повторяющиеся шаблоны последовательности. Если с помощью этого метода обнаруживаются несколько IP-адресов из одного домена, может произойти потенциальная DDoS-атака.

Методы

Метод, предложенный в этой статье, основан на структуре данных Suffix Array, используется для обнаружения всех повторяющихся шаблонов в последовательности.

Первый шаг – преобразовать найденные IP-адреса в реальные строки, которые будут использоваться для обнаружения повторяющихся шаблонов.

Второй шаг – отсортировать все IP-адреса в алфавитном порядке. Это необходимо для того, чтобы алгоритм ARPaD мог выполнить анализ, так как строки поступили непосредственно из структуры данных суффиксного массива.

Последний шаг – выполнение алгоритма ARPaD в отсортированном массиве строк IP-адресов и получение всех повторяющихся подстрок (IP-префиксы доменов или подсетей) или строк (полные IP-адреса).

Имея результаты, администратор сети может установить порог возникновения значения повторяющихся IP-адресов в зависимости от типа анализа, количества или времени попаданий.

Результаты обнаружения

Всего было проведено три крупных эксперимента с разным количеством обращений (100 000, 500 000 и 1 500 000 IP-адресов).

Для первого эксперимента алгоритм ARPaD запустился 15 раз, для второго – 3 раза, для третьего – 1 раз.

В результате выполнения экспериментов можно сделать вывод, что время необходимое алгоритму ARPaD для обнаружения всех повторяющихся шаблонов:

– в первом эксперименте – в среднем примерно 1,1 секунду, включая время процесса сортировки;

– в втором эксперименте – в среднем 4,3 секунды, для определения их положения на временной шкале требуется примерно 10,1 секунд;

– в третьем эксперимент – 43 секунды, в то время как время единичного обнаружения составляет 12 секунд.

Время, необходимое алгоритму ARPaD для обнаружения DDoS-атаки, оказалось линейным, в среднем сохраняется соотношение 1/10 от общего времени DDoS-атаки для простого обнаружения шаблонов и примерно 1/4 от общего времени DDoS-атаки для полного обнаружения шаблонов.

По результатам экспериментов время, необходимое для выявления запуска DDoS-атаки, составляет от 1,1 до 43 секунд в зависимости от исходных параметров, предоставленных алгоритму.

Методом уменьшения времени обнаружения DDoS-атаки является уменьшение времени между запусками алгоритма ARPaD или увеличение частоты запусков данного алгоритма, что показывает серия экспериментов, описанных выше.

Также был рассмотрен ряд работ [20–22], в заглавии которых заявлено ранее обнаружение DDoS-атак, но в тексте не объявлен параметр времени и не дано определения, что является ранним обнаружением.

Исходя из рассмотренных работ, можно сделать вывод, что формального определения, что является ранним обнаружением, не заявлено, методик расчета параметров раннего обнаружения также не представлено.

В связи с этим будем считать ранним обнаружением факт обнаружения атаки по прошествии 1–4 секунд после начала моделируемой атаки.

2. НАБОР ДАННЫХ ДЛЯ ОБУЧЕНИЯ МОДЕЛИ

2.1. Анализ датасетов

Далее проведем анализ наборов данных для обучения модели, определив критерии анализа:

1) количество атрибутов – количество столбцов датасета;
2) размер – количество строк в датасете;
3) виды DDoS-трафика – типы реализуемых DDoS-атак;
4) лицензия – лицензия, под которой распространяется этот набор данных (доступность датасета);
5) формат – формат файла датасета.

В соответствии с этими критериями, проведем дальнейший анализ.

Набор данных для оценки DDoS (CIC-DDoS2019) [23]

Набор данных CIC-DDoS2019 – это набор данных, специально разработанный для оценки и анализа DDoS-атак. Он содержит данные сетевого трафика, собранные из реалистичной и контролируемой среды. Данный набор данных создан Канадским институтом кибербезопасности (CIC) при Университете Нью-Брансуика.

Его первоначальные авторы: Иман Шарафальдов, Сакиб Хакак, Араш Хабиби Лашкари, Али Горбани.

Набор данных CIC-DDoS2019 состоит из образцов как безопасного, так и вредоносного трафика. Он включает в себя различные типы DDoS-атак, такие как TCP-флуд, UDP-флуд и ICMP-флуд, а также атаки прикладного уровня, такие как HTTP-атаки и SYN-флуд.

При создании датасета авторы использовали предложенную ими систему B-профиля для профилирования абстрактного поведения человеческих взаимодействий и генерации натуралистичного доброкачественного фонового трафика на предлагаемом испытательном стенде. Для этого набора данных было построено абстрактное поведение 25 пользователей на основе протоколов HTTP, HTTPS, FTP, SSH.

Набор данных предоставляет полный набор характеристик для каждого потока трафика, включая статистику сетевых потоков, информацию протокола транспортного уровня, информацию протокола прикладного уровня и различные статистические функции, извлеченные из полезной нагрузки. Эти характеристики включают IP-адреса источника и назначения, номера портов, размеры пакетов, временные метки и другие соответствующие атрибуты.

Набор данных CIC-DDoS2019 является общедоступным, состоит из 89 атрибутов и 431 327 строк, предоставляется в форматах CSV и ARFF.

Набор данных для оценки DoS (CIC-DoS2017) [24]

Поскольку DDoS-атаки является частным случаем DoS-атак, рассмотрим данный набор данных, содержащий в том числе данные DDoS-атак.

Набор данных CIC-DoS2017 – это набор данных о сетевом трафике, специально разработанный для изучения атак типа «отказ в обслуживании» (DoS).

Набор данных также, как и предыдущий, он разработан Канадским институтом кибербезопасности (CIC) при Университете Нью-Брансуика.

Его первоначальные авторы: Хоссейн Джази, Уго Гонсалесу, Наталья Стаханова, Али Горбани [25].

Набор данных содержит большое количество потоков сетевого трафика, как обычного, так и вредоносного, генерируемых в контролируемой среде. Он включает в себя различные типы DDoS-атаки, такие как TCP SYN-флуд, UDP-флуд и HTTP-флуд.

Каждый поток сетевого трафика в наборе данных представлен набором характеристик, таких как IP-адреса источника и назначения, порты источника и назначения, тип протокола, продолжительность потока, количество пакетов и байтов.

Набор данных CIC-DoS2017 содержит 77 атрибутов и 2 830 743 строк, при этом является общедоступным.

Набор данных DDoSDB [26]

Набор данных DDoSDB представляет собой набор данных сетевого трафика, связанных с атаками распределенного отказа в обслуживании (DDoS). Он содержит различные атрибуты и функции, которые предоставляют информацию о характеристиках DDoS-атак.

Каждая запись в наборе данных DDoSDB представляет собой конкретную DDoS-атаку и включает следующую информацию:

1) IP-адрес источника;
2) IP-адрес назначения;
3) тип DDoS-атаки;
4) время начала атаки;
5) время окончания атаки;
6) количество пакетов, задействованных в атаке;
7) количество байтов трафика, сгенерированного атакой;
8) продолжительность атаки в секундах;
9) сетевой протокол, используемый для атаки;
10) номер целевого порта на IP-адресе назначения.

Набор данных DDoSDB можно использовать для разработки и оценки методов обнаружения и смягчения последствий DDoS-атак, а также для изучения характеристик и моделей DDoS-атак в реальных сценариях.

CSE-CIC-IDS2018 [27]

Набор данных CSE-CIC-IDS2018 – это общедоступный набор данных, который обычно используется в области сетевой безопасности для исследования систем обнаружения вторжений. Он является совместным проектом Управления безопасности коммуникаций (CSE) и Канадского института кибербезопасности (CIC).

