# Раздел I. Кибератаки и их обнаружение

УДК 004.056 DOI 10.18522/2311-3103-2025-3-6-16

**А.В. Балыбердин**

## МУЛЬТИМОДАЛЬНЫЙ МЕТОД ИЗВЛЕЧЕНИЯ ПРИЗНАКОВ ДАННЫХ ДЛЯ КЛАССИФИКАЦИИ СЕТЕВЫХ АТАК

*Система обнаружения вторжений (СОВ) является важным компонентом защиты корпоративной сети передачи данных (КСПД). СОВ анализирует сетевой трафик и выявляет сетевые атаки. В зависимости от методов детектирования, СОВ можно классифицировать на следующие виды систем: система сигнатурного анализа, система обнаружения аномалий (СОА) и гибридная система, объединяющая ранее рассмотренные системы. В последнее время активно развиваются системы обнаружения аномалий (СОВ). Для систем обнаружения аномалий сетевые атаки представляют собой аномальное поведение сетевого трафика, состоящего из набора признаков или атрибутов событий. Современные СОВ опираются на методы машинного и глубокого обучения, в связи с чем обнаружение сетевых атак и аномалий формулируется как задача классификации и кластеризации. Для решения данных задач необходимы методы оптимизации признакового пространства сетевого трафика. Целью работы является разработка метода извлечения признаков на основе мультимодального подхода представления данных сетевого трафика для классификации сетевых атак. В работе рассмотрен анализ релевантных исследований по методам извлечения признаков из различных областей. Задача исследования – повысить эффективность классификации с помощью метода мультимодального представления признаков сетевого трафика. Результатом работы является метод извлечения признаков данных на основе двух модальностей: спектрального представления признаков сетевого трафика и матрицы признаков изображений. Новизна представленного метода заключается в применении метода оконного преобразования Фурье для событий сетевого трафика, с последующим вычислением спектральных признаков для дискретных сигналов, а также преобразованием признаков данных в матрицу изображений и её расширением для оптимизации пространства признаков с помощью сверточной нейронной сети (convolutional neural network, CNN). Оценка мультимодального метода показала, что данный метод повысил точность классификации для несбалансированных классов сетевых атак.*

*Система обнаружения вторжений; корпоративная сеть передачи данных; набор признаков; извлечение признаков; мультимодальность; сверточная нейронная сеть; задача классификации и кластеризации; признаковое пространство; сетевые атаки.*

**A.V. Balyberdin**

## MULTIMODAL DATA FEATURE EXTRACTION METHOD FOR NETWORK ATTACK CLASSIFICATION

*An intrusion detection system (IDS) is an important component of corporate data network (CDN) protection. IDS analyzes network traffic and detects network attacks. Depending on the detection methods, IDS can be classified into the following types of systems: signature-based analysis systems, anomaly detection systems (ADS), and hybrid systems combining the aforementioned approaches. Recently, anomaly detection systems (IDS) have been actively developing. For anomaly detection systems, network attacks are anomalous behavior of network traffic consisting of a set of features or event attributes. Modern IDS are based on machine and deep learning methods, and therefore the detection of network attacks and anomalies is formulated as a classification and clustering problem. To solve these problems, methods for optimizing the feature space of network traffic are required. The aim of the work is to develop a feature extraction method based on a multimodal approach to representing network traffic data for classifying network attacks. The paper considers the analysis of relevant studies on feature extraction methods from various fields. The objective of the study is to improve classification efficiency using a multimodal repre-*

*sentation of network traffic features. The result of the work is a method for extracting data features based on two modalities: a spectral representation of network traffic features and an image feature matrix. The novelty of the presented method lies in the application of the windowed Fourier transform method for network traffic events, followed by the calculation of spectral features for discrete signals, as well as the transformation of data features into an image matrix and its expansion to optimize the feature space using a convolutional neural network (CNN). Evaluation of the multimodal method showed that this method increased the classification accuracy for unbalanced classes of network attacks.*

*Intrusion detection system; enterprise data network; feature set; feature extraction; multimodality; convolutional neural network; classification and clustering problem; feature space; network attacks.*

**Введение.** В настоящее время особое внимание уделяется вопросам обеспечения информационной безопасности в организации. Для их защиты применяют комплекс организационных и технических мер. Построение корпоративной сети передачи данных неразрывно связано с соблюдением требований информационной безопасности. Одним из таких требований является обязательное применение средств защиты информации для обеспечения мониторинга, контроля сети и выявления сетевых атак [1].

Для обнаружения сетевых атак применяют системы обнаружения вторжений (СОВ) [2]. Классическим подходом детектирования сетевых атак является обнаружение злонамеренного поведения на основе заранее известных шаблонов, паттернов и сигнатур. В связи с постоянным изменением сетевых атак возникают проблемы с поддержкой в актуальном состоянии значительного объема сигнатур и их обработкой СОВ. Увеличение объема базы сигнатур приводит к построению более сложных архитектур СОВ, а также повышаются требования к вычислительным ресурсам системы. Для решения данной проблемы активно развиваются системы обнаружения аномалий (СОА). СОА рассматривает сетевую атаку как аномалию, то есть некое поведение, которое не соответствует нормальному поведению объекта.

