# РАСПРЕДЕЛЕННАЯ СИСТЕМА ОБНАРУЖЕНИЯ СЕТЕВЫХ АТАК НА ОСНОВЕ ФЕДЕРАТИВНОГО ТРАНСФЕРНОГО ОБУЧЕНИЯ

*Васильев В. И.[1], Вульфин А. М.[2], Картак В. М.[3], Башмаков Н. М.[4], Кириллова А. Д.[5]*

**DOI:** 10.21681/2311-3456-2024-6-117-129

***Цель исследования:*** *повышение эффективности обнаружения сетевых атак ботнетов за счет применения федеративного трансферного обучения, что позволит аккумулировать в рамках гибридной нейросетевой модели знания о сетевых атаках на различные клиентские корпоративные информационные инфраструктуры, обеспечивая конфиденциальность клиентского сетевого трафика.*

***Метод исследования:*** *для оперативной обработки и анализа сетевого трафика использованы методы машинного обучения. Применены методы построения моделей вложений и автоэнкодеров для извлечения признаков, методы построения бинарных классификаторов на основе глубоких нейронных сетей, включая сверточные нейронные сети и полносвязные сети прямого распространения. Использованы методы федеративного трансферного обучения.*

***Полученные результаты:*** *разработан прототип интеллектуальной системы обнаружения сетевых атак и вторжений на основе федеративного трансферного обучения. Предложена архитектура системы в составе центра мониторинга информационной безопасности, приведена структурная схема серверной и клиентской компонент системы, позволяющих решать задачи сбора и предобработки данных сетевых сессий и управлять жизненным циклом моделей анализа. Приводятся результаты сравнительной оценки эффективности обнаружения специализированных сетевых атак на примере управляющего трафика ботнетов. Сравниваются бинарные классификаторы на основе полносвязных глубоких нейронных сетей прямого распространения, сверточных нейронных сетей с одномерным входным слоем, ансамблевых моделей на основе деревьев решений, гибридных автоэнкодеров со слоем вложений и сверточным классификатором в сценариях централизованного и федеративного обучения. Гибридная нейросетевая модель в режиме федеративного обучения демонстрирует наилучшие показатели (F1-мера = 0,91) благодаря эффективной схеме представления признаков, но время ее обучения существенно возрастает (в 1,5–2 раза).*

***Научная новизна:*** *предложена гибридная нейросетевая модель классификации сетевых сессий, основанная на нейросетевых моделях вложений и моделях нейросетевых сверточных автоэнкодеров, отличающаяся алгоритмом кодирования разреженных категориальных и непрерывных признаков без использования размеченной обучающей выборки и применением федеративного трансферного обучения, что позволит обеспечить конфиденциальности данных локальных клиентов и возможность переноса обучения, а также повысить оперативность и достоверность обнаружения вредоносного сетевого трафика специалистами центров мониторинга информационной безопасности.*

***Вклад авторов:*** *Васильев В. И. – планирование исследований в области построения систем обнаружения атак с применением методов машинного обучения, проведение сравнительного анализа результатов моделирования, подготовка аналитического обзора. Вульфин А. М. – проведение экспериментального исследования на основе разработанного программного обеспечения. Картак В. М. – подготовка аналитического обзора, планирование эксперимента, проектирование программного обеспечения. Башмаков Н. М. – подготовка данных для моделирования, интерпретация результатов исследования; обобщение результатов исследования; формулировка выводов. Кириллова А. Д. – разработка программного обеспечения, оформление рукописи статьи; работа с графическим материалом.*

***Ключевые слова:*** *глубокое обучение, трафик управления ботнетами, сверточные нейросетевые классификаторы, автоэнкодеры, нейросетевые модели вложений.*

1 Васильев Владимир Иванович, доктор технических наук, профессор, Уфимский университет науки и технологий, г. Уфа, Россия. E-mail: vas0015@yandex.ru
2 Вульфин Алексей Михайлович, доктор технических наук, профессор, Уфимский университет науки и технологий, г. Уфа, Омский государственный технический университет, г. Омск, Россия. E-mail: vulfin.am@ugatu.su
3 Картак Вадим Михайлович, доктор физико-математических наук, профессор, Уфимский университет науки и технологий, г. Уфа, Россия. E-mail: kartak.vm@ugatu.su
4 Башмаков Наиль Маратович, аспирант, Уфимский университет науки и технологий, г. Уфа, Россия. E-mail: nail.bashmakov@gmail.com
5 Кириллова Анастасия Дмитриевна, кандидат технических наук, старший преподаватель, Уфимский университет науки и технологий, г. Уфа, Россия. E-mail kirillova.andm@gmail.com

## Введение

Одной из ключевых проблем использования методов искусственного интеллекта и машинного обучения (МО) при построении систем обнаружения атак (СОА) и обнаружения вторжений (СОВ) для распределенных автоматизированных объектов и систем является проблема формирования качественных наборов обучающих данных [1]. Данные могут иметь такие осложняющие особенности, как неполнота (т.е. отсутствие конкретных данных определенного вида), дисбаланс (неравномерность распределения данных для различных классов), недостаток размеченных данных. Попытки собрать и объединить разрозненные (локальные) обучающие данные, принадлежащие разным организациям, в единую, централизованную обучающую выборку, которая будет храниться на центральном сервере (центре обработки данных), входят в противоречие с нормативными требованиями обеспечения конфиденциальности этих данных.