Набор данных содержит данные о сетевом трафике, полученные из реальной сетевой среды с различными типами атак и обычной деятельностью. Он включает в себя как нормальный трафик, так и различные типы кибератак, такие как DoS, DDoS, сканирование и проникновение.

Набор данных CSE-CIC-IDS2018 состоит из нескольких файлов CSV, каждый из которых представляет определенный атрибут сетевого трафика:

1) IP-адрес источника;
2) IP-адрес пункта назначения;
3) номера портов;
4) типы протоколов;
5) количество пакетов;
6) количество байтов;
7) информацию о времени.

Всего в датасете 80 атрибутов и 1 048 576 строк.

В качестве результата анализа была сформирована сводная таблица 2 приложения Б с характеристиками наборов данных в соответствии с критериями, определенными перед анализом.

Исходя из данных, полученных при анализе наборов данных, можно сделать вывод, что наиболее подходящим набором данных для обучения модели является CIC-DDoS2019. Так как данный датасет был разработан специально под анализ DDoS-атак, содержит наибольшее количество данных по этому виду атак среди рассмотренных наборов данных, а также является часто используемым в подобных работах по обнаружению.

2.2. Предобработка выбранного датасета

Прежде чем использовать выбранный набор данных для обучения модели, необходимо провести предобработку данных. Стандартные процедуры предобработки данных включают:

1) удаление дубликатов;
2) обработку отсутствующих значений (пропусков);
3) нормализацию данных;
4) кодирование категориальных признаков;
5) масштабирование признаков;
6) удаление выбросов.

Набор данных CIC-DDoS2019 представляет собой 15 отдельных файлов формата CSV. Для дальнейшей работы был выбран файл «Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv».

Выбранный файл содержит 85 атрибутов и 225 745 строк, из которых 128 027 строк DDoS-трафика и 97 718 строк легитимного трафика. Данные файла охватывают промежуток времени продолжительностью в 90 минут (07.07.2017 г. с 03:30 по 05:02), в течение которого происходит 22 DDoS-атаки, данные о которых представлены в таблице 7.

Таблица 7 – Данные о DDoS-атаках файла в датасете CIC-DDoS2019

№	Время начала	Время окончания	Количество записей
1	07.07.2017 03:56	07.07.2017 03:56	2 438
2	07.07.2017 03:57	07.07.2017 03:57	8 244
3	07.07.2017 03:58	07.07.2017 03:58	6 576
4	07.07.2017 03:59	07.07.2017 03:59	6 607
5	07.07.2017 04:00	07.07.2017 04:00	6 533
6	07.07.2017 04:01	07.07.2017 04:01	6 977
7	07.07.2017 04:02	07.07.2017 04:02	6 889
8	07.07.2017 04:03	07.07.2017 04:03	6 778
9	07.07.2017 04:04	07.07.2017 04:04	6 364
10	07.07.2017 04:05	07.07.2017 04:05	6 273
11	07.07.2017 04:06	07.07.2017 04:06	5 954
12	07.07.2017 04:06	07.07.2017 04:06	3
13	07.07.2017 04:07	07.07.2017 04:07	6 053
14	07.07.2017 04:08	07.07.2017 04:08	6 305
15	07.07.2017 04:09	07.07.2017 04:09	6 466
16	07.07.2017 04:10	07.07.2017 04:10	6 264
17	07.07.2017 04:11	07.07.2017 04:11	6 300
18	07.07.2017 04:12	07.07.2017 04:12	6 552
19	07.07.2017 04:13	07.07.2017 04:13	6 403
20	07.07.2017 04:14	07.07.2017 04:14	6 436
21	07.07.2017 04:15	07.07.2017 04:15	6 265
22	07.07.2017 04:16	07.07.2017 04:16	1 348

Для выбранного файла набора CIC-DDoS2019 были проведены процедуры предобработки, в результате которых был получен датасет с характеристиками, представленными в листинге 1 приложения Д.

Как видно из листинга 1 приложения Д, датасет содержит большое количество атрибутов, в связи с этим был проведен отбор признаков с помощью оценщика признаков ExtraTreesClassifier и мета-трансформатора

SelectFromModel [30]. Оценщик на основе деревьев используется для вычисления значимости атрибутов на основе примесей, которые, в свою очередь, используются для отбрасывания нерелевантных признаков за счет мета-трансформатора.

В результате работы оценщика признаков ExtraTreesClassifier была получена столбчатая диаграмма, отражающая значимость атрибутов набора данных, которая представлена на рисунке 1 приложения В.

В результате отбора признаков были выделены 25 наиболее релевантных атрибутов, их название и описание представлены в таблице 8.

Таблица 8 – 25 наиболее релевантных атрибутов датасета

Атрибут	**Описание**
Timestamp	Время сеанса
Min Packet Length	Минимальная длина пакета
Source Port	Исходный порт, используемый сеансом
Protocol	Протокол, используемый сеансом
Packet Length Mean	Средняя длина пакета
Avg Bwd Segment Size	Средний размер обратного сегмента
Bwd Packet Length Std	Стандартное отклонение размера пакета в обратном направлении
Average Packet Size	Средний размер пакета
Bwd Packet Length Max	Максимальный размер пакета в обратном направлении
Bwd Packet Length Min	Минимальный размер пакета в обратном направлении
Fwd Packet Length Max	Максимальный размер пакета в прямом направлении
Fwd Packet Length Mean	Средний размер пакета в прямом направлении
Max Packet Length	Максимальная длина пакета
Bwd Packet Length Mean	Средний размер пакета в обратном направлении
Packet Length Std	Стандартное отклонение длины пакета
Packet Length Variance	Разброс длин пакетов данных
Flow Id	Внутренний числовой идентификатор потока
Down/Up Ratio	Коэффициент загрузки и выгрузки
Destination Port	Порт назначения, используемый сеансом
Source Ip	IP-адрес источника исходного сеанса
Destination Ip	IP-адрес пункта назначения сеанса
Min_seg_size_forward	Минимальный размер сегмента в прямом направлении
PSH Flag Count	Количество установленных флагов PSH (Push) в TCP пакетах, указывающих на срочную передачу данных
ACK Flag Count	Количество флагов ACK (подтверждение) в TCP заголовке пакета
URG Flag Count	Количество установленных флагов TCP, указывающих на необходимость обработки пакета с высоким приоритетом

Для дальнейшего определения метрик полученный в результате предобработки датасет был разделен на две части:

1) с начала датасета по 19-ую DDoS-атаку включительно – обучение и первичное тестирование модели;

2) с 20-ой по 22-ую DDoS-атаку – определение метрик модели.

В результате выполнения предобработки было получено четыре набора данных, состоящие из 26 атрибутов (25 наиболее релевантных атрибутов датасета и метка) и различного количества строк (таблица 9), которые будут использованы для обучения, тестирования и измерения времени обнаружения разрабатываемой модели.

Таблица 9 – Количество строк полученных наборов данных

Набор данных	**Описание**	**Количество строк**
DDoS	С начала датасета по конец 19-ой DDoS-атаки	173 132
DDoS_20	С конца 19-ой DDoS-атаки по конец 20-ой DDoS-атаки	11 210
DDoS_21	С конца 20-ой DDoS-атаки по конец 21-ой DDoS-атаки	9 104
DDoS_22	С конца 21-ой DDoS-атаки по конец 22-ой DDoS-атаки	4 097

Наборы данных DDoS_20–DDoS_22 состоят из легитимного трафика с последующей DDoS-атакой.

3. РЕАЛИЗАЦИЯ МОДЕЛИ

Поскольку целью данной работы является обнаружение DDoS-атак, то исследуемая задача сводится к задаче классификации, а именно к бинарной классификации, где 0 – это легитимный трафик, а 1 – трафик DDoS-атаки.

