DOI: 10.15514/ISPRAS-2025-37(6)-37

Перенос обучения в сетевых системах обнаружения вторжений: обзор методов и подходов

[1] *А.Ю. Покидько, ORCID: 0009-0008-8981-8429 <a.pokidko@ispras.ru>*
[1,2] *И.А. Степанов, ORCID: 0009-0003-1964-5001 <ivan_mipt@ispras.ru>*
[1,2,3,4] *А.И. Гетьман, ORCID: 0000-0002-6562-9008 <ever@ispras.ru>*

[1] *Институт системного программирования им. В.П. Иванникова РАН, Россия, 109004, г. Москва, ул. А. Солженицына, д. 25.*

[2] *Московский физико-технический институт, 141700, Россия, Московская область, г. Долгопрудный, Институтский пер., 9.*

[3] *Национальный исследовательский университет «Высшая школа экономики», 101978, Россия, г. Москва, ул. Мясницкая, д. 20.*

[4] *Московский государственный университет имени М.В. Ломоносова, Россия, 119991, Москва, Ленинские горы, д. 1.*

Аннотация. Статья представляет обзор современных методов переноса обучения (transfer learning) в сетевых системах обнаружения вторжений (СОВ), ориентируясь на проблему устойчивости моделей в условиях дрейфа сетевых данных, изменчивости трафика и появления новых типов атак. Рассматриваются основные парадигмы переноса – параметрический, признаковый и основанный на отношениях – и их адаптация к задаче обнаружения аномалий и классификации сетевых вторжений. Особое внимание уделено различиям между методами на основе анализа статистических свойств сетевых потоков и методами на основе анализа пакетов. На основе анализа существующих работ демонстрируется, что использование переноса обучения позволяет существенно повысить устойчивость сетевых СОВ к изменениям инфраструктуры и распределений данных, однако сталкивается с проблемами негативного переноса, недостатка репрезентативных источников домена и усложнения архитектур. В завершение формулируются ключевые направления дальнейших исследований, включая адаптивные модели с учётом дрейфа, перенос в условиях ограниченных данных и интеграцию с потоковыми методами машинного обучения.

Ключевые слова: сетевая система обнаружения вторжений (СОВ); перенос обучения.

Для цитирования: Покидько А.Ю., Степанов И.А., Гетьман А.И. Перенос обучения в сетевых системах обнаружения вторжений: обзор методов и подходов. Труды ИСП РАН, том 37, вып. 6, часть 3, 2025 г., стр. 73–90. DOI: 10.15514/ISPRAS–2025–37(6)–37.

Благодарности. Результаты получены с использованием услуг Центра коллективного пользования Института системного программирования им. В.П. Иванникова РАН – ЦКП ИСП РАН.

Transfer Learning in Network Intrusion Detection Systems: a Review of Methods and Approaches

[1] *A.Y. Pokidko, ORCID: 0009-0008-8981-8429 <a.pokidko@ispras.ru>*
[1,2] *I.A. Stepanov, ORCID: 0009-0003-1964-5001 <ivan_mipt@ispras.ru>*
[1,2,3,4] *A.I. Getman, ORCID: 0000-0002-6562-9008 <ever@ispras.ru>*

[1] *Ivannikov Institute for System Programming of the Russian Academy of Sciences, 25, Alexander Solzhenitsyn st., Moscow, 109004, Russia.*

[2] *Moscow Institute of Physics and Technology (National Research University), 9 Institutskiy per., Dolgoprudny, Moscow Region, 141701, Russia.*

[3] *National Research University «Higher School of Economics», 20, Myasnitskaya ulitsa, Moscow, 101000, Russia.*

[4] *Lomonosov Moscow State University, GSP-1, Leninskie Gory, Moscow, 119991, Russia.*

Abstract. This article provides an overview of modern transfer learning methods in network intrusion detection systems (IDS), focusing on the problem of model stability in conditions of network data drift, traffic variability, and the emergence of new types of attacks. The main transfer paradigms – parametric, feature-based, and relationship-based – and their adaptation to the task of anomaly detection and network intrusion classification are considered. Particular attention is paid to the differences between methods based on the analysis of statistical properties of network flows and methods based on packet analysis. Based on an analysis of existing work, it is demonstrated that the use of transfer learning can significantly improve the robustness of network IDSs to changes in infrastructure and data distributions, but faces problems of negative transfer, lack of representative domain sources, and architectural complexity. Finally, key directions for further research are formulated, including adaptive models that account for drift, transfer under limited data conditions, and integration with streaming machine learning methods.

Keywords: network intrusion detection system (NIDS); transfer learning.

For citation: Pokidko A.Y., Stepanov I.A., Getman A.I. Transfer learning in network intrusion detection systems: a review of methods and approaches. Trudy ISP RAN/Proc. ISP RAS, vol. 37, issue 6, part 3, 2025, pp. 73-90 (in Russian). DOI: 10.15514/ISPRAS-2025-37(6)-37.

Acknowledgements. The results were obtained using the services of the Ivannikov Institute for System Programming (ISP RAS) Data Center.

1. Введение

С ростом объемов и сложности сетевого трафика в современных информационных системах возрастает необходимость в эффективных методах его анализа. Такие задачи, как обнаружение аномалий, классификация трафика, выявление вредоносной активности сетевых атак, требуют высокой точности и адаптивности используемых сетевых систем обнаружения вторжений (СОВ, IDS).

Существующие сетевые СОВ делятся на 3 типа в зависимости от метода обнаружения атак [1]:

1. Обнаружение на основе сигнатур: метод, отслеживающий пакеты в сети и сравнивающий их с предварительно настроенными и заранее определенными шаблонами атак, известными как сигнатуры. Такие СОВ отличаются высокой точностью обнаружения известных атак, но требуют постоянного ручного обновления базы сигнатур.

2. Анализ протокола с сохранением состояния (Stateful protocol analysis, SPA): IDS может знать и отслеживать состояния протокола (например, связывая запросы с ответами). Как правило, модели сетевых протоколов в SPA изначально основаны на

стандартах протоколов международных организаций по стандартизации, например, IETF. Такие СОВ способны обнаруживать новые и неизвестные атаки (если они нарушают спецификацию), но имеют высокую ресурсоёмкость и могут не распознать атаку, которая формально корректна по спецификации, но содержит вредоносную полезную нагрузку (например, SQL-инъекция внутри корректного HTTP-запроса).

3. Статистическое обнаружение аномалий: сетевая СОВ, основанная на аномалиях, будет отслеживать сетевой трафик и сравнивать его характеристики с установленным базовым уровнем. Такие системы способны обнаруживать неизвестные или новые типы атак, но часто дают много ложных срабатываний.