Одним из эффективных путей решения этой проблемы является идея федеративного обучения, впервые предложенная в работах [2]. Федеративное обучение (ФО) – это новое направление в МО, когда несколько участников (клиентов) совместно обучают свои локальные модели МО под управлением центрального сервера, при этом не сообщая ему свои обучающие данные, т.е. сохраняя их конфиденциальность. Участники только информируют центральный сервер о своих промежуточных результатах обучения (настройках моделей), а центральный сервер, в свою очередь, обрабатывает эту информацию с целью обновления собственной (глобальной) модели МО и предоставляет эти обновления всем участникам для изменения настроек их моделей. В литературе представлен ряд подробных обзоров, посвященных рассмотрению методов ФО [3, 4] и применению этих методов в задачах построения СОА [5,6].

Отличительной особенностью многих работ, посвященных построению СОА на основе ФО, является использование открытых наборов обучающих данных (датасетов), таких как NSL-KDD, CICIDS 2017 и 2018, UNSW-NB15, N-BaIoT, BoT-IT, Edge-IoT dataset и др. В зависимости от признаков проявления различных классов атак с учетом специфики конкретной предметной области и способа их использования в процессе обучения различают следующие группы методов ФО: горизонтальное ФО, вертикальное ФО и федеративное трансферное обучение. Наибольший интерес из перечисленных методов представляют методы федеративного трансферного обучения (ФТО), предложенного в 2018 г. в работе [7], в основе которого заложено объединение двух подходов: федеративного обучения (ФО) и трансферного обучения (ТО). Основная идея ТО – перенос знаний (Knowledge Transfer) с некоторой предварительно обученной модели МО на другие модели (задачи), что позволяет дообучать (fine-tuning) другие проблемно-ориентированные модели на малом наборе данных, одновременно повышая точность их обучения [8]. Использование ТО позволяет в данном случае более полно воспользоваться располагаемыми гетерогенными данными и знаниями, имеющимися в арсенале участников, для восполнения возможного недостатка данных или меток.

Настоящая статья построена следующим образом. В первой главе приведены основные положения ФТО и анализ релевантных работ по рассматриваемой тематике. Во второй главе представлена архитектура предложенной СОА на основе ФТО, рассмотрены ее основные компоненты. Третья глава содержит изложение полученных результатов моделирования СОА и сравнительную оценку ее эффективности. В заключении приведены выводы по результатам исследований и направления будущих работ.

## 1. Федеративное трансферное обучение. Анализ релевантных работ

Рассмотрим формальную постановку задачи федеративного обучения (ФО) [4]. Пусть имеется $N$ клиентов (узлов, участников ФО), каждый из которых обладает собственным набором обучающих данных (датасетов) $D_i$. Каждый из этих наборов включает в себя определенное число объектов (образцов, samples) $I_i$, каждый из которых, в свою очередь, характеризуется некоторым множеством признаков (features) $\{X_i\}$ и метками (labels) $\{Y_i\}$, обозначающими принадлежность объекта $I_i$ тому или иному классу, т.е. $Di = <I_i, X_i, Y_i>$. Обозначим через $I = \{I_i\}$ множество объектов, $X = \{X_i\}$ – множество признаков, $Y = \{Y_i\}$ – множество меток объектов, а через $\mathrm{D} = \{D_i\}$ набор обучающих данных.

Применительно к задаче построения распределенной СОА под множеством объектов обычно понимается множество сетевых потоков (сегментов сетевого трафика, подлежащего распознаванию и классификации); множество признаков $X$ включает такие признаки, как длительность сетевого соединения, число передаваемых байтов или пакетов, используемый протокол, целевую службу и т.п.; множество $Y$ – это множество меток типа «Норма» или с указанием конкретного типа атак (в случае аномального сетевого трафика).

При использовании традиционного машинного обучения все наборы обучающих данных объединяются в полный набор $D = D_1 \cup \ldots \cup D_N$, на котором обучается модель $MD$ с некоторой точностью $A(MD)$. В случае ФО не происходит объединения наборов данных $D_i$, а глобальная модель $M_{FL}$ обучается

на основе локально обученных моделей $M_{D_i}$, использующих локальные наборы $D_i$, причем точность полученной глобальной модели $A(M_{ML})$ должна удовлетворять следующему требованию:

$$| A(M_D) - A(M_{ML}) | \le \delta, \quad (1)$$

где $\delta$ – малая положительная величина.

В зависимости от того, как обучающие данные распределяются между $N$ клиентами, участвующими в ФО, различают следующие категории ФО:

а) горизонтальное ФО (Horizontal Federated Learning, HFL) – наборы обучающих данных используют одно и то же множество признаков, но разные множества объектов:

$$X_i = X_j, \quad Y_i = Y_j, \quad I_i \ne I_j, \quad \forall D_i, D_j, \quad i \ne j; \quad (2)$$

б) вертикальное ФО (Vertical Federated Learning, VFL) – наборы обучающих данных используют одно и то же множество объектов, но различные множества признаков:

$$X_i \ne X_j, \quad Y_i \ne Y_j, \quad I_i = I_j, \quad \forall D_i, D_j, \quad i \ne j; \quad (3)$$

в) федеративное трансферное обучение (Federated Transfer Learning, FTL) – наборы данных отличаются как по объектам, так и по множеству признаков:

$$X_i \ne X_j, \quad Y_i \ne Y_j, \quad I_i \ne I_j, \quad \forall D_i, D_j, \quad i \ne j. \quad (4)$$

Особенность ФТО заключается в сочетании федеративного обучения, гарантирующего конфиденциальность обучающих данных клиентов, с трансферным обучением, с помощью которого осуществляется перенос знаний от одних клиентов, располагающих более богатой информацией, другим клиентам, имеющим недостаточно признаков или меток. Данная ситуация иллюстрируется рис. 1, где обучающие данные 2-х клиентов (A и B) перекрываются лишь в малой зоне как по множеству объектов (samples), так и по множеству признаков (features), но с помощью трансферного обучения происходит передача части признаков и меток от клиента  B клиенту A. Таким образом, обучаемая модель распространяется на непересекающуюся область данных в A (т. е. на рис. 1 фактически заполняется правый верхний угол).