Задача классификации – это тип задачи машинного обучения или статистической задачи, целью которой является присвоение категории или метки определенному набору входных данных на основе их характеристик или особенностей [31].

Классификация состоит в том, чтобы изучить сопоставление между входными данными и предопределенными классами (категориями), а затем использовать это сопоставление для прогнозирования меток классов новых, неопределенных точек данных.

Исходя из результатов, полученных в ходе анализа аналогичных решений, проведенного ранее, будем рассматривать метод случайного леса (RF), технику градиентного бустинга (GB), а также архитектуру сетей долгой краткосрочной памяти (LSTM).

3.1. Архитектура модели

Случайный лес

Случайный лес – ансамблевый метод машинного обучения, основанный на построении множества деревьев решений в процессе обучения и принятия решения путем усреднения или выбора большинства прогнозов всех деревьев [32]. Данный метод эффективен для задач классификации и регрессии.

Принцип работы

Алгоритм случайного леса состоит из набора деревьев решений, и каждое дерево в ансамбле состоит из выборки данных, взятой из обучающего набора данных. Затем посредством объединения признаков вводится еще один образец случайности, что добавляет больше разнообразия в набор

данных и уменьшает корреляцию между деревьями решений. Принцип работы случайного леса представлен на рисунке 4.

В зависимости от типа решаемой задачи определение прогноза будет различаться. Для задачи регрессии отдельные деревья решений будут усреднены, а для задачи классификации большинство голосов (то есть наиболее часто встречающаяся категориальная метка) даст прогнозируемый класс.

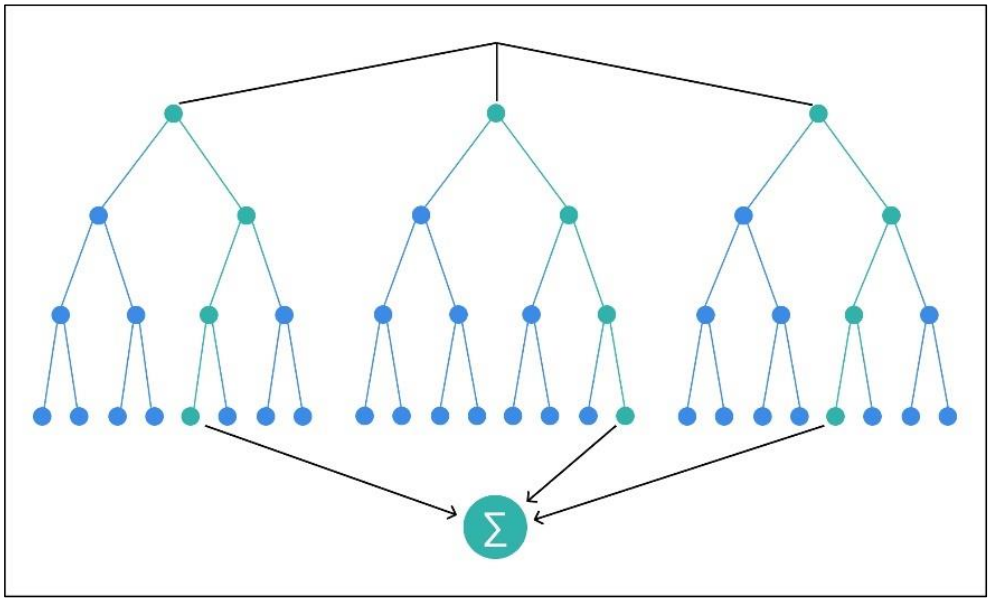

Рисунок 4 – Принцип работы алгоритма случайного леса

Преимущества и недостатки

К преимуществам алгоритма случайного леса относятся:

1) высокая точность предсказания;

2) устойчивость к переобучению;

3) возможность обрабатывать большие наборы данных с большим числом признаков.

К недостаткам можно отнести:

1) результаты могут быть сложными в интерпретации;

2) больший объем времени и ресурсов для обучения в сравнении с отдельными деревьями;

3) склонность к переобучению при случае неправильной настройки гиперпараметров модели.

Сети долгой краткосрочной памяти

Сеть долговременной краткосрочной памяти – это тип рекуррентной нейронной сети (RNN). LSTM преимущественно используются для изучения, обработки и классификации последовательных данных, поскольку эти сети могут изучать долгосрочные зависимости между временными шагами данных [33].

К основным компонентам LSTM относятся следующие элементы:

1) ячейка памяти (Cell State) – поддерживает информацию на протяжении всей последовательности данных и позволяет передавать информацию в следующий шаг времени без изменений;

2) ворота (Gates):

– ворота забывания (Forget Gate) – решает, какую информацию следует забыть из предыдущей ячейки памяти;

– ворота обновления (Update Gate) – определяет, какую информацию следует обновить в ячейке памяти;

– ворота вывода (Output Gate) – определяет, какую информацию из ячейки памяти следует использовать для предсказания;

3) входной вектор (Input Vector) – новая информация, которая должна быть добавлена в ячейку памяти;

4) скрытое состояние (Hidden State) – выходная информация сети, которая также может использоваться для прогнозирования следующего элемента последовательности.

Принцип работы

Принцип работы LSTM представлен на рисунке 5 и может быть описать следующими шагами:

1) получение входных данных и скрытого состояния, вычисленных на предыдущем шаге;

2) вычисление значений ворот на основе входных данных и скрытого состояния модели;

3) обновление ячейки памяти с использованием значений ворот и входного вектора;

4) вычисление нового скрытого состояния на основе обновленной ячейки памяти и входных данных;

5) передача нового скрытого состояния и ячейки памяти на следующий временной шаг.

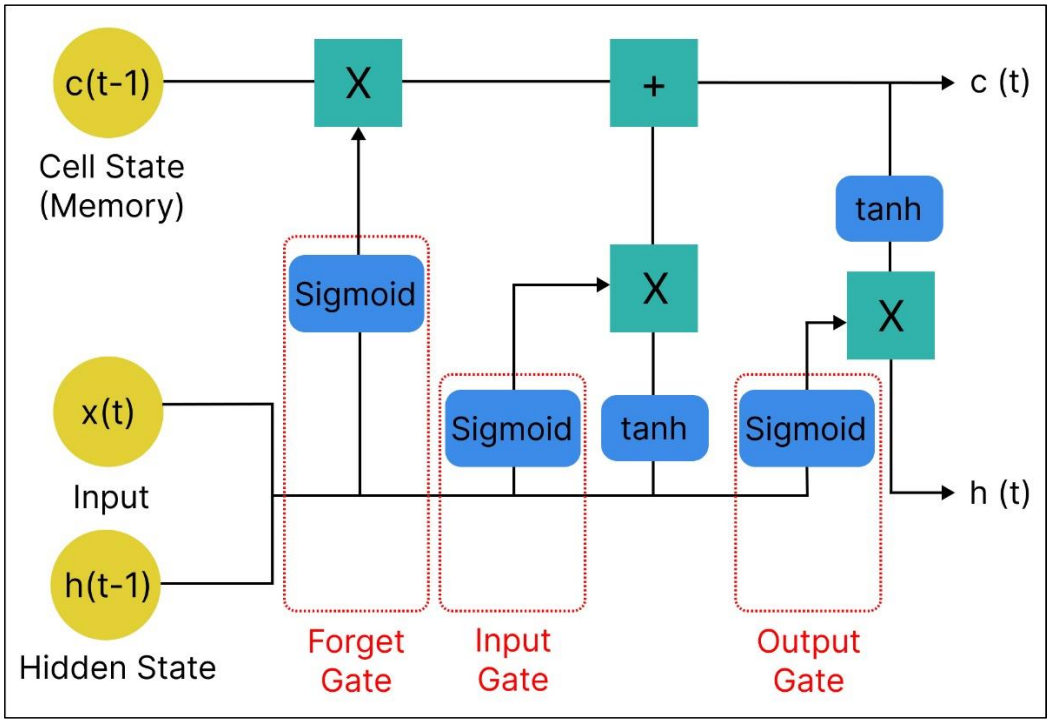

Рисунок 5 – Принцип работы сети долгой краткосрочной памяти

Преимущества и недостатки

К преимуществам LSTM можно отнести:

1) способность изучать долгосрочные зависимости;
2) избирательное сохранение памяти;
3) возможность обработки последовательностей переменной длины;
4) устойчивость к шуму и отсутствующим данным.