В этой статье будут рассмотрены статистические сетевые СОВ на основе нейронных сетей, так как они позволяют извлекать как простые статистические взаимосвязи, так и сложные нелинейные зависимости из данных. Однако обычно такие подходы зависят от большого количества размеченных данных и плохо переносятся на новые сетевые условия, где данные либо ограничены, либо имеют иную структуру.

Перенос обучения (Transfer learning или TL) предлагает эффективное решение этой проблемы, позволяя использовать знания, полученные в одной задаче или домене, для ускорения обучения и повышения точности в другой. Применение методов переноса обучения в анализе сетевого трафика открывает новые перспективы для создания универсальных и устойчивых систем сетевой безопасности, способных адаптироваться к изменяющимся условиям и различным типам сетей.

В данной статье рассматриваются ключевые подходы переноса обучения в контексте анализа сетевого трафика. Анализируются существующие исследования, обсуждаются преимущества и ограничения методов, а также направления для дальнейших исследований и применения на практике. Таким образом цель работы – провести систематический обзор методов переноса обучения, применимых к системам обнаружения вторжений и заложить основу для будущих исследований в этой области. В отличие от существующих обзоров, посвящённых либо общим методам переноса обучения в кибербезопасности, либо применению глубокого обучения в NIDS, данная работа систематизирует методы переноса обычения именно для сетевых систем обнаружения вторжений, предлагая сквозную классификацию по типу переносимых знаний и характеру сетевых данных (потоки или пакеты), а также оценивая применимость каждого подхода в практических сценариях

Остальная часть работы структурирована следующим образом. В разделе 2 содержатся основные определения и классификация методов переноса обучения. В разделе 3 приводится анализ существующих работ, относящихся к различным типам переноса обучения. В заключении приводятся выводы и возможные направления для будущих исследований.

2. Обзор

В этом разделе будут даны определения переноса обучения и сопутствующих терминов, а также представлена классификация методов переноса обучения.

2.1 Основные определения

В переносе обучения изученные знания передаются из исходной области (исходного домена) в целевую для улучшения процесса обучения целевой задачи. Таким образом, прежде всего будут даны определения «домена» и «задачи» [2].

Определение 1 (Домен) Доменом D является пара $D = \{\mathcal{X}, P(X)\}$, где $\mathcal{X}$ – признаковое пространство, $P(X)$ – маргинальное (или частное распределение), а X – множество объектов, т.е. $X = \{x \mid x_i \in \mathcal{X}, i = 1, \dots, n\}$.

Определение 2 (Задача) В домене D задача T состоит из двух частей: 1) пространства меток Y и 2) функции принятия решений $f(\cdot)$. Функция принятия решений обучается на основе пар векторов признаков и пространства меток, т.е. $\{x_i, y_i\}$, где $x_i \in X$ и $y_i \in Y$. Другими словами, задача определяется как $T = \{Y, f(\cdot)\}$.

В общем случае, функция принятия решений представляет собой предсказание соответствующей метки $f(x_i)$ для экземпляра x_i. В этом случае предсказательная функция может быть определена как $f(x_i) = \{P(y_k|x_i)|\ y_k \in Y, k = 1, \dots, |Y|\}$.

Определение 3 (Перенос обучения) Даны исходный домен D_S и задача T_S, а также целевой домен D_T и задача T_T. Перенос обучения направлен на улучшение качества обучения целевой предсказательной функции $f(\cdot)$ в области D_T за счет использования знаний, полученных в D_S и T_S, при условии, что $D_S \neq D_T$ или $T_S \neq T_T$.

2.2 Классификация

Перенос обучения представляет собой набор методов, позволяющих использовать знания, извлечённые из одной задачи или домена, для улучшения обучения в другой задаче. В отличие от традиционного машинного обучения, где обучающие и тестовые данные предполагаются независимыми и идентично распределёнными, TL допускает смещение распределений и различия между источником и целевой задачей [3]. Для достижения переноса обучения используется множество техник в зависимости от доменов и задач исходного и целевого доменов, их различий и способа переноса обучения. Таким образом, существует несколько вариантов классификации методов переноса [2].

2.2.1 Классификация по типу переносимых знаний

Первый и наиболее фундаментальный способ классификации методов переноса обучения – по типу переносимых знаний.

Перенос признаков (feature-based TL) предполагает, что перенос осуществляется через построение нового признакового пространства, в котором различия между исходным и целевым доменами минимизируются. Так, например, в задачах распознавания сигналов можно отобразить данные из разных частотных диапазонов в общее пространство признаков, что упрощает классификацию. Его можно разделить на две подкатегории: симметричный подход и асимметричный. В случае симметричного происходит преобразование признаков обеих областей в общее скрытое пространство. В ассиметричном напротив, преобразование только исходных признаков так, чтобы они соответствовали целевым.

В случае **переноса параметров** (parameter-based TL) перенос знаний осуществляется на уровне параметров или гиперпараметров моделей. Например, использование моделей, предобученных на ImageNet, для задач компьютерного зрения в смежных областях. В задачах анализа сетевого трафика это может означать перенос параметров модели, обученной на сетевом трафике одной организации, для анализа сетевого трафика другой.

Для **переноса отношений** (relation-based TL) перенос знаний осуществляется через использование связей и отношений между объектами. В отличие от предыдущих подходов, здесь внимание уделяется не только самим данным, но и структуре их взаимосвязей. Например, в социальных сетях это может быть перенос знаний о связях между пользователями; в беспроводных сетях – использование корреляций между узлами при решении задач маршрутизации или управления ресурсами.

Перенос экземпляров (instance-based TL) – знания переносятся на уровне исходных примеров (instances). Вместо прямого использования всех данных из исходного домена, методы этого типа направлены на перераспределение весов исходных образцов данных, чтобы сократить различия между маргинальными распределениями источника и цели.

2.2.2 Классификация в зависимости от соотношения доменов и задач

Вторым из наиболее распространённых критериев классификации переноса обучения является характер взаимосвязи между исходной и целевой задачами, а также наличие или отсутствие разметки в целевой области. На этой основе выделяют три ключевых типа переноса обучения: индуктивный, трансдуктивный и без учителя [3].

Определение 4 (Индуктивный перенос обучения, Inductive Transfer Learning). Пусть даны исходный домен D_S с соответствующей исходной задачей T_S и целевой домен D_T с соответствующей целевой задачей T_T. Индуктивный перенос обучения направлен на улучшение обучения целевой предсказательной функции $f_T(\cdot)$ целевого домена D_T на основе знаний, полученных в D_S и T_S, в которых исходная и целевая задача различаются, то есть $T_S \neq T_T$.