Обнаружение аномалий СОВ в академических исследованиях рассматривается как задача классификации и кластеризации на больших объемах данных [3]. Несмотря на значительное количество проведенных исследований, разработанные методы и методики не всегда показывают высокую эффективность обнаружения атак в реальной сети. Для построения эффективной модели используют различные способы и подходы, которые будут рассмотрены в исследовании.

В работе для повышения точности классификации представлен новый мультимодальный метод формирования признакового пространства. Для мультимодального метода проведена оценка и выполнен сравнительный анализ с классическим одномодальным представлением данных.

**1. Постановка задачи.** Одним из способов повышения эффективности модели является представление и оптимизация признаков набора данных [4]. Представление признаков зависит от методов, используемых в классификаторе. Признаки могут соответствовать атрибутам событий, а могут быть вычислены с применением различных методов [5].

Основной мотивацией данного исследования выступает проблема оптимизации и извлечения признаков сетевого трафика, которая влияет на точность классификации сетевых атак и аномалий.

В настоящее время применяют различные методы представления признакового пространства. Для извлечения признаков при классическом подходе используют один метод извлечения признаков [6], в то время как для гибридного подхода могут использовать два и более методов [7]. На основе анализа различных исследований можно сделать вывод, что извлечение признаков с помощью классического и гибридного подходов действительно повышают эффективность классификации.

В связи с развитием интеллектуальных методов вычислений активно развивается новый подход, в котором данные для классификатора представляются в виде различных модальностей с последующим объединением этих модальностей в единый вид.

Основной задачей работы является адаптация мультимодального подхода искусственного интеллекта (ИИ) к представлению признакового пространства сетевого трафика, с целью повысить эффективность классификации сетевых атак и аномалий.

**2. Релевантные работы.** Для обнаружения сетевых атак и аномалий СОВ используют различные виды событий: сетевой трафик и события пользовательской активности. Информативные признаки принято делить на две категории: параметрические и категориальные. К параметрическим признакам сетевого трафика относят признаки с числовыми значениями атрибутов такие как: количество переданных пакетов и байт, статистика сетевых соединений, порты для подключения, флаги, таймауты сессий и другие атрибуты сетевого трафика. Категориальные признаки представляют собой различные категории. К таким признакам можно отнести содержимое прикладных протоколов и журналов пользовательской активности.

Для формирования набора параметрических признаков используют различные технологии. К примеру, в работе [8] для анализа мобильного трафика и выделения набора признаков применяют нейронные сети с использованием многослойного автокодировщика. Для оценки качества сжатия передаваемых данных разработан интегральный показатель, предназначенный для выбора архитектуры многослойного автокодировщика. В работе [9] отмечается, что комбинирование статистических и кластерных методов для представления признакового пространства повышают эффективность обнаружения аномалий. Проведенный глубокий статистический анализ атрибутов трафика позволил на основе корреляционных свойств сформировать четыре атрибутных кластера. В результате удалось сократить признаковое пространство с 51 до 17 признаков.

Рассмотренные выше работы были направлены на сокращение признакового пространства, решая задачу снижения вычислительных ресурсов для анализа событий сетевого трафика. Для повышения эффективности модели можно использовать другой подход, направленный на увеличение количества признаков. К примеру, в работе [10] на основе оценок характеристики мультифрактального спектра фрактальной размерности сетевого трафика были введены новые экспериментальные признаки. В работе показано, что данные признаки повышают эффективность классификации компьютерных атак. В исследовании автор делает предположение об универсальности данного метода.

Как отмечалось выше, что в качестве источника категориальных признаков могут использоваться различные пользовательские журналы. В работе [10] для создания матрицы частоты событий применяют различные виды окон. С помощью окон выполняется разбиение логов на различные группы. Сформированная матрица подается на вход различных моделей машинного обучения. Другим способом извлечения признаков является разработанный метод на основе теории графов [11]. Отметим, что данный метод не требует большие вычислительные ресурсы и показал отличный результат при работе с моделью случайный лес.

В работе [12] для выделения признаков используется теорема Шеннона. С помощью метода производят вычисление энтропий для ранее выбранных признаков. Для обнаружения сетевых атак на вход классификатора направляются рассчитанные 14 энтропий. Применение вычисленных энтропийных свойств атрибутов данных повысил эффективность классификатора.

**3. Метод обнаружения сетевых атак на основе мультимодального представления признаков данных**. ***3.1. Мультимодальное представление признаков данных.*** В связи с развитием технологий искусственного интеллекта (ИИ) широко применяется мультимодальный подход представления данных для решения различных задач. Мультимодальность часто используется в больших языковых моделях (LLM) [13] и предполагает преобразование данных в различные виды модальностей такие как текст, изображение, звук и т.п..

Для СОВ также применяют мультимодальный подход. К примеру, в работе [14] представлена пространственно-временная модель формирования признакового пространства. Архитектура данной модели состоит из двух параллельных методов, построенных на нейронных сетях CNN и LSTM. Выходной слой набора признаков создается путем объединения извлеченных признаков из параллельных ветвей модели. Таким образом, формируется представление данных, состоящее из пространственно-временных признаков. Помимо пространственно-временных признаков могут использоваться другие мо-

дальности как это показано в работе [15]. Автор работы [15] для извлечения признаков использует две модальности: текст и изображение. Для текста извлечение признаков происходит с помощью Spark алгоритмов, а для изображения применяют глубокую сверточную нейронную сеть CNN. С помощью комбинации и объединения признаков разных модальностей удалось повысить точности классификации сетевых атак.