К недостаткам относятся:

1) затратность вычислений по ресурсам;
2) сложность интерпретации результатов;
3) склонность к забыванию при длинных последовательностях;

4) сложность обучения из-за большого числа параметров и риска переобучения, особенно если нет достаточного количества данных.

Градиентный бустинг

Градиентный бустинг – это продвинутый алгоритм машинного обучения для решения задач классификации и регрессии. Он строит предсказание в виде ансамбля слабых предсказывающих моделей, которыми в основном являются деревья решений [34].

Градиентный бустинг добавляет базовые модели в ансамбль последовательно, однако вместо обучения моделей с учетом весов на основе ошибок предшественников, в данном случае модели обучаются на остаточных ошибках (residual errors), допущенных предыдущими моделями.

Принцип работы

Принцип работы градиентного бустинга для классификации состоит в том, что к каждому уникальному классу необходимо применить one-hot encoding и перевести в вероятности с помощью softmax-функции, а к прогнозам деревьев добавить коэффициент, который регулирует степень вклада каждого нового дерева в общую модель для снижения переобучения. Также качество прогнозов градиентного бустинга можно существенно улучшить, применив концепцию K-class LogitBoost: для каждого дерева рассчитываются веса, а потом на их основе рассчитываются остатки.

Алгоритм строится следующим образом.

1. Для выборки тренировочных меток классов применяется one-hot encoding и первоначальному прогнозу присваиваются значения для каждого класса.

2. Преобразование прогнозов в вероятности с помощью softmax.

3. Рассчитываются остатки модели на основе антиградиента функции потерь и вероятностей (а также веса в случае LogitBoost).

4. Регрессионное дерево обучается на выборке тренировочных параметров и остатках (а в случае LogitBoost еще присваиваются веса), далее делается прогноз на этой же выборке данных.

5. Для каждого листа в дереве рассчитываются коэффициенты γ на основе остатков, взятых по позициям наблюдений, попавших в определенный в листовой узел.

6. Полученные прогнозы для каждого класса и сумма коэффициентов γ добавляются к первоначальным.

7. Шаги с второго по шестой повторяются для каждого дерева в каждом классе.

8. После обучения всех моделей создается первоначальный прогноз из первого шага алгоритма.

9. Делаются прогнозы для выборки тренировочных параметров на обученных деревьях по каждому классу и добавляются к первоначальным.

10. Конечным прогнозом являются класс с максимальной суммой.

Преимущества и недостатки

К преимуществам градиентного бустинга относятся:

1) работает с любыми функциями потерь;

2) устойчив к переобучению;

3) предсказания в среднем лучше, чем у других алгоритмов;

4) работа как с числовыми, так и с категориальными признаками;

5) самостоятельно справляется с пропущенными данными.

К минусам алгоритма можно отнести:

1) крайне чувствителен к выбросам;

2) при наличии выбросов необходимо большое количество ресурсов;

3) модель будет склонна к переобучению при слишком большом количестве деревьев;

4) более сложная настройка гиперпараметров по сравнению с другими методами;

5) требует больше времени на обучение из-за пошагового построения ансамбля.

3.2. Подбор гиперпараметров

Подбор параметров для моделей МО

С помощью GridSearchCV [37] был произведен подбор параметров для моделей RandomForestClassifier [35] и GradientBoostingClassifier [36].

GridSearchCV – метод поиска наилучших гиперпараметров для модели машинного обучения путем перебора всех возможных комбинаций значений гиперпараметров из заданного набора. Метод позволяет указать сетку возможных значений для различных гиперпараметров модели и перебрать все комбинации этих значений, подсчитывая метрики качества модели для каждой комбинации. После завершения процесса GridSearchCV возвращает наилучшие гиперпараметры, которые обеспечивают наилучшее качество модели.

Для подбора гиперпараметров модели RandomForestClassifier на вход GridSearchCV была предложена следующая сетка параметров:

1) n_estimators – количество деревьев в лесу: 10, 20, 50;
2) min_samples_leaf – минимальное количество выборок, которое должно находиться в листовом узле: 1, 3, 5;
3) min_samples_split – минимальное количество выборок, необходимое для разделения внутреннего узла: 5, 10, 50;
4) max_depth – максимальная глубина отдельных оценок регрессии, количество узлов в дереве: 3, 5, 7.

В результате работы GridSearchCV для RandomForestClassifier были получены следующие гиперпараметры:

1) max_depth – 3;
2) min_samples_leaf – 1;
3) min_samples_split – 5;
4) n_estimators – 20.

Для подбора гиперпараметров модели GradientBoostingClassifier на вход GridSearchCV была предложена следующая сетка параметров:

1) n_estimators – количество оценщиков ансамблевого метода: 100, 150, 200;

2) learning_rate – скорость обучения: 0,01, 0,1, 0,2;

3) min_samples_split – минимальное количество выборок, необходимое для разделения внутреннего узла: 5, 10, 50;

4) max_depth – максимальная глубина отдельных оценок регрессии, количество узлов в дереве: 3, 5, 7.

В результате работы GridSearchCV для GradientBoostingClassifier были получены следующие гиперпараметры:

1) learning_rate – 0,01;

2) max_depth – 3;

3) min_samples_split – 5;

4) n_estimators – 100.

Остальные параметры моделей были оставлены по умолчанию.

Подбор параметров для модели ИИ

Разрабатываемая модель BidirectionalLSTM состоит из четырех слоев:

1) двунаправленный слой LSTM [38] с 128 нейронами;

2) двунаправленный слой LSTM с 64 нейронами;

3) двунаправленный слой LSTM с 32 нейронами;

4) полносвязный слой Dense [39] с 1 нейроном и сигмоидной функцией активации.

Двунаправленность слоев LSTM осуществляется за счет использования двунаправленной оболочки Bidirectional [40]. Для объединения слоев в модели используется последовательная модель Sequential [41].

Подбор параметров для модели осуществлялся на основании параметров входных данных с помощью серии тестов.

Структура модели BidirectionalLSTM представлена в виде листинга 1.