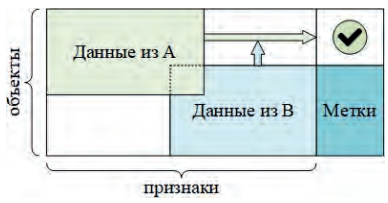


*Рис. 1. Схема федеративного трансферного обучения [4]*

Рассмотрению различных подходов к реализации ФТО посвящены обзоры [9]. На рис. 2 представлена типовая архитектура системы ФТО с $N$ клиентами.


*Рис. 2. Архитектура системы ФТО*

Согласно рис. 2, процесс ФТО содержит следующие основные этапы:

1. **Инициализация** – на центральном сервере создается глобальная модель МО, которая предварительно обучается на открытом датасете, содержащем большое количество обучающих данных. Эта модель затем распространяется в качестве базовой модели для построения локальных моделей всех $N$ клиентов.
2. **Локальное обучение** – каждый клиент дообучает полученную им предварительно обученную модель (fine-tuning) на своих собственных обучающих данных (локальном датасете) для решения своей конкретной задачи.
3. **Передача параметров (настроек) серверу** – результаты обучения, т.е. параметры настроек локальных моделей (веса, градиенты изменения весов) пересылаются в зашифрованном виде центральному серверу.
4. **Агрегирование** – центральный сервер усредняет полученные значения параметров локальных моделей и вносит соответствующие изменения в свою глобальную модель, обновляя ее таким образом с учетом полученной информации.
5. **Передача обновлений клиентам** – улучшенная глобальная модель возвращается клиентам, которые, в свою очередь, дообучают ее на своих обучающих данных (далее шаги 2-5 повторяются до тех пор, пока не будут достигнуты заданные показатели точности модели ФТО (см. условие (1)).

Использование ФТО при этом обеспечивает такие преимущества, как:

– сохранение конфиденциальности данных клиентов (поскольку обучающие данные клиентов остаются у них, они не передаются на центральный сервер, а передаются только параметры локальных моделей);

– эффективное использование данных (нет необходимости использования клиентами больших датасетов, публичный датасет с большим набором данных используется только на этапе предварительного обучения (pre-training) глобальной модели);
– трансфер знаний (фактически имеет место для совместного обучения локальных моделей клиентов, с переносом знаний от одних клиентов к другим);
– снижение нагрузки на сервер (распределяя вычисления по отдельным клиентам, ФТО позволяет уменьшить вычислительную нагрузку на центральный сервер).

На сегодня предложено значительное количество алгоритмов агрегирования параметров локальных моделей (настроек ФТО) [6]. Наиболее популярным из них является алгоритм федеративного усреднения FedAvg (Federated Averaging), согласно которому обновленные значения векторов параметров вычисляются по формуле

$$W_{t+1}^k = \sum_{k=1}^{N} \frac{n_k}{N} W_t^k, \quad (5)$$

где $W_t^k$ и $W_{t+1}^k$ – векторы настраиваемых параметров (весов, градиентов) локальной модели $k$-го клиента соответственно в моменты времени $t$ и $(t + 1)$; $t$ – дискретное время (итерация обучения); $n_k$ – размерность используемого $k$-м клиентом набора обучающих данных; $N$ – общее количество клиентов. Другие известные алгоритмы агрегирования параметров локальных моделей МО – FedSGD, FedProx, Fed+, FedCM, DWFed, FedMA [6].

Для разработки и реализации прикладных систем ФТО обычно используются специализированные проблемно-ориентированные фреймворки (frameworks). я кодом на основе языка Python, такие как Tensor Flow Federated (TFF), FATE, PFL, PySyft, FL&DP [10], а также Java-ориентированные программные продукты, например, Federated Learning for Java (FL4J) [11].

Исследованию особенностей применения ФТО для построения распределенных СОА в последние годы посвящено достаточно много публикаций. Значительное внимание, в частности, уделяется вопросам обеспечения защищенности IoT. Так, в [12] представлены результаты разработки фреймворка для построения СОА на основе принципов ФТО на примере 3-х узлов (клиентов) IoT. Произведена оценка эффективности СОА с использованием специально собранного экспериментального стенда и датасета CSE-CICIDS 2018 (с разделением этого датасета на непересекающиеся подмножества, имеющие различные признаки и метки, для отдельных клиентов). В качестве вариантов построения глобальной и локальных моделей рассмотрены полносвязная глубокая нейронная сеть (НС) и сверточная НС.

В [13] аналогичные базовые модели МО (полносвязная НС и сверточная МО) использовались при построении СОА для медицинской IoT-системы (Internet of Medical Things) с 3-мя клиентами. При обучении этих моделей с помощью ФТО предполагалось использование сервером и клиентами 4-х различных подмножеств датасета Edge-IIoT set.