***3.2. Описание методики обнаружения сетевых атак и аномалий с помощью мультимодального представления признаков сетевого трафика.*** Рассмотренные выше модальности применялись для анализа сетевого трафика, но для решения задачи классификации сетевых атак необходимо рассматривать задачу шире. Для этого поток событий сетевого трафика представляем в виде непрерывных или дискретных сигналов. Основываясь на данном предположении, можно рассмотреть другие области, для которых решается аналогичная задача извлечения признаков. К примеру, в медицинской сфере исследуются сигналы, полученные при проведении ЭКГ для выявления патологий. Так в работе [16] представлен широкий обзор методов извлечения признаков в медицинской области. В данной работе технологии извлечения признаков условно объединены в пять категорий или областей: временная область, частотная/спектральная область, частотно-временная область, область декомпозиции и глубокие признаки. В работе отмечается, что извлечение признаков с помощью частотно-временных методов является наиболее эффективным способом выявления патологии при анализе сигналов ЭКГ.

В текущей работе для представления признакового пространства используются две модальности: спектральное представление сетевого трафика и матрица изображений. Методика извлечения признаков с помощью мультимодального подхода представлена на рис. 1.


*Рис. 1. Методика представления признаков данных с помощью разных модальностей*

Сетевой трафик поступает на сенсор, который выполняет сбор и передачу «сырых» событий сетевого трафика для последующей нормализации и агрегации. В процессе нормализации «сырой» сетевой трафик парсится и раскладывается в соответствующие атрибуты проприетарного протокола Flow. При агрегации событий Flow происходит их объединение и удаление дублирующих данных. Далее нормализованный поток событий направляется на спектральный анализ и на преобразование признаков в изображение с последующим их извлечением. В дальнейшем полученные признаки объединяются и направляются в классификатор для обнаружения сетевых атак, таких как DDoS, U2R, R2L и BotNet.

***3.3. Выделение спектральных признаков представления данных.*** Сетевой трафик представляет собой нестационарный процесс [17]. Известно достаточное количество работ, в которых рассматривается фрактальный анализ самоподобия нестационарных свойств сетевого трафика.

Данное исследование направлено на анализ спектральных характеристик сигналов сетевого трафика, вычисленных с помощью метода оконного преобразования Фурье. Для данного метода выберем следующие признаки:

- Суммарное количество переданных байт.
- Суммарное количество полученных байт.
- Суммарное количество переданных пакетов.
- Суммарное количество полученных пакетов.

Сетевой трафик будем рассматривать как дискретные сигналы, для которых можно использовать оконное преобразование Фурье (Short-Time fourier transform, STFT):

$$X(m,k) = \sum_{n=0}^{N-1} x[n] \cdot \omega[n-m] \cdot e^{-j\frac{2\pi}{N}kn},$$

где $x\,[n]$ – дискретный сигнал, $\omega[n-m]$ – оконная функция, $N$ – длина окна, $m$ – временной индекс, $k$ – частотный индекс.

Для STFT существуют несколько видов окон: прямоугольное окно, окна Ханна, Хэмминга и Блэкмана. Для экспериментальной оценки мультимодального метода будем использовать окно Ханна.

На основе оконного преобразования Фурье вычисляем следующие признаки для представления данных:

- Спектральная мощность

$$P = \sum_{k=1}^{N} |X(k)|^2,$$

$X(k)$ – амплитуда частотного компонента $k$

- Спектральная плотность как показатель шумоподобного сигнала

$$F = \frac{\text{Среднегеометрическое } (|X(k)|)}{\text{Среднеарифметическое } (|X(k)|)}$$

- Спектральная энтропия мера хаотичности и неопределенности спектра

$$E = -\sum_{k=1}^{N} p(k) \cdot log_2 \ \big(p(k)\big),$$

где $p(k) = \frac{|X(k)|^2}{\sum_{k=1}^{N}|X(k)|^2}$ – нормированная спектральная мощность

***3.4. Выделение признаков из изображений.*** В настоящее время активно применяют нейронные сети глубокого обучения для задач классификации [18]. Помимо этого, данные методы используют для извлечения признаков сетевого трафика. К примеру, сверточная сеть CNN эффективно извлекает признаки из изображений [19]. Анализируя структуру сетевого трафика, можно сделать вывод, что событие можно представить в виде вектора признаков. В работе [20] атрибуты набора данных KDD представлены в виде вектора признаков изображений. Матрица признаков изображений подается на вход сверточной сети CNN для выделения признаков и оптимизации признакового пространства.

В данном случае, события сетевого трафика можно записать в виде вектора признаков: $X_i = \{x_j, \dots, x_n\}$

Тогда матрица изображений признаков будет: $Y = \begin{matrix} X_i \\ X_n \end{matrix}$

В результате получаем матрицу признаков изображений nxn, которую можно использовать как входные данные для нейронной сети CNN.