Листинг 1 – Summary модели BidirectionalLSTM

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 bidirectional (Bidirection  (None, 25, 256)           133120
 al)

 bidirectional_1 (Bidirecti  (None, 25, 128)           164352
 onal)

 bidirectional_2 (Bidirecti  (None, 64)                41216
 onal)

 dense (Dense)               (None, 1)                 65

=================================================================
```

Для компиляции модели были подобраны следующие параметры:

1) оптимизатор – Адам (Adam);
2) функция потерь – бинарная кросс-энтропия (Binary Cross-Entropy);
3) метрики для модели во время обучения и тестирования – меткость (Accuracy).

Для обучения модели были подобраны следующие параметры:

1) количество эпох обучения – 3;
2) количество образцов при обновлении градиента (batch_size) – 64.

4. ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

4.1. Метрики оценки качества работы модели

Для оценки качества работы модели бинарной классификации используются различные метрики.

Наиболее популярными метриками являются [42]:

1) матрица ошибок (Confusion matrix);
2) меткость (Accuracy);
3) точность (Precision);
4) полнота (Recall);
5) специфичность (Specificity);
6) F1-мера (F1-score).

Данные метрики основаны на возможных исходах классификации:

1) истинно положительные (True Positive, TP) – истинные значения, которые были верно определены как истинные;
2) истинно отрицательные (True Negative, TN) – ложные значения, которые были верно определены как ложные;
3) ложно положительные (False Positive, FP) – ложные значения, которые были неверно определены как истинные;
4) ложно отрицательные (False Negative, FN) – истинные значения, которые были неверно определены как ложные.

Матрица ошибок (confusion matrix)

Матрица ошибок используется для визуализации производительности алгоритма классификации путем сравнения фактических значений классов с предсказанными моделью.

В случае бинарной классификации матрица ошибок представляет собой матрицу размерностью 2x2, где строки – эталонные метки, а столбцы – предсказанные. Общий вид матрицы ошибок представлен на рисунке 6.

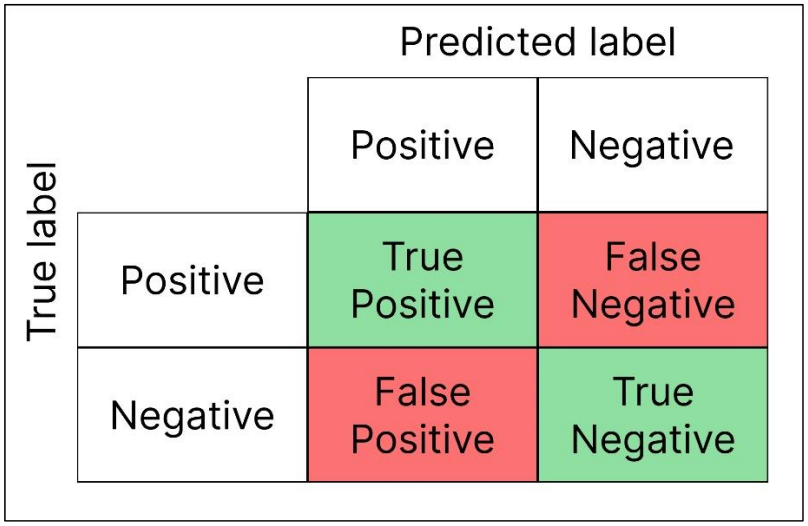

Рисунок 6 – Общий вид матрицы ошибок

Наибольшую ценность для задачи обнаружения DDoS-атак представляют значения показателей True Positive и False Negative.

Меткость (Accuracy)

Accuracy – это показатель, который описывает общую точность предсказания модели по всем классам. Меткость рассчитывается как отношение количества правильных прогнозов к их общему количеству.

Accuracy вычисляется по формуле (1):

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}. \quad (1)$$

Точность (Precision)

Precision – показатель, который описывает производительность модели по отношению к положительным результатам классификации. Данная величина часто упоминается как positive predictive value (PPV) или положительное прогностическое значение. Точность равна доле истинноположительных классификаций к общему числу положительных классификаций.

Precision вычисляется по формуле (2):

$$Pr = PPV = \frac{TP}{TP + FP}. \quad (2)$$

Полнота (Recall)

Recall – это метрика оценки производительности модели, которая отражает способность модели обнаруживать все положительные образцы в данном наборе данных.

Полнота, известная еще как чувствительность или доля истинноположительных примеров, определяется как число истинноположительных классификаций относительно общего числа положительных наблюдений.

Recall вычисляется по формуле (3):

$$Pe = TPR = \frac{TP}{TP + FN}. \quad (3)$$

Специфичность (Specificity)

Specificity – показатель, который описывает производительность модели по отношению к отрицательным результатам классификации. Специфичность определяется как доля истинноотрицательных классификаций в общем числе отрицательных классификаций.

Specificity вычисляется по формуле (4):

$$Sp = TNR = \frac{TN}{TN + FP}. \quad (4)$$

F1-мера (F1-score)

F1-score – это метрика, которая учитывает точность в случае, когда классы не сбалансированы, сосредоточив внимание на точности положительных прогнозов и фактически положительных записей.

Данная метрика учитывает как ложноположительные, так и ложноотрицательные результаты, обеспечивая баланс между значениями метрик точности и полноты модели.

F1-score вычисляется по формуле (5):

$$F1 = \frac{2 \cdot PPV \cdot TPR}{PPV + TPR} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}. \quad (5)$$

Время обнаружения

Время обнаружения DDoS-атаки – это момент, когда система, подвергающаяся атаке, обнаруживает необычный или аномальный уровень сетевого трафика, который может свидетельствовать о DDoS-атаке.

В исследовании временем обнаружения будем считать время в секундах от начала моделируемой DDoS-атаки до момента предсказания моделью метки DDoS-атаки.

Значения перечисленных метрик, полученные при тестировании реализованных моделей, представлены в виде таблицы 3 приложения Г.

Из анализа полученных метрик можно сделать следующие выводы.

1. Наилучшие значения метрики «Время обнаружения» были получены при тестировании модели GradientBoostingClassifier. Лучшее из значений данной метрики, равное 0,0014 секунды, было получено на наборе данных DDoS_20.

2. Лучшие значения качественных метрик были получены при тестировании модели RandomForestClassifier. Значения данных метрик равняются 100% на всех наборах данных.

3. Лучшие значения показателя True Positive из матрицы ошибок было получено при тестировании модели RandomForestClassifier.

4. Лучшее значение показателя False Negative из матрицы ошибок было получено при тестировании модели RandomForestClassifier, оно равняется 0 на всех наборах данных.

4.2. Сравнение времени обнаружения с другими моделями

Для анализа полученных значений метрики времени обнаружения DDoS-атак было произведено сравнение лучшего времени обнаружения, полученного при тестировании разрабатываемых моделей, с временем обнаружения, полученным в результате тестирования аналогичных реализованных моделей для обнаружения DDoS-атак, обученных на датасете CIC-DDoS2019.

Для сравнения были выбраны следующие модели:

1) KNeighborsClassifier [43, 44];

2) C-SupportVectorClassifier (SVC) [43, 45];

3) модель ИИ на основе полносвязного слоя Dense со слоем Dropout, структура которой представлена в виде листинга 2 [46, 47].

Листинг 2 – Summary модели ИИ на основе Dense с Dropout

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_20 (Dense)            (None, 25, 64)            128

 dropout_15 (Dropout)        (None, 25, 64)            0

 dense_21 (Dense)            (None, 25, 64)            4160

 dropout_16 (Dropout)        (None, 25, 64)            0

 dense_22 (Dense)            (None, 25, 64)            4160

 dropout_17 (Dropout)        (None, 25, 64)            0

 dense_23 (Dense)            (None, 25, 2)             130

 dense_24 (Dense)            (None, 25, 1)             3

=================================================================
Total params: 8581 (33.52 KB)
Trainable params: 8581 (33.52 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Поскольку лучшее время обнаружения было получено при измерении на наборе данных DDoS_20, сравнение времени обнаружения моделей будет производится на этом наборе данных (таблица 10).

Таблица 10 – Время обнаружения DDoS-атаки различными моделями

Модель	**Время обнаружения DDoS-атаки, с**
GradientBoostingClassifier	0,0014
KNeighborsClassifier	0,013
SVC	0,0048
Модель ИИ на основе Dense с Dropout	Модель не определила DDoS-атаку

На основании полученной таблицы можно сделать вывод, что лучшее время обнаружения DDoS-атаки на наборе данных DDoS_20 было получено при тестировании разработанной модели GradientBoostingClassifier.

ЗАКЛЮЧЕНИЕ

В современных условиях быстро развивающегося общества обеспечение информационной безопасности становится одним из важнейших приоритетов. С каждым годом увеличивается количество успешно совершенных кибератак [48]. Одной из самых распространенных атак сетевого соединения является распределенная атака типа «отказ в обслуживании».

В рамках данной работы была разработана модель раннего обнаружения DDoS-атак на основе анализа сетевого трафика. При этом были решены следующие задачи.

1. Произведен анализ предметной области и аналогичных решений.
2. Произведен анализ и выбор набора данных.
3. Определена архитектура разрабатываемой модели.
4. Проведено обучение разрабатываемой модели.
5. Проведена оценка результатов работы полученной модели.

Основными результатами данного исследования являются:

1) использование большего количества признаков из набора данных CIC-DDoS2019 для обучения и тестирования разработанной модели по сравнению с другими аналогичными моделями;

2) измерение времени обнаружения DDoS-атаки с использованием разработанной модели;

3) сравнение времени обнаружения DDoS-атаки с помощью разработанной модели и аналогичных моделей, обученных на наборе данных CIC-DDoS2019.

Направление дальнейших исследований включает реализацию проверки наличия DDoS-атаки по временному ряду (например, при трех подряд идущих инцидентов, классифицированных как потенциально опасные), доработку модели в контексте анализа признаков временных рядов сетевого трафика, дальнейшую интеграцию модели в системы администрирования компьютерных сетей.

ЛИТЕРАТУРА

1. Что такое кибератака? [Электронный ресурс] URL: https://www.microsoft.com/ru-ru/security/business/security-101/what-is-a-cyberattack (дата обращения: 10.02.2024 г.).
2. 45 Global DDOS Attack Statistics 2023. [Электронный ресурс] URL: https://www.getastra.com/blog/security-audit/ddos-attack-statistics/ (дата обращения: 10.02.2024 г.).
3. What is a DDoS (distributed denial of service) attack? [Электронный ресурс] URL: https://www.ibm.com/topics/ddos/ (дата обращения: 10.02.2024 г.).
4. Radware Global Application & Network Security 2016-2017 Report. [Электронный ресурс] URL: https://web.tierpoint.com/radware-ert-2016-2017-report-w/ (дата обращения: 10.02.2024 г.).
5. Кадыров Р.Р. Методы обнаружения и предотвращения DDoS-атак. // Политехнический молодежный журнал, 2019. – No 07. – 9 с.
6. Попов И.Ю. Методы и алгоритмы защиты от распределенных сетевых атак типа «отказ в обслуживании»: Диссертация на соискание ученой степени кандидата технических наук. – Санкт-Петербург, 2020. – 240 с.