Индуктивный перенос обучения делится на два вида: самообучение (Self-taught learning) и многозадачное обучение (Multitask learning). Самообучение используется, когда в исходной области нет размеченных данных. Используются неразмеченные данные источника, чтобы построить более высокоуровневое представление признаков (снижается размерность пространства), а затем классификация проводится с использованием размеченных данных целевой области. Многозадачное обучение – когда и в исходной, и в целевой области есть размеченные данные. Обе задачи обучаются одновременно, что позволяет улучшать результаты друг друга.

В индуктивном переносе обучения знания могут быть перенесены через перенос признаков, перенос параметров, перенос отношений или перенос экземпляров.

Перенос признаков создаёт новое пространство X_{new}признаков для перевода объектов исходного и целевого домена в X_{new} . Перенос параметров предполагает, что отдельные модели для связных задач имеют общие параметры или распределение гиперпараметров. Подход на основе переноса отношений не предполагает, что данные, из каждого домена, являются независимыми. Наконец, перенос экземпляров позволяет переносить часть данных из исходного домена в целевой домен.

Определение 5 (Трансдуктивный перенос обучения, Transductive Transfer Learning) Пусть даны исходный домен D_S с соответствующей исходной задачей T_S и целевой домен D_T с соответствующей целевой задачей T_T. Трансдуктивный перенос обучения направлен на улучшение обучения целевой предсказательной функции $f_T(\cdot)$ целевого домена D_T на основе знаний, полученных в D_S и T_S, где исходный домен и целевой домен различны, т. е. $D_S \neq D_T$, а исходная и целевая задача одинаковы, т. е. $T_S = T_T$.

В данном сценарии в исходном домене имеется большой объем размеченных данных, в то время как в целевом домене метки отсутствуют вовсе. Ключевое различие доменов при идентичности задач обычно обусловлено либо несовпадением пространств признаков (например, разные языки в текстовых задачах), либо различием в маргинальных распределениях вероятностей входных данных, когда $P(X_S) \neq P(X_T)$. Данный подход широко известен как адаптация домена (Domain Adaptation). В трансдуктивном переносе основное внимание уделяется минимизации расхождения между доменами. Это достигается либо через взвешивание экземпляров исходного домена (Instance-based), чтобы они больше соответствовали распределению целевого, либо через поиск инвариантного представления признаков (Feature-based), которое одинаково эффективно описывает данные обоих доменов, несмотря на их исходные различия

Определение 6 (Перенос обучения без учителя, Unsupervised Transfer Learning) Пусть даны исходный домен D_S с соответствующей исходной задачей T_S и целевой домен D_T с соответствующей целевой задачей T_T. Перенос обучения без учителя направлен на улучшение обучения целевой предсказательной функции $f_T(\cdot)$ целевого домена D_T на основе знаний, полученных в D_S и T_S, в которых исходная и целевая задача различаются, т. е. $T_S \neq T_T$, а размеченные данные Y_S и Y_T не поддаются наблюдению.

В переносе обучения без учителя знания могут быть перенесены только через перенос признаков. В этом случае переносимыми знаниями являются структурные и представительные характеристики данных. Модель, выявляющая скрытые закономерности или латентные структуры в одном наборе данных, может быть использована для анализа другого набора данных без явных меток. Такой тип переноса полезен для кластеризации, снижения размерности или предварительного обучения представлений в целевом домене.

Сравнение методов переноса с точки зрения переносимых знаний представлено в табл. 1.

Табл. 1. Сравнение методов переноса знаний.
Table 1. Comparison of knowledge transfer methods.

	Перенос признаков	Перенос параметров	Перенос отношений	Перенос экземпляров
Индуктивный перенос обучения	+	+	+	+
Трансдуктивный перенос обучения	+	-	-	+
Перенос обучения без учителя	+	-	-	-

Таким образом наиболее гибким является индуктивным перенос обучения, поддерживающий перенос любых типов знаний. Менее гибким является трансдуктивный перенос, который не поддерживает перенос параметров и перенос отношений.

2.2.3 Стратегии переноса в глубоком обучении

Перенос в глубоком обучении или глубокий перенос обучения (англ. Deep Transfer Learning, DTL) – рассматривает перенос знаний для глубоких нейронных сетей, обученных на одном домене или задаче, в другой домен/задачу. В отличие от классического TL, здесь используются предобученные глубокие модели.

Разделяют 4 основных типа [4]:

- Готовые предварительно обученные модели (Off-the-shelf pretrained models);
- Извлечение признаков с помощью предобученной модели (Feature extraction);
- Дообучение (Fine-tuning);
- Гибридные методы.

Стратегия готовых предварительно обученных моделей представляет из себя прямое использование предварительно обученной модели для целевой задачи, без модификаций. Этот метод эффективен, если источник и целевая задача совпадают или очень близки, т.е. $D_S \cong D_T$ и $T_S \cong T_T$.

При использовании предобученной модели для извлечения признаков обученная модель используется для получения более абстрактных и обобщенных представлений признаков. Этот метод эффективен при ограниченных данных в целевом домене и снижает вычислительные затраты при обучении.

Глубокое обучение можно рассматривать как иерархическое обучение. В частности, отдельный слой DNN обучается различным признакам, от общих, т. е. низкого уровня, до более конкретных, т. е. высокого уровня, по мере углубления в DNN. Было доказано, что

признаки, обученные DL, более переносимы, что означает, что их легче повторно использовать в похожих доменах [5].

Дообучение представляет собой стратегию передачи знаний, при которой предварительно обученная нейронная сеть дообучается под конкретную целевую задачу. Этот подход заключается в том, что веса предварительно обученной модели используются как начальная точка для обучения новой модели, что значительно сокращает время и ресурсы, необходимые для достижения высокой точности. У данной стратегии есть 2 основных подхода: инициализация весов и выборочное дообучение. При инициализации весов веса предварительно обученной модели используются как начальные значения для целевой модели, а затем происходит обучение модели на целевых данных. В выборочном дообучении предполагается дообучение только части модели, оставляя другие слои замороженными. Это особенно эффективно при ограниченных данных.

3. Применимость к задаче обнаружения вторжений

Одной из ключевых проблем построения сетевых систем обнаружения вторжений, основанных на методах машинного обучения, является ограниченная переносимость обученных моделей. Наборы данных для обучения классификаторов сетевого трафика часто создаются в строго определённой среде (например, в конкретной корпоративной сети, в лабораторных условиях или с использованием специализированных бенчмарков, таких как CICIDS или UNSW-NB15 [6]). Однако в реальных сценариях эксплуатации трафик подвержен постоянным изменениям: появляются новые протоколы, меняется поведение приложений, возникают ранее неизвестные атаки. Также возникает проблема при использовании пред обученных моделей так как они обучались в иной среде. В этой ситуации возникает проблема смещения распределений, которая резко снижает эффективность детекторов, обученных в одной среде и перенесённых в другую [7].