В [14] представлена разработка федеративной системы IoT Defender, представляющей собой фреймворк для построения распределенной СОА на основе ФТО применительно к IoT с использованием телекоммуникаций нового поколения 5G. В качестве базовой модели (сервер и 4 клиента) рассматривается сверточная НС. Всего были использованы 5 различных датасетов: 2 публичных датасета CICIDS 2017 и NSL-KDD, а также 3 частных (специально подобранных) датасета для различных групп «умных» устройств IoT. Алгоритм агрегирования – FedAvg.

В [15] задача построения СОА на основе ФТО решалась с использованием в качестве базовой модели гибридной НС (полносвязная НС + сверточная НС). Исходный датасет – Edge-IIoT set, с разделением по различным клиентам; алгоритм агрегирования – FedSGD (Federated Stochastic Gradient Descent).

Работа [16] посвящена разработке СОА на основе ФТО для IIoT с использованием перспективных беспроводных сетей 6-го поколения (6G) – Mobile Edge Computing. На стороне сервера (глобальная модель) и клиентской стороне (локальные модели, 10 клиентов) используется сверточная НС; 2 датасета NSL-KDD и UNSW-NB15 разбиты на отдельные непересекающиеся подмножества; алгоритм агрегирования – FedSGD.

Другая группа работ, в отличие от перечисленных, использует в качестве базовых достаточно редкие классы моделей МО, пока не столь характерные для рассматриваемой предметной области. Так, в [17] при построении СОА на основе ФТО базовая модель (на серверной и клиентской стороне) выбрана в классе НС специального вида – машин экстремального обучения (Extreme Learning Machine, ELM). При обучении моделей использовались датасеты NSL-KDD, KDD99, ISCX 2012; алгоритм агрегирования – FedAvg.

В [18] глобальная и локальные модели МО строятся в классе генеративно-состязательных сетей (Generative Adversarial Networks, GAN). Клиенты представляют собой информационные системы промышленных предприятий; датасеты сформированы на основе собранных клиентами реальных данных об атаках; алгоритм агрегирования – FedAvg.

В [19] для защиты IoT от атак ботнетов предложена СОА, в которой в качестве базовой, предварительно обучаемой модели МО используется нейросетевая модель трансформера. Исходный датасет – N-BaIoT,

собран клиентами с 9 коммерческих IoT-устройств, атакуемых ботнетами Mirai и BASHLITE. Алгоритм агрегирования – FedAvg.

Все рассмотренные СОА, построенные с применением ФТО, показали высокую точность обнаружения атак по сравнению с традиционными децентрализованными СОА, в том числе при обнаружении неизвестных ранее для клиентов атак (немаркированных, отсутствующих в локальных датасетах), при сохранении конфиденциальности обучающих данных клиентов, что является главным преимуществом ФТО.

## 2. Архитектура системы обнаружения атак на основе федеративного трансферного обучения

Современным этапом развития комплексного подхода к обеспечению безопасности информационной инфраструктуры является создание центров мониторинга информационной безопасности (ЦМИБ, Security Operation Center, SOC). Организационно-технические процедуры ЦМИБ направлены на обнаружение и предотвращение киберугроз с учетом ключевых принципов проактивной защиты как сочетания тактической и стратегической аналитики на основе инженерии знаний и интеллектуальной обработки гетерогенных слабоструктурированных данных, получаемых из внутренних и внешних источников.

**Архитектура СОА/СОВ.** Гибридная архитектура распределенной СОА/СОВ, предназначенной для использования в составе ЦМИБ, представлена на рис. 3. Здесь выделены клиентские компоненты СОА/СОВ, роль которых заключается в оперативном мониторинге сетевого трафика в пределах клиентской инфраструктуры, а также серверная компонента, предназначенная для агрегации накапливаемых клиентскими компонентами знаний о реализации сетевых атак и их ключевых признаках и их интеграции с внешними базами знаний. Подобное разделение позволяет снизить как объемы передаваемых в ЦМИБ для анализа данных (минимизируя передачу чувствительных данных), так и необходимые вычислительные ресурсы.

Рассмотрим более подробно структурную организацию предлагаемой распределенной СОА/СОВ (рис. 4).

В состав клиентской компоненты системы входят следующие подсистемы:

- подсистема (IK) сбора и предобработки данных сетевых сессий – выполняет непрерывный мониторинг сетевой активности в пределах клиентской информационной инфраструктуры, размещая данные сетевых сессий в формате Netflow в хранилище на основе локальной колоночной базы данных;
- подсистема (IIK) – обеспечивает взаимодействие с локальными компонентами SIEM ЦМИБ для создания контекста сетевого взаимодействия и разметки сетевой активности на основе совокупности событий и инцидентов ИБ;
- подсистема (IIIK) подготовки данных для обучения локальных моделей анализа сетевого трафика – формирует обучающий набор данных для контролируемого обучения локальной модели в процессе ФТО;