Для повышения эффективности классификации расширяем признаковое пространство путем увеличения размерности матрицы изображений с помощью метода бикубической интерполяции. В таком случае новую матрицу можно представить в следующем виде:

$$f(x,y) = \sum_{i=0}^{N}\sum_{j=0}^{N} a_{ij}x^i y^i,$$

где $x, y$ – элементы матрицы, $a_{ij}$ – коэффициент определяется на основе ближайших значений.

***3.5. Оценка эффективности мультимодального метода представления признаков сетевого трафика.*** Для проведения эксперимента будем использовать общедоступный набор данных CSE-CICIDS-2018. Набор данных CSE-CICIDS-2018 является обновлением CICIDS2017. Набор данных также имеет дисбаланс классов и структурно похож на прошлую версию набора. CSE-CICIDS-2018 состоит из большего количества клиентских и атакующих машин [21]. Суммарное количество экземпляров трафика составляет 16 233 002 объектов. Сетевой трафик собирался в течение 10 дней. Процентное распределение типа трафика представлено в табл. 1. Набор данных распределен по 10 csv файлам. Девять файлов состоят из событий сетевого трафика с 79 признаками и один файл с 83 признаками.

Таблица 1

**Распределение сетевого трафика для CSE-CIC-IDS2018**

| № | Тип трафика | Распределение (%) |
|---|---|---|
| 1 | Безопасный | 83,07 |
| 2 | DDoS | 7,786 |
| 3 | DoS | 4,031 |
| 4 | Брутфорс | 2,347 |
| 5 | Ботнет | 1,763 |
| 6 | Проникновение | 0,997 |
| 7 | Веб-атака | 0,006 |

Для оценки эффективности метода экспертно выделим 11 признаков из набора данных CSE-CICIDS-2018: Total Fwd Packet, Total Bwd packets, Total Length of Fwd Packet, Total Length of Bwd Packet, Fwd Header Length, Bwd Header Length, Flow Duration, Flow Bytes/s, Flow Packets/s, Fwd Packets/s, Bwd Packets/s.

Разделим 11 признаков на две модальности следующим образом:

1) Количественные данные пакетов для спектрального анализа: Total Fwd Packet, Total Bwd packets, Total Length of Fwd Packet, Total Length of Bwd Packet, Fwd Header Length, Bwd Header Length, Flow Duration.
2) Скорость потока пакетов для матрицы изображений признаков: Flow Bytes/s, Flow Packets/s, Fwd Packets/s, Bwd Packets/s.

Согласно разработанной методике (см. рис. 1) события сетевого трафика с набором признаков спектрального анализа преобразуются с помощью STFT метода и вычисляются новые признаки: спектральная мощность, спектральная плотность и спектральная энтропия.

Для набора признаков скорости потока пакетов формируется матрица признаков изображений с бикубическим преобразованием размерности матрицы.

В качестве классификатора используем нейронную сеть LSTM с одним общим входом. Структура LSTM представляет собой двухуровневую архитектуру с Dropout-слоем для снижения риска переобучения и адаптации выходного слоя для многоклассовой классификации. Обучающие и тестовые данные делятся соответственно на 70 и 30 от общего набора данных.

Результат работы классификатора LSTM при одномодальном представлении признаков сетевого трафика представлен в табл. 2.

Таблица 2

**Классификатор LSTM с одномодальным представлением признаков сетевого трафика. 11 признаков CSE-CICIDS-2018**

| № | Class | Precision | Recall | F1-score | Support |
|---|---|---|---|---|---|
| 1 | BENIGN | 0.78 | 0.97 | 0.86 | 1183 |
| 2 | Botnet | 0.58 | 0.47 | 0.52 | 104 |
| 3 | Botnet - Attempted | 0.00 | 0.00 | 0.00 | 253 |
| 4 | DDoS | 0.99 | 1.00 | 0.99 | 71 |
| 5 | Portscan | 0.83 | 0.40 | 0.54 | 48 |
| | | | | | |
| 6 | accuracy | | | 0.78 | 1659 |
| 7 | macro avg | 0.63 | 0.57 | 0.58 | 1659 |
| 8 | weighted avg | 0.66 | 0.78 | 0.70 | 1659 |

На рис. 2 представлен график изменений Accuracy и функции потерь в зависимости от количества эпох.

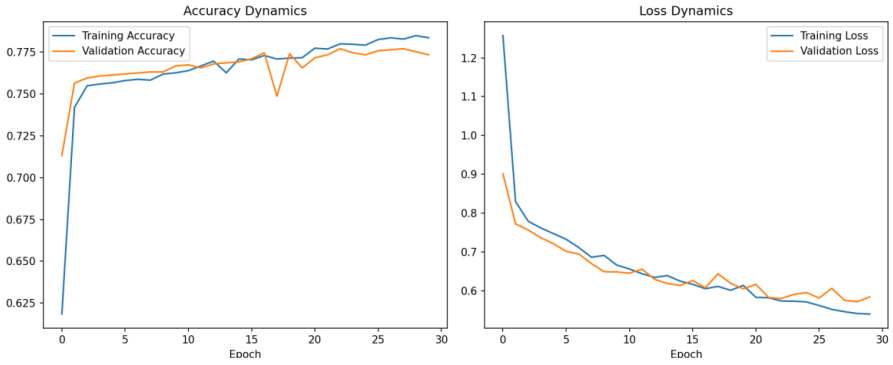


*Рис. 2. Accuracy и функция потерь с одномодальным представлением сетевого трафика в зависимости от эпох. 11 признаков CSE-CICIDS-2018*

В табл. 2 показаны результаты классификации с мультимодальным представлением признаков сетевого трафика.