7. What is the OSI Model? [Электронный ресурс] URL: https://www.cloudflare.com/learning/ddos/glossary/open-systems-interconnection-model-osi/ (дата обращения: 15.02.2024 г.).
8. DDoS Attacks. [Электронный ресурс] URL: https://www.imperva.com/learn/ddos/ddos-attacks/ (дата обращения: 20.02.2024 г.).
9. What Is a Network Packet? [Электронный ресурс] URL: https://www.liveaction.com/resources/blog-post/what-is-a-network-packet/ (дата обращения: 25.02.2024 г.).

10. Защита от DDoS-атак. [Электронный ресурс] URL: https://www.evraas.ru/solutions/ddos-protection/ (дата обращения: 25.02.2024 г.).

11. Perakovic D. Model for Detection and Classification of DDoS Traffic Based on Artificial Neural Network. / D. Perakovic, M. Perisa, I. Cvitic, S. Husnjak // Telfor Journal, 2017. – Vol. 9, No. 1. – 26–31 pp.

12. Saied A. Artificial Neural Networks in the Detection of Known and Unknown DDoS Attacks: Proof-of-Concept. / A. Saied, R.E. Overill, T. Radzik // Commun. Comput. Inf. Sci, 2014. – Vol. 430. – 300–320 pp.

13. Li Y. DDoS attack detection method based on feature extraction of deep belief network. / Y. Li, B. Liu, S. Zhai, M. Chen // IOP Conf. Series: Earth and Environmental Science 252, 2019. – 5 p.

14. Zeinalpour A. Addressing the Effectiveness of DDoS-Attack Detection Methods Based on the Clustering Method Using an Ensemble Method. / A. Zeinalpour, H.A. Ahmed // Electronic, 2022. – No. 11, Article 17. – 16 p.

15. David J. DDoS Attack Detection using Fast Entropy Approach on Flow-Based Network Traffic. / J. David, C. Thomas // Procedia Computer Science 50, 2015. – 30–36 pp.

16. Lopez A. Network Traffic Behavioral Analytics for Detection of DDoS Attacks. / A.D. Lopez, A.P. Mohan, S. Nair // SMU Data Science Review, 2019. – 25 p.

17. Mishra A. Prediction Approach against DDoS Attack based on Machine Learning Multiclassfier [Электронный ресурс] // arXiv.org, 2017. Дата обновления: 27.04.2022 г. URL: https://arxiv.org/abs/2204.12855 (дата обращения: 01.03.2024 г.).

18. Alduailij M. Machine-Learning-Based DDoS Attack Detection Using Mutual Information and Random Forest Feature Importance Method. / M. Alduailij, Q.W. Khan, M. Tahir, M. Sardaraz // Cloud Computing and Symmetry: Latest Advances and Prospects, 2022. – 15 p.

19. Xylogiannopoulos K. Early DDoS Detection Based on Data Mining Techniques. / K. Xylogiannopoulos, P. Karampelas, R. Alhajj // 8th IFIP International Workshop on Information Security Theory and Practice (WISTP), 2014. –190–199 pp.

20. Gaur V. Analysis of Machine Learning Classifers for Early Detection of DDoS Attacks on IoT Devices. / V. Gaur, R. Kumar // Arabian Journal for Science and Engineering, 2021. – 23 p.

21. Kirtas M. Early Detection of DDoS Attacks using Photonic Neural Networks / M. Kirtas, N. Passalis, D. Kalavrouziotis, D. Syrivelis, P. Bakopoulos, N. Pleros, A. Tefas // IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP), 2022. – 5 p.

22. Kaushal K. Early Detection of DDoS Attack in WSN / K. Kaushal, V. Sahni // International Journal of Computer Applications 134, 2016. – 14–18 pp.

23. DDoS Evaluation Dataset (CIC-DDoS2019). [Электронный ресурс] URL: https://www.unb.ca/cic/datasets/ddos-2019.html (дата обращения: 15.03.2024 г.).

24. CIC DoS dataset (2017). [Электронный ресурс] URL: https://www.unb.ca/cic/datasets/dos-dataset.html (дата обращения: 15.03.2024 г.).

25. Jazi H. Detecting HTTP-based application layer DoS attacks on web servers in the presence of sampling / H. Jazi, H. Gonzalez, N. Stakhanova, A. Ghorbani // Computer Networks Volume 121, 2017. – 25–36 pp.

26. Vos M. Characterizing infrastructure of DDoS attacks based on DDoSDB fingerprints // University of Twente, 2019. – 1–6 pp.

27. CSE-CIC-IDS2018 on AWS. [Электронный ресурс] URL: https://www.unb.ca/cic/datasets/ids-2018.html (дата обращения: 20.03.2024 г.).

28. Liu Z. A DDoS Detection Method Based on Feature Engineering and Machine Learning in Software-Defined Network / Z. Liu, Y. Wang, F. Feng, Y. Liu, Z. Li, Y. Shan // Sensors 23, 2023. – 24 p.

29. Leevy J. Detecting cybersecurity attacks across different network features and learners. / J. Leevy, J. Hancock, R. Zuech, T. Khoshgoftaar // Journal of Big Data 8, 2021. – 29 p.

30. Выбор признаков. [Электронный ресурс] URL: https://scikit-learn.ru/1-13-feature-selection/ (дата обращения: 25.03.2024 г.).

31. Binary Classification with TensorFlow Tutorial. [Электронный ресурс] URL: https://www.freecodecamp.org/news/binary-classification-made-simple-with-tensorflow/ (дата обращения: 10.04.2024 г.).

32. What is random forest? [Электронный ресурс] URL: https://www.ibm.com/topics/random-forest (дата обращения: 10.04.2024 г.).

33. LSTM – сети долгой краткосрочной памяти. [Электронный ресурс] URL: https://habr.com/ru/companies/wunderfund/articles/331310/ (дата обращения: 10.04.2024 г.).

34. Градиентный бустинг. Реализация с нуля на Python и разбор особенностей его модификаций (XGBoost, CatBoost, LightGBM). [Электронный ресурс] URL: https://habr.com/ru/articles/799725/ (дата обращения: 10.04.2024 г.).

35. Инструмент sklearn.model_selection.GridSearchCV. [Электронный ресурс] URL: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html (дата обращения: 15.04.2024 г.).

36. Классификатор sklearn.ensemble.RandomForestClassifier. [Электронный ресурс] URL: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html (дата обращения: 15.04.2024 г.).

37. Классификатор sklearn.ensemble.GradientBoostingClassifier. [Электронный ресурс] URL: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html (дата обращения: 15.04.2024 г.).

38. LSTM layer. [Электронный ресурс] URL: https://keras.io/api/layers/recurrent_layers/lstm/ (дата обращения: 15.04.2024 г.).

39. Dense layer. [Электронный ресурс] URL: https://keras.io/api/layers/core_layers/dense/ (дата обращения: 15.04.2024 г.).

40. Слой tf.keras.layers.Bidirectional. [Электронный ресурс] URL: https://www.tensorflow.org/api_docs/python/tf/keras/layers/Bidirectional (дата обращения: 15.04.2024 г.).

41. The Sequential model. [Электронный ресурс] URL: https://keras.io/guides/sequential_model/ (дата обращения: 15.04.2024 г.).

42. Метрики в задачах машинного обучения. [Электронный ресурс] URL: https://habr.com/ru/companies/ods/articles/328372/ (дата обращения: 15.04.2024 г.).

43. Kaggle: expermint-2. [Электронный ресурс] URL: https://www.kaggle.com/code/hananmohamedhafiz/expermint-2#Data-balansing-using-Over-sampling (дата обращения: 20.04.2024 г.).

44. Классификатор sklearn.neighbors.KNeighborsClassifier. [Электронный ресурс] URL: https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html (дата обращения: 20.04.2024 г.).

45. Классификатор sklearn.svm.SVC¶. [Электронный ресурс] URL: https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html (дата обращения: 20.04.2024 г.).

46. GitHub: omossad/cmpt980. [Электронный ресурс] URL: https://github.com/omossad/cmpt980/blob/master/code/ML_category_classifier.py (дата обращения: 20.04.2024 г.).

47. Dense layer. [Электронный ресурс] URL: https://keras.io/api/layers/core_layers/dense/ (дата обращения: 20.04.2024 г.).

48. Актуальные угрозы и ключевые тенденции в сфере кибербезопасности в 2023 году. [Электронный ресурс] URL: https://www.tadviser.ru/index.php (дата обращения: 05.05.2024 г.).

ПРИЛОЖЕНИЯ

Приложение А. Анализ аналогичных решений

Таблица 1 – Анализ исследовательских работ

Исследование	Цель	Входные данные	Метод классификации	Результаты классификации
Perakovic, D. Model for Detection and Classification of DDoS Traffic Based on Artificial Neural Network	Обнаружение и классификация DDoS-атак	Не указано	MLP	95,6 %
Saied, A. Artificial Neural Networks in the Detection of Known and Unknown DDoS Attacks	Обнаружение DDoS-атак	Не указано	ANN	98%
Li, Y. DDoS attack detection method based on feature extraction of deep belief network	Обнаружения DDoS-атак	Не указано	LSTM	Не указано
Zeinalpour, A. Addressing the Effectiveness of DDoS-Attack Detection Methods Based on the Clustering Method Using an Ensemble Method	Обнаружения DDoS-атак	CIC-DoS2017	K-means, SOMs	98,8%
David, J. DDoS Attack Detection using Fast Entropy Approach on Flow-Based Network Traffic	Обнаружения DDoS-атак	CAIDA	Fast Entropy	Не указано
Lopez, A. Network Traffic Behavioral Analytics for Detection of DDoS Attacks	Обнаружения DDoS-атак	CIC-DoS2017	Random Forest	99%
Mishra, A. Prediction Approach against DDoS Attack based on Machine Learning Multiclassfier	Обнаружение, предсказание DDoS-атак	CIC-DDoS2019	Random Forest, SVM Classifier	99,99%
Alduailij, M. Machine-Learning-Based DDoS Attack Detection Using Mutual Information and Random Forest Feature Importance Method	Обнаружения DDoS-атак	CIC-DoS2017, CIC-DDoS2019	Random Forest	99,9977%

Приложение Б. Анализ наборов данных

Таблица 2 – Анализ наборов данных

Название набора данных	Количество атрибутов	Размерность	Типы атак	Лицензия	Формат файла	Работы
CIC-DDoS2019	89	431327	PortMap, NetBIOS, LDAP, MSSQL, UDP, UDP-Lag, SYN, NTP, DNS, SNMP	Общедоступный	CSV, ARFF	[17], [18]
CIC-DoS2017	77	2830743 (DDoS записей – 128027)	TCP SYN, UDP, HTTP	Общедоступный	CSV	[14], [16], [18]
DDoSDB	10	–	UDP, TCP, SSDP, NTP, DNS, Chargen и др.	Общедоступный	–	–
CSE-CIC-IDS2018	80	1048575	LOIC-HTTP, LOIC-UDP, HOIC, SQL Injection и др.	Общедоступный	CSV	[28], [29]

Приложение В. Значимость атрибутов CIC-DDoS2019

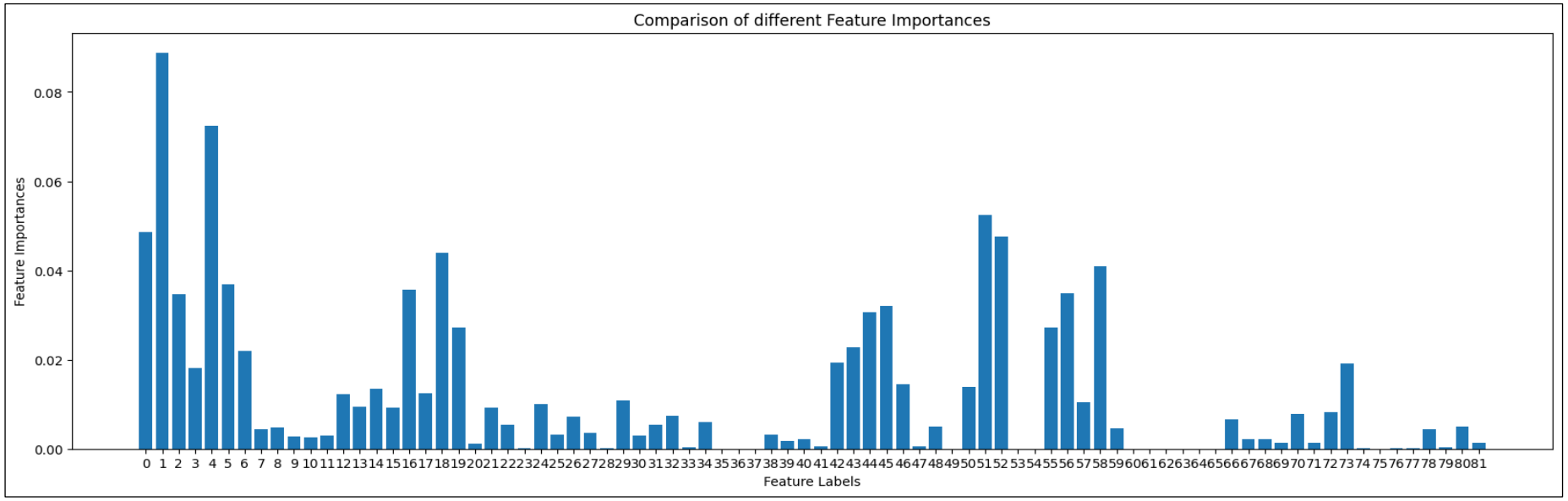

Рисунок 1 – Результат работы оценщика ExtraTreesClassifier

Приложение Г. Метрики разработанных моделей

Таблица 3 – Метрики разработанных моделей

Метрики	**RandomForestClassifier**			**GradientBoostingClassifier**			**BidirectionalLSTM**		
	DDoS_20	**DDoS_21**	**DDoS_22**	**DDoS_20**	**DDoS_21**	**DDoS_22**	**DDoS_20**	**DDoS_21**	**DDoS_22**
Confusion Matrix	[4775 0] [0 6435]	[2839 0] [0 6265]	[2749 0] [0 1348]	[4775 0] [0 6435]	[2835 4] [0 6265]	[2748 1] [0 1348]	[4758 17] [3 6432]	[2780 59] [0 6265]	[2521 228] [0 1348]
Accuracy, %	100	100	100	100	99,96	99,98	99,82	99,35	94,43
Precision, %	100	100	100	100	99,94	99,93	99,74	99,07	85,53
Recall, %	100	100	100	100	100	100	99,95	100	100
Specificity, %	100	100	100	100	99,86	99,96	99,64	97,92	91,71
F1-score, %	100	100	100	100	99,97	99,96	99,84	99,53	92,20
Время обнаружения, с	0,0023	0,0025	0,0022	0,0014	0,0020	0,0017	0,022	0,025	0,021

Приложение Д. Характеристики набора данных

Листинг 1 – Характеристики датасета после предобработки