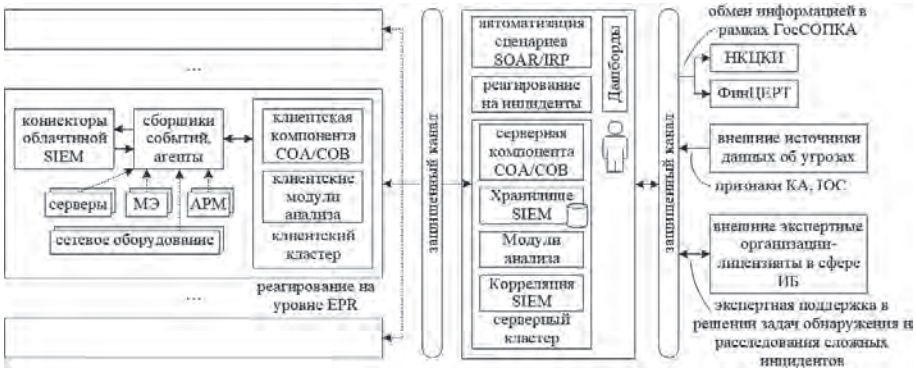


*Рис. 3. Гибридная архитектура СОА/СОВ в составе ЦМИБ (МЭ – межсетевой экран, КА – компьютерные атака, IOC – индикаторы компрометации)*


*Рис. 4. Структурная схема клиентской и серверной компонент СОА/СОВ*

– подсистема (IVK) предназначена для управления жизненным циклом цепочки моделей в процессе ФТО: «глобальная модель»-«локальная модель»-«обновления для глобальной модели».

В состав подсистемы IVK входят следующие модули:

– модуль (1K) подготовки локальной модели анализа;
– модуль (2K) оперативного анализа клиентского сетевого окружения;
– модуль (3K) управления жизненным циклом локальной модели анализа;
– модуль (4K) для обновления локальной модели на основе серверной глобальной модели, а также пересылки на сервер обновлений для глобальной модели в рамках процесса ФТО.

Серверная компонента распределенной СОА/СОВ включает в себя следующие подсистемы:

– подсистема (IC) и подсистема (VC) обеспечивают сбор, предобработку и систематизацию данных сетевого взаимодействия, полученных из открытых источников (наборы данных PCAP с разметкой) и доступных клиентских сессий в серверном хранилище;
– подсистема (IIIC) обеспечивает верификацию и валидацию собираемых данных с целью обнаружения возмож ных атак на модель ФТО;
– подсистема (IIC) в ходе взаимодействия с базой знаний ЦМИБ позволяет создавать и обогащать контекст собираемых сетевых сессий;
– подсистема (IVC) реализует серверную часть процесса ФТО.

В состав подсистемы IVC серверной компоненты входят: модуль (1C) подготовки и обновления глобальной модели, модуль двустороннего взаимодействия с локальными моделями (2C) и модуль (3C) для управления жизненным циклом моделей ФТО.

**Гибридная нейросетевая модель классификации сетевых сессий на основе ФТО.** Основные трудности классификации сетевых сессий заключаются в следующем:

– существенный дисбаланс количества доступных сетевых сессий обычного взаимодействия конечных систем и подтвержденной вредоносной сетевой активности [20];
– неидентичность и зависимость в распределении данных (nonIDD) [1, 21], собираемых с различных клиентских инфраструктур;
– разнообразие способов выделения и кодирования ключевых признаков;
– необходимость учитывать как «дрейф данных», так и «дрейф концепции» для оценки горизонта пригодности обучаемых моделей [22];
– проверка эффективности работы системы при использовании нескольких наборов данных с разными способами реализации атак одного класса;
– как правило, основные результаты в известных работах получены с использованием различных типов широко распространенных сетевых атак, вопросы обнаружения узкоспециализированных атак анализируются очень редко.

Исходя из вышеперечисленного, предлагается использовать гибридную нейросетевую модель классификации сетевых сессий (рис. 5).

Блок (1) в составе модели обобщает предложенный в [23–25] способ представления разреженных категориальных и непрерывных численных переменных в виде компактного векторного представления. Отличительной особенностью является возможность эффективного кодирования признаков разного типа