Таблица 2

**Классификатор LSTM с мультимодальным представлением признаков сетевого трафика. 11 признаков CSE-CICIDS-2018**

| № | Class | Precision | Recall | F1-score |
|---|---|---|---|---|
| 1 | BENIGN | 0.78 | 0.96 | 0.86 |
| 2 | Botnet | 0.69 | 0.39 | 0.50 |
| 3 | Botnet - Attempted | 0.57 | 0.10 | 0.17 |
| 4 | DDoS | 0.99 | 0.97 | 0.98 |
| 5 | Portscan | 0.76 | 0.40 | 0.52 |
| | | | | |
| 6 | Accuracy | | | 0.78 |
| 7 | Macro Avg | 0.76 | 0.57 | 0.61 |
| 8 | Weighted Avg | 0.75 | 0.78 | 0.73 |

На графике Accuracy и функции потерь можно заметить снижение времени обучения модели для набора данных CSE-CICIDS-2018 (рис. 3). Примерно с 15 эпохи начинается этап переобучения.

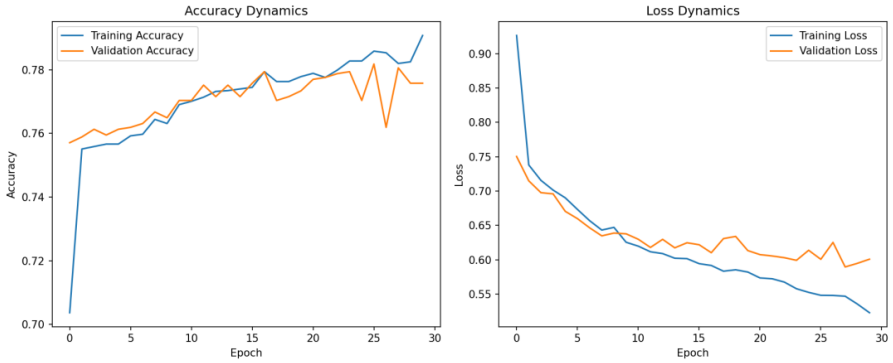


*Рис. 3. Accuracy и функция потерь мультимодального представления сетевого трафика с одним общим входом в зависимости от эпох. 11 признаков CSE-CICIDS-2018*

В табл. 3 представлен сравнительный анализ методов выделения признаков различной модальности. Мультимодальный метод извлечения признаков повысил оценку precision на 18,9% для классификации сетевой атаки Botnet. Также отметим значительное повышение точности обнаружения сетевой атаки Botnet – Attempted. Оценка precision показывает увеличение точности классификации при мультимодальном методе на 57%, для Recall – 10% и для F1-score – 17%.

Таблица 3

**Сравнительный анализ одномодального и мультимодального метода**

| № | Class | Precision | Recall | F1-score | Precision | Recall | F1-score |
|---|---|---|---|---|---|---|---|
| | | Одномодальный метод | | | Мультимодальный | | |
| 1 | BENIGN | 0.78 | 0.97 | 0.86 | 0.78 | 0.96 | 0.86 |
| 2 | Botnet | 0.58 | 0.47 | 0.52 | **0.69** | 0.39 | 0.50 |
| 3 | Botnet - Attempted | 0.00 | 0.00 | 0.00 | **0.57** | **0.10** | **0.17** |
| 4 | DDoS | 0.99 | 1.00 | 0.99 | 0.99 | 0.97 | 0.98 |
| 5 | Portscan | 0.83 | 0.40 | 0.54 | 0.76 | 0.40 | 0.52 |
| | | | | | | | |
| 6 | accuracy | | | 0.78 | | | 0.78 |
| 7 | macro avg | 0.63 | 0.57 | 0.58 | **0.76** | 0.57 | **0.61** |
| 8 | weighted avg | 0.66 | 0.78 | 0.70 | **0.75** | 0.78 | **0.73** |

Применение мультимодального метода представления данных со спектральными признаками и матрицей признаков изображений повысил точность классификации для несбалансированных классов сетевых атак: Botnet и Botnet-Attempted. Оценки по другим классам атак практически не изменились. На основе проведенных экспериментов можно сделать вывод о том, что мультимодальный метод можно использовать в качестве одного из способов повышения эффективности классификации сетевых атак и аномалий.

**Заключение.** В работе проведен анализ основных подходов и методов представления признакового пространства. Отмечается, что мультимодальный подход является новым способом повышения эффективности классификации. Данный подход используется во многих областях, где применяют большие языковые модели (LLM). В работе отмечается важность оптимизации признакового пространства и её влияния на классификацию.