Методы переноса обучения позволяют существенно смягчить эту проблему, обеспечивая адаптацию моделей к новым условиям без необходимости полного переобучения «с нуля». Их применимость в контексте СОВ можно рассмотреть через призму используемых представлений входных данных, так как именно они во многом определяют эффективность и тип используемого переноса знаний.

3.1 Статистические признаки потоков

В традиционных системах сетевого мониторинга часто используются агрегированные статистические признаки потоков (flow-based features), такие как количество пакетов, средний размер сегмента, длительность соединения, дисперсия межпакетных интервалов и др. Данные признаки относительно компактны, легко интерпретируемы и позволяют применять широкий спектр классических алгоритмов машинного обучения. Однако такие признаки подвержены высокой зависимости от среды: одни и те же типы трафика могут иметь существенно разные статистические характеристики в разных сетях.

Здесь перенос обучения на основе переноса экземпляров (например, методы повторного взвешивания образцов или фильтрации нерелевантных потоков) позволяет адаптировать распределения признаков источника к целевой сети. Кроме того, перенос обучения на основе переноса признаков (например, доменная адаптация через автоэнкодеры) помогает проектировать более устойчивое представление потоков, снижая чувствительность к конкретным условиям трафика.

Примером переноса обучения на основе переноса экземпляров является работа [8], предложившая перенос обучения на основе перераспределения весов образцов между доменами. Авторы рассматривают проблему, типичную для NIDS на основе анализа статистических признаков потоков: статистические параметры потоков в целевой сети могут

существенно отличаться от тех, что использовались при обучении модели, т.е. домены различны. Для решения этой проблемы предлагается **перенос обучения на основе экземпляров**, в котором каждому образцу исходного домена назначается параметр-вес, определяющий его релевантность целевой задаче. В основе метода лежит идея о том, что часть данных источника близка к данным целевой сети, а часть – вводит шум и снижает точность. Авторы не рассматривают случаи появления новых типов трафика, т.е. $T_S \neq T_T$. Следовательно, рассматривается **трансдуктивный перенос обучения**.

Для вычисления оптимальных весов авторы применяют критерий максимального среднего расхождения (MMD), который минимизирует расстояние между распределениями признаков источника и цели. Далее классификатор обучается с учётом сгенерированных весов: релевантные примеры усиливают влияние при обучении, а нерелевантные – подавляются. Такой подход позволяет адаптировать модель без изменения её архитектуры, используя только перераспределение важности отдельных потоков. Ключевым преимуществом является высокая совместимость с классическими методами анализа данных со статистическими признаками сетевых потоков (градиентный бустинг, SVM, нейросети малой глубины). Итоги экспериментов показали повышение устойчивости к сетевым сдвигам, особенно при переносе между наборами данных с разной плотностью трафика. Однако точность сильно зависит от качества оценки расстояния между доменами: при высокой неоднородности потоков в целевой сети подбор весов становится менее стабильным.

В работе [9] был предложен гибридный подход, сочетающий генеративно-состязательную сеть (GAN), многоядерное максимальное среднее расхождение (MK-MMD) и оптимизацию признаков для выявления аномалий в сетевом трафике. На первом этапе выполняется оптимизация признаков с помощью ансамбля агентов на основе обучения с подкреплением (Collaborative Learning Automata) которые взаимодействуют с классификатором как со стохастической средой и итеративно корректируют вероятности выбора отдельных признаков в зависимости от полученной точности классификации. Такая совместная адаптация нескольких автоматов позволяет исключить коррелированные и избыточные признаки, сократить размерность пространства и улучшить информативность входных данных. Полученное оптимальное подмножество (около 9 признаков) используется на втором этапе, где применяется генеративно-состязательная сеть (GAN). В её структуре генератор создаёт синтетические примеры сетевого трафика на основе случайного шума и условных меток, а дискриминатор учится одновременно различать реальные и сгенерированные образцы и классифицировать их по типу атаки. Таким образом, GAN выполняет роль **переноса обучения без учителя** (semi-supervised transfer-learning), расширяя выборку и повышая устойчивость модели при ограниченном количестве размеченных данных. Для дальнейшего повышения обобщающей способности в архитектуру внедрён модуль многоядерного варианта меры максимального среднеквадратичного расхождения (multiple kernel variant of maximum mean discrepancy, MK-MMD) [10], минимизирующий различия между распределениями исходной (source) и целевой (target) областей в пространстве признаков; тем самым реализуя доменную адаптацию – перенос знаний о нормальном и атакующем трафике между различными сетевыми средами. Обучение всех компонентов проводится совместно: формирование оптимизированного вектора признаков, генерация дополнительных образцов, выравнивание их распределений с помощью MK-MMD и затем классификатор на выходе (16-мерный Softmax) выдаёт окончательные вероятности классов. Итоговая система продемонстрировала высокую точность (около 91,7 % для бинарной и 91,5 % для многоклассовой классификации), низкий уровень ложных срабатываний и устойчивость при переносе между доменами, объединяя в едином конвейере принципы обучения с подкреплением, генеративного моделирования и адаптивного переноса признаков.

Таким образом сильной стороной предложенного авторами метода является эффективная компенсация недостатка данных об аномалиях и общее улучшение устойчивости

классификатора к сдвигам во входных данных. Однако генерация синтетических признаков потоков может не полностью отражать реальные зависимости между признаками при эксплуатации данного метода в реальных условиях.

К аналогичному направлению относится работа [11], однако здесь авторами предложен метод обнаружения неизвестных сетевых атак на основе **трансдуктивного переноса признаков**, направленный на уменьшение зависимости от конкретной среды и набора данных. Авторы используют набор NSL-KDD, содержащий 41 признак, и формируют два домена: исходный, включающий известные атаки, и целевой, содержащий другие типы атак, которые моделируют «неизвестные» угрозы. Основная идея работы заключается в создании общего латентного пространства, в котором данные из исходного и целевого доменов становятся сопоставимыми. Первым вариантом метода является HeTL [12], который может находить общее скрытое подпространство двух разных атак и обучаться оптимизированному представлению, которое было инвариантным к изменениям поведения атак. Он находит две матрицы отображения, которые минимизируют искажение исходных данных и одновременно уменьшают расстояние между представлениями двух доменов. Однако HeTL чувствителен к выбору гиперпараметра, определяющего близость доменов, и требует ручной настройки. Чтобы устранить эту проблему, авторы создают улучшенный метод CeHTL, который автоматически определяет соответствие между доменами с помощью кластеризации. Этот подход позволяет избежать ручной настройки параметров, делает метод более устойчивым и позволяет работать даже при разных наборах признаков в доменах.