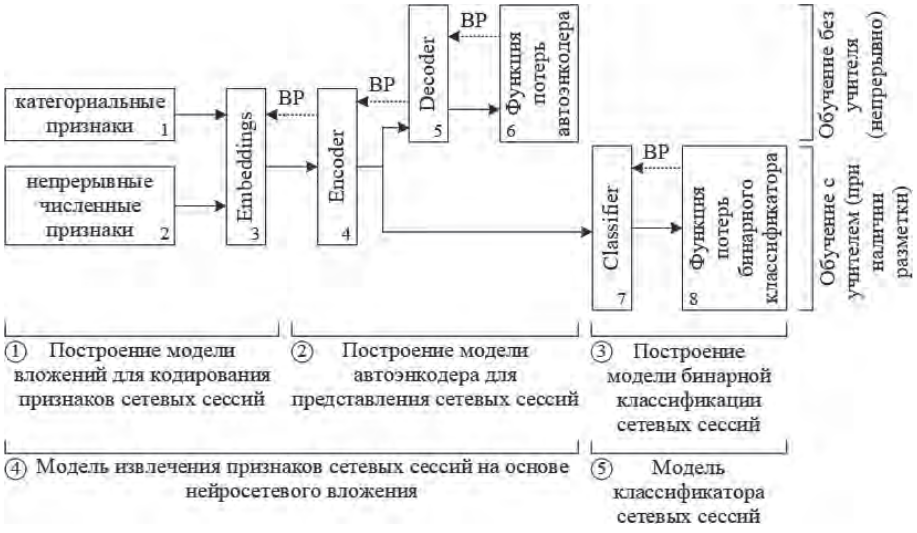


*Рис. 5. Гибридная нейросетевая модель классификации сетевых сессий*

(особенно важно при использовании глубоких нейросетевых моделей) с помощью модели вложений, настройка параметров которой осуществляется с помощью автоэнкодера (2), для обучения которого не требуется размеченной выборки – схема построения «автоэмбеддера» [26]. Обучение слоев вложений (Embedder) и симметричных слоев автоэнкодера (Encoder-Decoder) позволяет использовать все доступные сетевые сессии – нормальной работы, различных типов атак и т.д., игнорируя дисбаланс распределения примеров по классам. Блок (3) представляет собой классификатор, использующий в качестве вектора признаков подготовленные входными слоями высокоуровневые признаки. «Заморозка» параметров входных слоев позволяет дообучить слои классификатора на имеющихся, ограниченных по количеству, размеченных сетевых сессиях. Таким образом, гибридная нейросетевая модель классификации сетевых сессий включает две составляющие: модель извлечения признаков сетевых сессий (4) и модель классификатора (5).

**Федеративное трансферное обучение гибридной нейросетевой модели классификации сетевых сессий.** Рассмотрим схему обучения предлагаемой гибридной нейросетевой модели (рис. 6) в схеме ФТО:

Шаг 0. На основе доступных данных с частичной разметкой на сервере создаются и обучаются последовательно два блока модели: автоэнкодер (Embedder и Encoder-Decoder – используются все доступные верифицированные данные без разметки) и классификатор (на ограниченном размеченном наборе обучается блок Classifier).

Шаг 1. Обученная глобальная модель передается по защищенным каналам на клиентские компоненты.

Шаг 2. Клиентские модели на локальных данных поэтапно продолжают обучение блока автоэнкодера (повышая эффективность извлечения признаков) и, при наличии размеченных данных, – блока классификатора. По истечении заданного количества итераций локального обучения выполняется передача оценок градиентов на сервер.

Шаг 3. Полученные оценки градиентов агрегируются на сервере и используются для трансферного обучения блоков глобальной модели. Весовые коэффициенты глобальной модели передаются клиентским моделям.

Шаг 4. Процедура продолжается либо до достижения сходимости оценок (FedAVG), либо по достижении заданных критериев (FedAVG+), обеспечивая устойчивость к неидентичности и зависимости в распределении данных (non-IID) на клиентских подсистемах.

### 3. Оценка эффективности СОА

Для оценки эффективности работы прототипа распределенной СОА/СОВ на основе ФТО обратимся к задаче обнаружения сетевого трафика командных центров ботнетов на ранних стадиях проникновения в корпоративные информационные инфраструктуры. Сетевой трафик инфраструктур управления и контроля ботнетов (Command & Control, C&C) характеризует взаимодействие специализированных серверов злоумышленника со скомпрометированными устройствами и является узкоспециализированной сетевой атакой, обнаружение которой сопряжено с рядом трудностей [14].

Для серии экспериментов были выбраны наборы данных NF-UNSW-NB15 и NF-CSE-CIC-IDS2018, преобразованные к единому формату представления признаков NetFlow, а также специализированные наборы

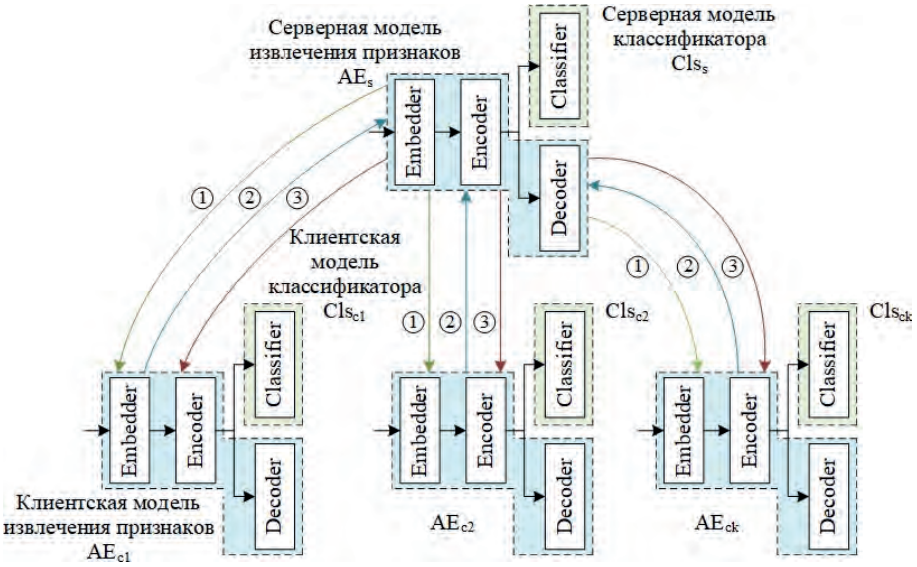


*Рис. 6. Схема федеративного трансферного обучения гибридной нейросетевой модели*

Таблица 1.