В работе представлен новый метод извлечение признаков сетевого трафика, который отличается от других методов следующим:

- Признаковое пространство формируется на основе мультимодального представления данных сетевого трафика.
- Сформированы новые признаки на основе спектрального анализа оконного преобразования Фурье с помощью признаков транспортного уровня сетевого трафика.
- Матрица признаков изображений второй модальности формируется из признаков сетевого трафика с последующим повышением размерности с помощью метода бикубической интерполяции.

В работе выполнена проверка нового мультимодального метода представления признаков данных. Оценка метода показала, что новый метод извлечения признаков повысил точность классификации для несбалансированных классов сетевых атак на наборе CSE-CICIDS-2018.

Дальнейшие исследования будут направлены на снижение дисбаланса данных в классах с помощью методов аугментации данных с применением генеративных нейронных сетей.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК


1. *Васильев В.И., Кириллова А.Д., Кухарев С.Н.* Кибербезопасность автоматизированных систем управления промышленных объектов (современное состояние, тенденции) // Вестник УрФО. Безопасность в информационной сфере. – 2018. – № 4 (30). – С. 66-74. – DOI: 10.14529/secur180410. – EDN YUNKEP.
2. *Шелухин О.И., Сакалема Д.Ж., Филинова А.С.* Обнаружение вторжений в компьютерные сети (сетевые аномалии): учеб. пособие / под ред. О.И. Шелухина. – М.: Горячая линия-Телеком, 2018. – 220 с. – ISBN 978-5-9912-0323-4. Лань: электронно-библиотечная система. – URL: https://e.lanbook.com/book/111119.
3. *Шелухин О.И.* Сетевые аномалии. Обнаружение, локализация, прогнозирование. – М.: Горячая линия-Телеком, 2019. – 447 с. – ISBN 978-5-9912-0756-0.
4. *Varun Chandola, Arindam Banerjee, and Vipin Kumar.* Anomaly detection: A survey. ACM Comput. // Surv. – 2009. – 41, 3, Article 15 (July 2009). – 58 p. – https://doi.org/10.1145/1541880.1541882.
5. *Шелухин О.И., Судариков Р.А.* Анализ информативных признаков в задачах обнаружения аномалий трафика статистическими методами // T-Comm: Телекоммуникации и транспорт. – 2014. – Т. 8, № 3. – С. 14-18. – EDN SGIHFZ.
6. *Xin R., Liu H., Chen P. et al.* Robust and accurate performance anomaly detection and prediction for cloud applications: a novel ensemble learning-based framework // J Cloud Comp. – 2023. – 12, 7. – https://doi.org/10.1186/s13677-022-00383-6.
7. *Alsaffar A.M., Nouri-Baygi M. & Zolbanin H.M.* Shielding networks: enhancing intrusion detection with hybrid feature selection and stack ensemble learning // J Big Data. – 2024. – 11, 133. – https://ezpro.fa.ru:2117/10.1186/s40537-024-00994-7.
8. *Шелухин О.И., Маторин Ф.А*. Снижение размерности массивов данных с помощью многослойных автокодировщиков в задаче классификации мобильных приложений // Тр. учебных заведений связи. – 2024. – Т. 10, № 6. – С. 111-120. – DOI: 10.31854/1813-324X-2024-10-6-111-120. – EDN TOPDUA.
9. *Шелухин О.И., Раковский Д.И.* Выбор метрических атрибутов редких аномальных событий компьютерной системы методами интеллектуального анализа данных // T-Comm: Телекоммуникации и транспорт. – 2021. – Т. 15, № 6. – С. 40-47. – DOI: 10.36724/2072-8735-2021-15-6-40-47. – EDN YJDUYV.
10. *Шелухин О.И., Рябинин В.С., Фармаковский М.А.* Обнаружение аномальных состояний компьютерных систем средствами интеллектуального анализа данных системных журналов // Вопросы кибербезопасности. – 2018. – № 2 (26). – С. 33-43. – DOI: 10.21681/2311-3456-2018-2-33-43. – EDN XYHQUP.
11. *Слипенчук П.В.* Алгоритм извлечения характерных признаков из данных пользовательских активностей // Вопросы кибербезопасности. – 2019. – № 1 (29). – С. 53-58. – DOI: 10.21681/2311-3456-2019-1-53-58. – EDN YZFWPZ.
12. *Do E.H. and Gadepally V.N*. Classifying Anomalies for Network Security // ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020. – P. 2907-2911. – DOI: 10.1109/ICASSP40776.2020.9053419.