Экспериментальная часть показывает, что обычные классификаторы (SVM, KNN, деревья решений) плохо справляются с обнаружением новых атак, в то время как HeTL и особенно CeHTL существенно повышают accuracy, F1 и AUC.

В свою очередь, исследование [13] развивает идею **переноса отношений**. Статья предлагает модель, предназначенную для того, чтобы устранить ключевую проблему современных систем обнаружения вторжений на основе машинного обучения: их статичность и неспособность своевременно адаптироваться к новым атакам. Авторы используют два типа данных – признаки потоков и структурированные CTI-отчёты (Cyber Threat Intelligence, Аналитика киберугроз) VirusTotal. Первый набор данных включает почти три миллиона сетевых потоков с большим набором статистических признаков, полученных из корпоративного трафика, а второй состоит более чем из двух тысяч отчётов, содержащих до сотни атрибутов, включая временные метрики, репутацию IP и результаты анализа различными анализаторами. На основе этих данных создаётся гибридная модель обнаружения вторжений, сочетающая метод опорных векторов (SVM) и метод ближайших соседей (K-means): первая отвечает за классификацию известных угроз, вторая выявляет аномалии и формирует класс выбросов (outlier), который запускает обращение к CTI. При обнаружении неопределённого трафика из него извлекается IoC, производится CTI поиск (CTI-lookup), после чего CTI Transfer Model на основе метода ближайших соседей анализирует отчёт, определяет характер IoC и сопоставляет его с исходными сетевыми наблюдениями, автоматически формируя новое обучающее наблюдение. Этот процесс обеспечивает непрерывное дообучение IDS. Использование CTI приводит к росту F1-метрики на 9,29% по сравнению с моделью без CTI, а применение ML внутри CTI-модуля даёт выигрыш 30,92% по сравнению с простыми эвристическими правилами. Особенно значимыми оказываются данные из класса выбросов, поскольку именно по ним модель испытывает наибольшую неопределённость и получает максимальное улучшение после интеграции CTI. В результате DICI демонстрирует способность распознавать сложные динамические атаки, которые ускользают от традиционных IDS, оставаясь при этом вычислительно лёгкой и пригодной для онлайн-обновления.

Такой подход обеспечивает постоянную адаптацию к текущей обстановке в сети, снижая задержку между появлением новой угрозы и её детектированием. В работе реализованы

механизмы контроля качества обновлений и отката к предыдущим версиям модели, что делает систему эксплуатационно надёжной. Вместе с тем, зависимость от полноты и достоверности CTI-источников создаёт риск некорректных обновлений, а интеграция такого подхода в корпоративную инфраструктуру требует значительных ресурсов.

В статье [14] рассматривается построение IDS на основе **параметрического переноса обучения**, при котором из исходного домена в целевой переносится не структура данных и не примеры, а параметры глубокой модели – архитектура и обученные веса CNN-LSTM. Данные в обоих доменах представлены одинаково – как статистические признаки сетевых потоков, однако распределения отличаются: исходный домен использует обучающую и валидационную выборки, а целевой – полностью невидимые тестовые данные.

Такой перенос относится к **индуктивному переносу обучения**: задача классификации одинакова в обоих доменах, но данные не идентичны, поэтому модель использует заранее выученные параметрические представления признаков для работы в новой среде. В исходном домене CNN-LSTM обучается на полном наборе данных и формирует внутренние представления трафика; в целевом домене она применяет эти же веса без дообучения, что позволяет сохранить точность и существенно ускорить обработку при минимальных ресурсах.

Результаты показывают, что параметрический перенос обеспечивает точность выше 98% и повышает скорость инференса в целевом домене, демонстрируя, что заранее обученная глубинная модель способна эффективно работать в условиях ограниченных ресурсов и умеренного сдвига распределений.

В работе [15] авторы используют форму переноса обучения, которая сочетает перенос параметров и перенос представлений. С точки зрения соотношения доменов и задач их подход относится к индуктивному переносу, поскольку исходная и целевая задачи остаются одинаковыми – это многоклассовая классификация сетевого трафика, однако домены различаются, потому что Bot-IoT и TON-IoT формируют разные распределения признаков и содержат отличающиеся варианты атак. Модель, обученная на Bot-IoT, служит источником уже сформированных признаковых представлений, и эти представления переносятся на целевой домен, где присутствуют новые типы поведения тех же классов и иная статистика трафика. Замороженная свёрточная основа фактически переносит знания о структуре сетевого трафика, зафиксированные в весах, тогда как новый классификатор обучается на TON-IoT, адаптируя общее представление к характеристикам целевого домена.

Преимущество такого подхода заключается в том, что он позволяет использовать накопленные знания об общих паттернах сетевого поведения и при этом адаптировать модель к изменениям без переобучения всей сети. Он уменьшает требования к объёму данных и вычислительным ресурсам и компенсирует нехватку размеченных примеров в целевом датасете. Однако эта стратегия уязвима к сильному расхождению доменов: если распределения в Bot-IoT и TON-IoT различаются слишком существенно, перенесённая основа может кодировать устаревшие или нерелевантные признаки, ограничивая качество обновлённой модели. Кроме того, за счёт заморозки основной части сети остаётся риск того, что модель будет хуже адаптироваться к принципиально новым видам аномалий, если они требуют изменения низкоуровневых признаков, а не только обновления классификатора.

3.2 Последовательность пакетов

Современные подходы также предлагают представление сетевых данных в виде «сырых» пакетов, рассматривая первые n байт полезной нагрузки и заголовков в качестве входных данных для нейронных сетей. Такой подход ближе к компьютерному зрению или обработке текста, где модели могут самостоятельно выявлять значимые паттерны без ручной инженерии признаков.

В статье [16] анализируются атаки на автомобильную шину CAN (Controller Area Network, сеть контроллеров), где новые варианты вторжений появляются часто, а данные для обучения минимальны. Авторы используют реальные массивы CAN-трафика и преобразуют каждый кадр в компактное числовое представление из одиннадцати признаков, включая интерпакетный интервал и нормализованные байты полезной нагрузки. Из таких последовательностей формируются временные окна, которые затем переводятся в двумерное пространственно-временное представление, подходящее для сверточной модели с долгой краткосрочной памятью. Такая модель обучается на нормальном трафике и известной DoS-атаке, после чего используется один из методов **индуктивного переноса экземпляров**, а именно одномоментный перенос обучения (one-shot transfer learning), позволяющий дообучить систему на единственном примере нового типа атаки. Такой тип переноса позволяет переносить уже извлечённые закономерности на новую задачу, в которой имеется крайне мало размеченных данных. После обучения на известных вторжениях и дообучения на одной выборке новой атаки модель достигла F1=88,47% при обнаружении новых атак, что демонстрирует потенциал переноса экземпляров в задачах сетевой безопасности. Его достоинства проявляются в резком увеличении точности обнаружения новых атак, снижении количества ложных срабатываний и отсутствии необходимости собирать крупные датасеты для каждого нового класса. Основным недостатком является высокая зависимость результата от качества исходного представления данных и предварительного обучения, а также риск того, что единственный пример новой атаки может оказаться нерепрезентативным, что приведёт к деградации обобщающей способности модели.