Характеристики наборов данных C&C

| Название набора данных | Источник | Количество признаков | Тип разметки | Количество сетевых сессий | Роль в схеме обучения моделей |
|---|---|---|---|---|---|
| NF-UNSW-NB15 | Университет Нового Южного Уэльса (UNSW), Австралия | 43 | Netflow | Benign – 1550712<br>Backdoor – 1782 | Тестовое множество для имитации новых реализаций атаки |
| BH-KSU23 | Университет имени Короля Сауда (KSU), Саудовская Аравия | 79 | CICFlowmeter | Benign – 257691<br>Malicious – 209539 | Обучение базовой серверной модели |
| NF-CSE-CIC-IDS2017 | Канадский институт кибербезопасности (CIC), Канада | 43 | Netflow | Benign – 7373198<br>Bot – 15683 | Имитация данных «Клиент 1» |
| Trojan Detection | Университет Дрексела (Drexel), США | 79 | CICFlowmeter | Benign – 86799<br>Trojan – 90683 | Имитация данных «Клиент 2» |
| MTA-KDD19 | Университет Л'Акуила (L'Aquila), Италия | 50 | Netflow + модификации | Benign – 31926<br>Malware – 39544 | Имитация данных «Клиент 3» |

данных BH-KSU23, Trojan Detection и MTA-KDD19, содержащие сессии нормальной работы и размеченный трафик для 15 схем реализации C&C взаимодействия (табл. 1).

Проекция исходного признакового пространства с помощью алгоритма UMAP (Uniform Manifold Approximation and Projection) для сетевых сессий нормальной работы и C&C трафика из подмножеств NF-UNSW-NB15 и NF-CSE-CIC-IDS2017 приведена на рис. 7.

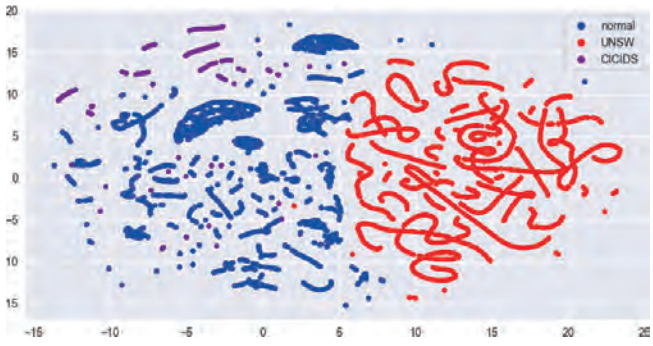


*Рис. 7. Проекция исходного признакового пространства для сетевых сессий нормальной работы и C&C трафика двух наборов*

Как видно, сетевые сессии, помеченные как «нормальный трафик», располагаются компактно, выраженными группами. Данные C&C из набора данных UNSW достаточно хорошо отделимы от обычного трафика. Данные C&C из набора CICIDS, напротив, отделимы хуже и перемешаны с примерами нормальных сетевых сессий.

Схема проведения серии экспериментов включает построение четырех типов моделей бинарных классификаторов (табл. 2).

Модели DNN, CNN1D построены в данном случае как с помощью ФТО (эксперименты 1, 3), так и в качестве одиночных глобальных моделей (эксперименты 2, 4), обучаемых на объединенном наборе данных. Модель XGBoost строится только как глобальная модель (эксперимент 7), а модель AE-CNN1D – только с помощью ФТО (эксперименты 5 и 6).

Из всех имеющихся наборов данных существенный дисбаланс по числу примеров в классах наблюдается только для NF-UNSW-NB15 и NF-CSE-CIC-IDS2017 – для них применяется схема случайного удаления примеров мажоритарного класса. Обучающая выборка после семплирования в каждом из наборов включает 75 % примеров, тестовая – 25 %. Кодировщики категориальных и непрерывных признаков для всех моделей, кроме AE-CNN1D: RS – Robust Scaler, OE – Ordinal Encoder, OHE – One hot encoder. С помощью фреймворка Optuna выполнена предварительная оптимизация гиперпараметров бинарных классификаторов на основе алгоритма TPE.

В качестве алгоритма передачи обновлений градиентов в ходе обучения локальных моделей

*Таблица 2.*

*Схема проведения серии экспериментов*

| Модель | Характеристика модели | Параметры модели | | Схема построения | Особенность набора данных | № эксперимента |
|---|---|---|---|---|---|---|
| DNN | Полносвязная глубокая нейронная сеть прямого распространения | Количество слоев | 6 | ФТО | Сбалансированный | 1 |
| | | Количество нейронов по слоям | 98, 128, 64, 32, 4, 1 | | | |
| | | Исключение (dropout) | 3, 4 слои | Глобальная | Сбалансированный | 2 |
| | | Функция активации | ReLU, последний – sigmoid | | | |
| | | Функция потерь | Binary Cross-Entropy With Logits | | | |
| | | Коэффициент скорости обучения | 0,085 | | | |
| | | Количество эпох обучения | 64 | | | |
| CNN1D | Сверточная нейронная сеть с входным слоем в виде одномерного кортежа | Слои: conv1d, dropout, conv1d, dropout, flatten, dropout, batch_normalization, fully_connected, fully_connected, fully_connected | | ФТО | Сбалансированный | 3 |
| | | Размер ядер свертки | 3 | | | |
| | | Количество фильтров | 5 | Глобальная | Сбалансированный | 4 |
| | | Параметр исключения (dropout) по слоям | 0,1, 0,1 | | | |
| | | Количество нейронов в полносвязных слоях | 128, 64, 1 | | | |
| | | Функция активации | ReLU, последний – sigmoid | | | |
| | | Функция потерь | Binary Cross-Entropy With Logits | | | |
| | | Количество эпох обучения | 32 | | | |
| AE-CNN1D | Гибридный автоэнкодер со слоем вложений в сочетании с классификатором на основе полносвязной сети с пакетной нормализацией | Слои Encoder/Decoder (симметрично): (embeddings1 + embeddings2), concat_embeddings, conv1d, dropout, conv1d, dropout, flatten | | ФТО | Сбалансированный | 5 |
| | | Слои Classifier: flatten, dropout, batch_normalization, fully_connected, fully_connected, fully_connected | | | | |
| | | Размер ядер свертки | 3 | | Не сбалансированный + все доступные сессии прочих атак – для автоэнкодера | 6 |
| | | Количество фильтров | 5 | | | |
| | | Параметр исключения (dropout) по слоям | 0,1, 0,1 | | | |
| | | Количество нейронов в полносвязных слоях | 128, 64, 1 | | | |
| | | Функция активации | ReLU, последний слой Classifier – sigmoid | | | |
| | | Количество эпох обучения автоэнкодера | 32 | | | |
| | | Количество эпох обучения классификатора | 16 | | | |
| XGBoost | Ансамблевый метод объединяет слабые модели на основе деревьев решений | Максимальная глубина дерева | 16 | Глобальная | Сбалансированный | 7 |
| | | Скорость обучения | 0,029 | | | |
| | | Количество слабых классификаторов в ансамбле | 316 | | | |
| | | Соотношение подвыборок обучающих экземпляров | 0,997 | | | |
| | | Коэффициенты регуляризации L1, L2 | 1,391, 2,840 | | | |