13. *Wu J., Gan W., Chen Z., Wan S. and Yu P.S.* Multimodal Large Language Models: A Survey // 2023 IEEE International Conference on Big Data (BigData), Sorrento, Italy, 2023. – P. 2247-2256. – DOI: 10.1109/BigData59044.2023.10386743.
14. *Shi S., Han D., & Cui M.* A multimodal hybrid parallel network intrusion detection model // Connection Science. – 2023. – 35 (1). – https://doi.org/10.1080/09540091.2023.2227780.
15. *Ullah F., Turab A., Ullah S., Cacciagrano D., Zhao Y.* Enhanced Network Intrusion Detection System for Internet of Things Security Using Multimodal Big Data Representation with Transfer Learning and Game Theory // Sensors. – 2024. – 24 (13):4152. – https://doi.org/10.3390/s24134152.
16. *Singh A.K., Krishnan S.* ECG signal feature extraction trends in methods and applications // BioMed Eng OnLine. – 2023. – 22. – https://doi.org/10.1186/s12938-023-01075-1.
17. *Kotenko I., Saenko I., Lauta O., Kribel A.* An approach to detecting cyber attacks against smart power grids based on the analysis of network traffic self-similarity // Energies. – 2020. – Vol. 13, No. 19. – P. 5031. – DOI: 10.3390/en13195031. – EDN YVERBA.
18. *Гетьман А.И., Горюнов М.Н., Мацкевич А.Г. [и др.].* Применение глубокого обучения для обнаружения компьютерных атак в сетевом трафике // Тр. Института системного программирования РАН. – 2023. – Т. 35, № 4. – С. 65-92. – DOI: 10.15514/ISPRAS-2023-35(4)-3. – EDN CSLHAE.
19. *Jogin M., Mohana, Madhulika M.S., Divya G.D., Meghana R.K. and Apoorva S.* Feature Extraction using Convolution Neural Networks (CNN) and Deep Learning // 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 2018. – P. 2319-2323. – DOI: 10.1109/RTEICT42901.2018.9012507.
20. *Xiao Y., Xing C., Zhang T. and Zhao Z.* An Intrusion Detection Model Based on Feature Reduction and Convolutional Neural Networks // in IEEE Access. – 2019. – Vol. 7. – P. 42210-42219. – DOI: 10.1109/ACCESS.2019.2904620.
21. *Thakkar A., Lohiya R.* A review of the advancement in intrusion detection datasets // Procedia Comput Sci. – 2020. – 167. – P. 636-645.

REFERENCES

1. *Vasil'ev V.I., Kirillova A.D., Kukharev S.N.* Kiberbezopasnost' avtomatizirovannykh sistem upravleniya promyshlennykh ob"ektov (sovremennoe sostoyanie, tendentsii) [Cybersecurity of automated control systems of industrial facilities (current status, trends)], *Vestnik UrFO. Bezopasnost' v informatsionnoy sfere* [Bulletin of the Ural Federal District. Security in the Information Sphere**],** 2018, No. 4 (30), pp. 66-74. DOI: 10.14529/secur180410. EDN YUNKEP.
2. *Shelukhin O.I., Sakalema D.Zh., Filinova A.S.* Obnaruzhenie vtorzheniy v komp'yuternye seti (setevye anomalii): ucheb. posobie [Detection of intrusions in computer networks (network anomalies): a tutorial], ed. by O.I. Shelukhina. Moscow: Goryachaya liniya-Telekom, 2018, 220 p. ISBN 978-5-9912-0323-4. Lan': elektronno-bibliotechnaya sistema. Available at: https://e.lanbook.com/book/111119.
3. *Shelukhin O.I.* Setevye anomalii. Obnaruzhenie, lokalizatsiya, prognozirovanie [Network anomalies. Detection, localization, forecasting]. Moscow: Goryachaya liniya-Telekom, 2019, 447 p. ISBN 978-5-9912-0756-0.
4. *Varun Chandola, Arindam Banerjee, and Vipin Kumar.* Anomaly detection: A survey. ACM Comput., Surv., 2009, 41, 3, Article 15 (July 2009), 58 p. Available at: https://doi.org/10.1145/1541880.1541882.
5. *Shelukhin O.I., Sudarikov R.A.* Analiz informativnykh priznakov v zadachakh obnaruzheniya anomaliy trafika statisticheskimi metodami [Analysis of informative features in the problems of detecting traffic anomalies by statistical methods], *T-Comm: Telekommunikatsii i transport* [T-Comm: Telecommunications and transport], 2014, Vol. 8, No. 3, pp. 14-18. EDN SGIHFZ.
6. *Xin R., Liu H., Chen P. et al.* Robust and accurate performance anomaly detection and prediction for cloud applications: a novel ensemble learning-based framework, *J Cloud Comp*., 2023, 12, 7. Available at: https://doi.org/10.1186/s13677-022-00383-6.
7. *Alsaffar A.M., Nouri-Baygi M. & Zolbanin H.M.* Shielding networks: enhancing intrusion detection with hybrid feature selection and stack ensemble learning, *J Big Data,* 2024, 11, 133. Available at: https://ezpro.fa.ru:2117/10.1186/s40537-024-00994-7.
8. *Shelukhin O.I., Matorin F.A.* Snizhenie razmernosti massivov dannykh s pomoshch'yu mnogosloynykh avtokodirovshchikov v zadache klassifikatsii mobil'nykh prilozheniy [Reducing the dimensionality of data arrays using multilayer autoencoders in the problem of mobile application classification], *Tr. uchebnykh zavedeniy svyazi* [Proceedings of educational institutions of communication], 2024, Vol. 10, No. 6, pp. 111-120. DOI: 10.31854/1813-324X-2024-10-6-111-120. EDN TOPDUA.
9. *Shelukhin O.I., Rakovskiy D.I.* Vybor metricheskikh atributov redkikh anomal'nykh sobytiy komp'yuternoy sistemy metodami intellektual'nogo analiza dannykh [Selection of metric attributes of rare anomalous events of a computer system using data mining methods], *T-Comm: Telekommunikatsii i transport* [T-Comm: Telecommunications and transport], 2021, Vol. 15, No. 6, pp. 40-47. DOI: 10.36724/2072-8735-2021-15-6-40-47. EDN YJDUYV.