Дополнительно возможен перенос из смежных доменов – например, использование архитектур, обученных на задаче классификации протоколов или идентификации приложений, для задачи обнаружения аномалий. В таких сценариях извлечение признаков оказывается особенно полезным: первые слои сети фиксируются как универсальные извлекатели признаков, а финальные слои перенастраиваются на задачу обнаружения атак.

В более поздних работах идея переноса репрезентаций для задач сетевой безопасности получила развитие в контексте интернета вещей (IoT), где ограниченность вычислительных ресурсов и разнообразие сетевых топологий требуют особой гибкости моделей.

В работе [17] авторы строят IDS-фреймворк на основе **индуктивного переноса признаков**, где перенос осуществляется между двумя наборами сетевого трафика, имеющими одинаковое пространство признаков (15 общих параметров), но различное распределение и различное происхождение. В этом подходе предварительно обученная на BoT-IoT[18] модель CNN служит источником сверточных фильтров, а затем эти слои переносятся в целевой домен UNSW-NB15[19], где обучается только классификатор.

В статье подчеркивается, что использование переноса обучения позволяет извлечь устойчивые закономерности сетевого поведения из масштабного BoT-IoT и перенести их на менее репрезентативный UNSW-NB15, где часть атак относится к атакам нулевого для (zero-day). В результате предварительно обученная сверточная база служит универсальным извлекателем признаков, а дообучаемые полносвязные слои адаптируются к целевому распределению. Такой подход обеспечивает высокую точность при обнаружении неизвестных атак, поскольку модель опирается на общие низкоуровневые сетевые паттерны, а не только на специфические сигнатуры.

Достоинством метода является то, что он позволяет компенсировать нехватку данных в целевом домене и значительно улучшает показатели обнаружения новых атак благодаря переносу универсальных признаков, извлечённых из источника с большим количеством экземпляров. Дополнительным преимуществом становится уменьшение переобучения: замороженные сверточные слои создают устойчивое основание для классификатора, что снижает чувствительность к шуму и дисбалансу данных. Недостатком является то, что перенос возможен только при совпадении признакового пространства, поэтому авторы были

вынуждены сократить данные до набора из 15 общих признаков, что автоматически означает потерю потенциально важной информации. Ограничением остаётся и то, что оба датасета являются синтетическими; распределения сетевого трафика в реальных IoT-средах значительно сложнее, и перенос, выполненный в однородном признаковом пространстве, может оказаться менее эффективным при переходе от симулированного домена к реальному. Кроме того, рассматривается бинарная классификация, и такой тип TL не решает задачу точного различения множества конкретных семейств атак, что ограничивает практическое применение на уровне оперативного реагирования.

Таким образом, среди исследуемых работ пакетный анализ представлен двумя направлениями: анализ последовательностей низкоуровневых кадров в CAN-сетях и анализ IoT-пакетов на уровне байтовых структур. Оба подхода демонстрируют преимущества в контекстах, где статистические потоки либо невозможно сформировать, либо они теряют важную информацию о вредоносных изменениях в полезной нагрузке. Однако пакетные методы требуют больших вычислительных ресурсов, чувствительны к изменению формата пакетов и часто менее устойчивы при переносе между доменами по сравнению с потоково-ориентированными системами, что делает задачу переноса обучения особенно важной.

Другим примером индуктивного обучения является работа [20]. В этой работе авторы решают смежную с обнаружением вторжений задачу, а именно классификацию веб-приложений. На этапе предобучения решаются вспомогательные задачи самоконтролируемого обучения (self-supervised tasks) – распознавание зашифрованных пакетов по значению энтропии и определение непрерывности потока, тогда как целевая задача – это многоклассовая классификация приложений. Таким образом, переносимые знания представляют собой неспецифические высокоуровневые представления, отражающие статистические и структурные закономерности зашифрованного трафика: распределение байтовых значений, временные паттерны, слабые межпакетные зависимости. Эти признаки извлекаются свёрточной нейронной сетью (CNN) на уровне отдельных пакетов и обобщаются кодировщиком на основе архитектуры BERT на уровне последовательностей пакетов.

Достоинством такого подхода является высокая адаптивность: модель не требует размеченных данных на этапе предобучения, а дообучение (fine-tuning) на целевом наборе данных возможно даже при небольшом объёме размеченных примеров (20–40%), что особенно важно в практических сценариях, где аннотация сетевого трафика трудоёмка и дорогостояща. Кроме того, отсутствие зависимости от заголовков пакетов – в том числе от пятёрки ключевых полей (5-tuple: исходный IP, исходный порт, целевой IP, целевой порт, транспортный протокол) – делает перенос между доменами более устойчивым: например, становится возможной адаптация между наборами данных, собранными в различных сетевых средах.

Недостатком является высокая вычислительная сложность самой архитектуры CBD: большое число параметров, особенно при использовании 12-уровневого кодировщика BERT, делает обучение чувствительным к объёму размеченных данных в целевом домене; без предобучения модель склонна к переобучению или несходимости даже на относительно простых задачах. Также следует отметить, что предложенный подход к переносу обучения остаётся замкнутым по отношению к неизвестным классам поскольку как предобучение, так и дообучение ориентированы на фиксированный набор классов, и полученные обобщённые признаки могут оказаться недостаточными для обнаружения принципиально новых, ранее не встречавшихся типов трафика.

3.3 Преимущества и ограничения переноса обучения для NIDS

Классификация приведенных выше примеров с указанием типа анализируемых данных (полужирным выделены методы на основе анализа статистических признаков) приведена в табл. 2.

Табл. 2. Классификация примеров.
Table 2. Classification of examples.

	Перенос признаков	Перенос параметров	Перенос отношений	Перенос экземпляров
Индуктивный перенос обучения	[17], [20], **[15]**	**[14]**	**[13]**	[16]
Трансдуктивный перенос обучения	**[11]**	–	–	**[8]**
Перенос обучения без учителя	**[9]**	–	–	–

Методы, основанные на анализе отдельных сетевых пакетов, обладают явным преимуществом в полноте информации: они позволяют извлекать признаки непосредственно из структуры заголовков и полезной нагрузки, что обеспечивает выявление сложных атак, скрытых на уровне байтовых последовательностей. Благодаря этому подходы на основе анализа пакетов особенно эффективны против угроз ориентированных на полезную нагрузку, таких как внедрение SQL-кода, инъекции кода и атаки на протоколы прикладного уровня. Дополнительным преимуществом является высокая чувствительность к малозаметным закономерностям в поведении атакующих, которые могут быть потеряны при агрегации в сетевые потоки. В контексте переноса обучения наличие богатого низкоуровневого представления создаёт условия для переноса универсальных признаков, особенно связанных с протокольной структурой и вредоносными шаблонами полезной нагрузки.