```
RangeIndex: 225745 entries, 0 to 225744
Data columns (total 83 columns):
 #   Column                       Non-Null Count   Dtype
---  ------                       --------------   -----
 0   Flow ID                      225745 non-null  int64
 1    Source IP                   225745 non-null  int64
 2    Source Port                 225745 non-null  int64
 3    Destination IP              225745 non-null  int64
 4    Destination Port            225745 non-null  int64
 5    Protocol                    225745 non-null  int64
 6    Timestamp                   225745 non-null  int64
 7    Flow Duration               225745 non-null  int64
 8    Total Fwd Packets           225745 non-null  int64
 9    Total Backward Packets      225745 non-null  int64
 10   Total Length of Fwd Packets  225745 non-null  int64
 11   Total Length of Bwd Packets  225745 non-null  int64
 12   Fwd Packet Length Max       225745 non-null  int64
 13   Fwd Packet Length Min       225745 non-null  int64
 14   Fwd Packet Length Mean      225745 non-null  float64
 15   Fwd Packet Length Std       225745 non-null  float64
 16   Bwd Packet Length Max        225745 non-null  int64
 17   Bwd Packet Length Min       225745 non-null  int64
 18   Bwd Packet Length Mean      225745 non-null  float64
 19   Bwd Packet Length Std       225745 non-null  float64
 20   Flow IAT Mean               225745 non-null  float64
 21   Flow IAT Std                225745 non-null  float64
 22   Flow IAT Max                225745 non-null  int64
 23   Flow IAT Min                225745 non-null  int64
 24   Fwd IAT Total                225745 non-null  int64
 25   Fwd IAT Mean                225745 non-null  float64
 26   Fwd IAT Std                 225745 non-null  float64
 27   Fwd IAT Max                 225745 non-null  int64
 28   Fwd IAT Min                 225745 non-null  int64
 29   Bwd IAT Total                225745 non-null  int64
 30   Bwd IAT Mean                225745 non-null  float64
 31   Bwd IAT Std                 225745 non-null  float64
 32   Bwd IAT Max                 225745 non-null  int64
 33   Bwd IAT Min                 225745 non-null  int64
 34   Fwd PSH Flags                225745 non-null  int64
 35   Bwd PSH Flags               225745 non-null  int64
 36   Fwd URG Flags               225745 non-null  int64
 37   Bwd URG Flags               225745 non-null  int64
 38   Fwd Header Length           225745 non-null  int64
 39   Bwd Header Length           225745 non-null  int64
 40   Fwd Packets/s                225745 non-null  float64
 41   Bwd Packets/s               225745 non-null  float64
 42   Min Packet Length           225745 non-null  int64
 43   Max Packet Length           225745 non-null  int64
 44   Packet Length Mean          225745 non-null  float64
 45   Packet Length Std           225745 non-null  float64
 46   Packet Length Variance      225745 non-null  float64
 47   FIN Flag Count               225745 non-null  int64
 48   SYN Flag Count              225745 non-null  int64
 49   RST Flag Count              225745 non-null  int64
 50   PSH Flag Count              225745 non-null  int64
 51   ACK Flag Count              225745 non-null  int64
 52   URG Flag Count              225745 non-null  int64
 53   CWE Flag Count              225745 non-null  int64
 54   ECE Flag Count              225745 non-null  int64
```