10. *Shelukhin O.I., Ryabinin V.S., Farmakovskiy M.A.* Obnaruzhenie anomal'nykh sostoyaniy komp'yuternykh sistem sredstvami intellektual'nogo analiza dannykh sistemnykh zhurnalov [Detection of abnormal states of computer systems by means of intelligent analysis of system log data], *Voprosy kiberbezopasnosti* [Cybersecurity Issues], 2018, No. 2 (26), pp. 33-43. DOI: 10.21681/2311-3456-2018-2-33-43. EDN XYHQUP.
11. *Slipenchuk P.V.* Algoritm izvlecheniya kharakternykh priznakov iz dannykh pol'zovatel'skikh aktivnostey [Algorithm for extracting characteristic features from user activity data], Voprosy kiberbezopasnosti [Cybersecurity Issues], 2019, No. 1 (29), pp. 53-58. DOI: 10.21681/2311-3456-2019-1-53-58. EDN YZFWPZ.
12. *Do E.H. and Gadepally V.N*. Classifying Anomalies for Network Security, *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020,* pp. 2907-2911. DOI: 10.1109/ICASSP40776.2020.9053419.
13. *Wu J., Gan W., Chen Z., Wan S. and Yu P.S*. Multimodal Large Language Models: A Survey, *2023 IEEE International Conference on Big Data (BigData), Sorrento, Italy, 2023,* pp. 2247-2256. DOI: 10.1109/BigData59044.2023.10386743.
14. *Shi S., Han D., & Cui M.* A multimodal hybrid parallel network intrusion detection model, *Connection Science,* 2023, 35 (1). Available at: https://doi.org/10.1080/09540091.2023.2227780.
15. *Ullah F., Turab A., Ullah S., Cacciagrano D., Zhao Y.* Enhanced Network Intrusion Detection System for Internet of Things Security Using Multimodal Big Data Representation with Transfer Learning and Game Theory, *Sensors,* 2024, 24 (13):4152. Available at: https://doi.org/10.3390/s24134152.
16. *Singh A.K., Krishnan S.* ECG signal feature extraction trends in methods and applications, *BioMed Eng OnLine,* 2023, 22. Available at: https://doi.org/10.1186/s12938-023-01075-1.
17. *Kotenko I., Saenko I., Lauta O., Kribel A.* An approach to detecting cyber attacks against smart power grids based on the analysis of network traffic self-similarity, *Energies,* 2020, Vol. 13, No. 19, pp. 5031. DOI: 10.3390/en13195031. EDN YVERBA.
18. *Get'man A.I., Goryunov M.N., Matskevich A.G. [i dr.].* Primenenie glubokogo obucheniya dlya obnaruzheniya komp'yuternykh atak v setevom trafike [Application of deep learning to detect computer attacks in network traffic], *Tr. Instituta sistemnogo programmirovaniya RAN* [Proceedings of the Institute for System Programming of the Russian Academy of Sciences], 2023, Vol. 35, No. 4, pp. 65-92. DOI: 10.15514/ISPRAS-2023-35(4)-3. EDN CSLHAE.
19. *Jogin M., Mohana, Madhulika M.S., Divya G.D., Meghana R.K. and Apoorva S.* Feature Extraction using Convolution Neural Networks (CNN) and Deep Learning, *2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 2018,* pp. 2319-2323. DOI: 10.1109/RTEICT42901.2018.9012507.
20. *Xiao Y., Xing C., Zhang T. and Zhao Z.* An Intrusion Detection Model Based on Feature Reduction and Convolutional Neural Networks, *in IEEE Access,* 2019, Vol. 7, pp. 42210-42219. DOI: 10.1109/ACCESS.2019.2904620.
21. *Thakkar A., Lohiya R.* A review of the advancement in intrusion detection datasets, *Procedia Comput Sci*., 2020, 167, pp. 636-645.

**Балыбердин Алексей Викторович –** Финансовый университет при Правительстве РФ**;** e-mail:balyberdinav@gmail.com**;** г. Москва, Россия; аспирант.

**Balyberdin Alexey Viktorovich** – Financial University under the Government of the Russian Federation; e-mail:balyberdinav@gmail.com; Moscow, Russia; graduate student.

УДК 004.89 DOI 10.18522/2311-3103-2025-3-16-31

**М.А. Лапина, Р.А. Дымуха, Н.Н. Кучеров, Е.С. Басан**

## ИССЛЕДОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ СПУФИНГ-АТАК В ДЕЦЕНТРАЛИЗОВАННЫХ СЕТЯХ

*Беспилотные летательные аппараты всё больше и больше появляются в нашей жизни и используются для различных целей, таких как доставка грузов, мониторинг, управление хозяйством, мониторинг и развлечения. Но вместе с ростом их популярности, увеличивается и число людей, которые намеренно хотят помешать работе БВС (беспилотным воздушным судам) и использовать в своих интересах и целях. Они используют различные виды атак, чтобы любыми способами*