Однако высокая детализация ведёт к существенным минусам. Анализ пакетов требует значительно большего объёма вычислительных ресурсов, как для хранения, так и для обработки данных. Модели на таких входах склонны сильнее деградировать при смене домена, поскольку структура пакетов может существенно различаться между сетями, прошивками, устройствами и версиями протоколов. Это снижает устойчивость переноса обучения на основе анализа пакетов и требует дополнительного дообучения. Кроме того, многие современные сети используют шифрование, что делает полезную нагрузку недоступной для анализа и резко снижает эффективность методов, основанных на содержимом пакета.

Методы, основанные на признаках потоков, напротив, обладают высокой масштабируемостью и существенно меньшими требованиями к ресурсам. Они дают компактное обобщение поведения соединения, что делает такие модели более стабильными и переносимыми между различными сетями. Они устойчивы к шифрованию, поскольку используются только метаданные (размеры пакетов, время межпакетных интервалов, количество направленных пакетов), что позволяет эффективно работать в условиях современного TLS-ориентированного трафика. На практике именно методы на основе анализа признаков потоков чаще демонстрируют хорошую адаптивность к новым сетям и реже требуют переработки признаков.

Однако агрегированная природа потоков неизбежно приводит к потере информации. Потери касаются прежде всего полезной нагрузки и последовательности пакетов внутри соединения, что делает такие модели уязвимыми к атакующим, использующим нестандартные атаки или атакующие последовательности, проявляющиеся только на уровне отдельных пакетов. Таким

образом, такие методы часто уступают в обнаружении сложных и ранее неизвестных атак, особенно тех, что не изменяют метаданные соединения.

Отдельно следует учитывать риск негативного переноса [21]. При использовании моделей, обученных на пакетных данных, негативный перенос проявляется особенно часто: низкоуровневые признаки, извлечённые из заголовков и полезной нагрузки, могут быть тесно связаны с конкретными конфигурациями сети, типами устройств, версиями прошивок или даже характерными шаблонами трафика отдельных инфраструктур. При переносе на новую сеть такие признаки перестают быть релевантными и могут вводить модель в заблуждение, снижая точность и увеличивая число ложных срабатываний. В потоковых методах негативный перенос выражен слабее, однако и здесь обобщённые статистические характеристики соединений могут систематически отличаться между доменами: например, варьируются профили задержек, средние размеры пакетов или характер распределения длительностей соединений. В результате модель начинает интерпретировать безвредный трафик как аномальный или игнорировать признаки реальных атак. Таким образом, негативный перенос остаётся фундаментальной проблемой для обоих подходов, особенно при использовании данных из сильно разнородных сетевых сред.

3.4 Обобщение методов

Обобщив примеры и анализ выше, рассмотрим основные виды переноса, применяемые в NIDS, их ключевые идеи, преимущества и ограничения. Такое сравнение позволяет оценить, какие методы наиболее практичны в условиях ограниченных данных, меняющихся распределений трафика и появления неизвестных атак.

Индуктивный перенос признаков использует предобученную CNN с замороженными слоями, обученную на BoT-IoT и адаптированную на UNSW-NB15. Такой подход улучшает качество при нехватке данных и снижает переобучение, но требует совпадения признаков и в основном подходит для бинарных задач.

Индуктивный перенос параметров основан на передаче знаний от модели, обученной на исходной задаче, к новой задаче через инициализацию или частичную донастройку весов. Веса свёрточной части CNN, предобученной на Bot-IoT, фиксируются и используются как обобщённые признаковые экстракторы, а верхние полносвязные слои дообучаются на небольшом целевом датасете (TON-IoT), что позволяет эффективно адаптировать модель к новым атакам без полного переобучения и с минимальными вычислительными затратами.

Индуктивный перенос отношений сочетает SVM, K-means и KNN, используя корпоративный трафик и отчёты VirusTotal. Он улучшает F1 и качество обнаружения новых атак, но сильно зависит от качества CTI-данных и требует значительных ресурсов.

Индуктивный перенос экземпляров применяет Conv-LSTM с обучением по одному примеру атаки (one-shot) на реальном CAN-трафике. Он обеспечивает высокую точность и низкое число ложных срабатываний, но чувствителен к качеству единственного обучающего примера.

Трансдуктивный перенос признаков (HeTL и CeHTL) работает на NSL-KDD, выравнивая скрытые пространства исходного и целевого доменов. Он улучшает обнаружение новых атак; CeHTL не требует ручного тюнинга, тогда как HeTL чувствителен к параметрам.

Трансдуктивный перенос экземпляров использует перевзвешивание данных через MMD в сочетании с классическими моделями. Он помогает адаптироваться к дрейфу распределений, но зависит от точности измерения расстояния между доменами.

Перенос без учителя объединяет GAN, RL-оптимизацию признаков, MMD и Softmax-классификацию, применяя это к многоклассовому трафику. Он обеспечивает высокую точность и снижает смещение распределений, хотя синтетические данные не всегда полностью отражают реальные зависимости.

Рассмотренные подходы к переносу обучения демонстрируют, что использование знаний из других доменов позволяет существенно повысить качество и устойчивость систем обнаружения вторжений. Каждый тип переноса решает свою задачу: одни помогают справиться с нехваткой данных, другие – адаптироваться к новым атакам или изменению распределений трафика. Несмотря на имеющиеся ограничения, методы переноса уменьшают потребность в полном переобучении моделей и делают NIDS более гибкими и эффективными в условиях быстро меняющейся сетевой среды.

4. Заключение

Перенос обучения в сетевых системах обнаружения вторжений представляет собой ключевое направление, позволяющее преодолеть фундаментальные ограничения традиционных NIDS, связанные с изменчивостью трафика, дрейфом распределений и появлением новых типов атак. Анализ представленных работ показывает, что различные парадигмы переноса – от переноса признаков и моделей до переноса отношений и экземпляров – по-разному раскрывают потенциал адаптации систем обнаружения вторжений к новым сетевым средам, обеспечивая устойчивость и повышенную обобщающую способность при работе в условиях ограниченной разметки и неоднородности доменов.