```
 55   Down/Up Ratio               225745 non-null  int64
 56   Average Packet Size         225745 non-null  float64
 57   Avg Fwd Segment Size        225745 non-null  float64
 58   Avg Bwd Segment Size        225745 non-null  float64
 59   Fwd Header Length.1         225745 non-null  int64
 60   Fwd Avg Bytes/Bulk           225745 non-null  int64
 61   Fwd Avg Packets/Bulk        225745 non-null  int64
 62   Fwd Avg Bulk Rate           225745 non-null  int64
 63   Bwd Avg Bytes/Bulk          225745 non-null  int64
 64   Bwd Avg Packets/Bulk        225745 non-null  int64
 65   Bwd Avg Bulk Rate            225745 non-null  int64
 66   Subflow Fwd Packets          225745 non-null  int64
 67   Subflow Fwd Bytes           225745 non-null  int64
 68   Subflow Bwd Packets         225745 non-null  int64
 69   Subflow Bwd Bytes           225745 non-null  int64
 70   Init_Win_bytes_forward       225745 non-null  int64
 71   Init_Win_bytes_backward     225745 non-null  int64
 72   act_data_pkt_fwd            225745 non-null  int64
 73   min_seg_size_forward        225745 non-null  int64
 74   Active Mean                  225745 non-null  float64
 75   Active Std                  225745 non-null  float64
 76   Active Max                  225745 non-null  int64
 77   Active Min                  225745 non-null  int64
 78   Idle Mean                    225745 non-null  float64
 79   Idle Std                    225745 non-null  float64
 80   Idle Max                    225745 non-null  int64
 81   Idle Min                    225745 non-null  int64
 82   Label                       225745 non-null  int64
dtypes: float64(22), int64(61)
memory usage: 143.0 MB
```