Методы, основанные на статистических признаках потоков, демонстрируют высокую переносимость и устойчивость к шифрованию, что делает их особенно привлекательными для практического применения. При этом существенным ограничением остаётся неизбежная потеря информации о структуре отдельных пакетов, что снижает способность обнаруживать сложные аномалии. Напротив, подходы, использующие последовательности пакетов и их низкоуровневое представление, обладают большей детализированностью и позволяют выявлять тонкие закономерности поведения атакующих, однако значительно чувствительнее к смене домена и требуют больших вычислительных ресурсов.

Показательно, что в современных исследованиях всё чаще встречаются гибридные и многоуровневые стратегии переноса, сочетающие методы доменной адаптации, генеративные модели, оптимизацию признаков и непрерывное обновление на основе CTI-данных. Такие системы демонстрируют способность к эволюции в реальном времени, что критически важно для противодействия динамическим и малоизвестным угрозам.

Тем не менее, перенос обучения для NIDS сталкивается с рядом ограничений: риском негативного переноса, зависимостью от степени близости доменов, высокой сложностью архитектур, а также отсутствием универсальных наборов данных, адекватно отражающих разнообразие реальных сетевых условий. Решение этих проблем требует развития методов, способных учитывать дрейф, работать с частично размеченными или полностью неразмеченными данными, а также интегрироваться с потоковыми алгоритмами машинного обучения.

Таким образом, перенос обучения формирует основу для нового поколения интеллектуальных NIDS, способных адаптироваться к быстро меняющимся условиям и противостоять угрозам, ранее недоступным для классических методов, однако дальнейший прогресс в этой области требует как теоретических исследований, так и систематических практических оценок в реалистичных сетевых сценариях.

Список литературы / References

[1]. Liao H. J. et al. Intrusion detection system: A comprehensive review //Journal of network and computer applications. – 2013. – Т. 36. – №. 1. – С. 16-24.

[2]. Zhuang F. et al. A comprehensive survey on transfer learning //Proceedings of the IEEE. – 2020. – Т. 109. – №. 1. – С. 43-76.

[3]. Pan S. J., Yang Q. A survey on transfer learning //IEEE Transactions on knowledge and data engineering. – 2009. – Т. 22. – №. 10. – С. 1345-1359.

[4]. Nguyen C. T. et al. Transfer learning for wireless networks: A comprehensive survey //Proceedings of the IEEE. – 2022. – Т. 110. – №. 8. – С. 1073-1115.
[5]. Yosinski J. et al. How transferable are features in deep neural networks? //arXiv preprint arXiv:1411.1792. – 2014.
[6]. Ring M. et al. A survey of network-based intrusion detection data sets //Computers & security. – 2019. – Т. 86. – С. 147-167.
[7]. Wang M. et al. On the robustness of ML-based network intrusion detection systems: An adversarial and distribution shift perspective //Computers. – 2023. – Т. 12. – №. 10. – С. 209.
[8]. Wu P., Guo H., Buckland R. A transfer learning approach for network intrusion detection //arXiv preprint arXiv:1909.02352. – 2019.
[9]. Ma W. et al. Abnormal traffic detection based on generative adversarial network and feature optimization selection //International Journal of Computational Intelligence Systems. – 2021. – Т. 14. – №. 1. – С. 1170-1188.
[10]. Gretton A. et al. A kernel two-sample test //The journal of machine learning research. – 2012. – Т. 13. – №. 1. – С. 723-773.
[11]. Zhao J. et al. Transfer learning for detecting unknown network attacks //EURASIP Journal on Information Security. – 2019. – Т. 2019. – №. 1. – С. 1-13.
[12]. Zhao J., Shetty S., Pan J. W. Feature-based transfer learning for network security //MILCOM 2017-2017 IEEE Military Communications Conference (MILCOM). – IEEE, 2017. – С. 17-22.
[13]. Lin Y. D. et al. Evolving ML-based Intrusion Detection: Cyber Threat Intelligence for Dynamic Model Updates //IEEE Transactions on Machine Learning in Communications and Networking. – 2025.
[14]. Dhillon H., Haque A. Towards network traffic monitoring using deep transfer learning //2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom). – IEEE, 2020. – С. 1089-1096.
[15]. Idrissi I., Azizi M., Moussaoui O. Accelerating the update of a DL-based IDS for IoT using deep transfer learning //Indones. J. Electr. Eng. Comput. Sci. – 2021. – Т. 23. – №. 2. – С. 1059-1067.
[16]. Tariq S., Lee S., Woo S. S. CANTransfer: Transfer learning based intrusion detection on a controller area network using convolutional LSTM network //Proceedings of the 35th annual ACM symposium on applied computing. – 2020. – С. 1048-1055.
[17]. Rodríguez E. et al. Transfer-learning-based intrusion detection framework in IoT networks //Sensors. – 2022. – Т. 22. – №. 15. – С. 5621.
[18]. BoT IoT Dataset, Available at: https://research.unsw.edu.au/projects/bot-iot-dataset, accessed 20.11.2025.
[19]. Moustafa N., Slay J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set) //2015 military communications and information systems conference (MilCIS). – IEEE, 2015. – С. 1-6.
[20]. Hu X. et al. CBD: A deep-learning-based scheme for encrypted traffic classification with a general pre-training method //Sensors. – 2021. – Т. 21. – №. 24. – С. 8231.
[21]. Wang Z. et al. Characterizing and avoiding negative transfer //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2019. – С. 11293-11302.

Информация об авторах / Information about authors

Антон Юрьевич ПОКИДЬКО – стажер-исследователь отдела компиляторных технологий ИСП РАН. Научные интересы: дрейф в машинном обучении и нейронных сетях, перенос обучения, анализ сетевого трафика.

Anton Yurevich POKIDKO – research intern at Compiler Technology department of ISP RAS. Research interests: drift in machine learning and neural networks, transfer learning, network traffic analysis.

Иван Александрович СТЕПАНОВ – аспирант, стажёр-исследователь ИСП РАН, ассистент кафедры информатики и вычислительной математики МФТИ. Сфера научных интересов:анализ сетевого трафика с помощью машинного обучения.

Ivan Alexandrovich STEPANOV – postgraduate student of the ISP RAS, intern researcher at ISP RAS, an assistant at the Department of Computer Science and Computational Mathematics at MIPT. Research interests: network traffic analysis using machine learning.

Александр Игоревич ГЕТЬМАН – кандидат физико-математических наук, старший научный сотрудник ИСП РАН, ассистент ВМК МГУ, доцент ВШЭ и МФТИ. Сфера научных интересов: анализ бинарного кода, восстановление форматов данных, анализ и классификация сетевого трафика.

Aleksandr Igorevich GETMAN – Cand. Sci. (Phys.-Math.), senior researcher at ISP RAS, assistant at CMC MSU, associate professor at HSE and MIPT. Research interests: binary code analysis, data format recovery, network traffic analysis and classification.