*Таблица 3.*

*Параметры сервера для запуска моделей*

| Параметр | Характеристика |
|---|---|
| GPU | 4 GPU Tesla V100 |
| Объем видеопамяти GPU | 128 ГБ |
| Процессор | Intel Xeon E5-2698 v4 2,2 ГГц (20-ядерный) |
| Объем оперативной памяти | 256 ГБ RDIMM DDR4 |

использована модификация алгоритма FedAvg+ в версии [8] фреймворка FATE. Размер пакета при обучении серверной и клиентских моделей составляет 64, агрегация локальных градиентов осуществлялась через каждые 2 эпохи обучения моделей.

Для обучения использовался высокопроизводительный сервер, параметры которого представлены в табл. 3. Каждая из моделей в изолированном окружении использовала выделенные GPU, обмен данными между моделями осуществлялся через каналы в RAM.

При оценке качества бинарной классификации были использованы следующие метрики:

– Precision (Точность) – доля правильно предсказанных положительных случаев среди всех предсказанных положительных случаев;
– Recall (Полнота) – доля правильно предсказанных положительных случаев среди всех реальных положительных случаев;
– F1-мера – является гармоническим средним точности и полноты.

Результаты оценки качества моделей приведены в табл. 4.

Схема с централизованным обучением моделей на всех имеющихся данных (2, 4, 7) продемонстрировала ожидаемые высокие показатели F1-меры. Причем, сверточная модель по показателям F1-меры несущественно опережает классическую полносвязную DNN; модель XGBoost в задаче опережает на отдельных тестовых наборах данных нейросетевые модели, но лишена преимуществ обучения распределенных моделей.

В целом, применение схемы федеративного обучения оказалось весьма успешным: модель способна классифицировать трафик исходного набора данных (BH-KSU23), клиентских наборов данных и «новых» сетевых сессий (NF-UNSW-NB15) командного трафика ботнетов.

С точки зрения повышения эффективности классификации оказалось целесообразным использовать гибридную нейросетевую модель AE-CNN1D – модель демонстрирует наилучшие показатели благодаря эффективной схеме представления признаков, но время

*Таблица 4.*

*Основные результаты серии экспериментов*

| Набора данных | Метрики | ФТО | | | | Глобальная модель | | |
|---|---|---|---|---|---|---|---|---|
| | | DNN | CNN1D | AE-CNN1D | | DNN | CNN1D | XGB |
| | Эксперимент | 1 | 3 | 5 | 6 | 2 | 4 | 7 |
| | Время обучения, мин | 98 | 173 | 247 | 362 | 24 | 51 | 16 |
| BH-KSU23 (тестовая) | Precision | 0,9605 | 0,9674 | 0,9708 | 0,9963 | 0,9612 | 0,9771 | 0,972 |
| | Recall | 0,9689 | 0,9721 | 0,972 | 0,9936 | 0,9703 | 0,9859 | 0,9723 |
| | F1-мера | 0,9647 | 0,9698 | 0,9714 | 0,995 | 0,9654 | 0,9815 | 0,9722 |
| NF-CSE-CIC-IDS2017 (тестовая) | Precision | 0,9737 | 0,9755 | 0,9811 | 0,9962 | 0,9755 | 0,9771 | 0,9788 |
| | Recall | 0,9629 | 0,9723 | 0,9867 | 0,9934 | 0,9652 | 0,9778 | 0,9773 |
| | F1-мера | 0,9683 | 0,9739 | 0,9839 | 0,9948 | 0,9703 | 0,9774 | 0,9781 |
| Trojan Detection (тестовая) | Precision | 0,8558 | 0,8623 | 0,892 | 0,9086 | 0,8569 | 0,8673 | 0,8854 |
| | Recall | 0,8512 | 0,86 | 0,9067 | 0,9175 | 0,8534 | 0,863 | 0,9103 |
| | F1-мера | 0,8535 | 0,8612 | 0,8993 | 0,913 | 0,8551 | 0,8652 | 0,8977 |
| MTA-KDD19 (тестовая) | Precision | 0,987 | 0,9883 | 0,9941 | 0,9956 | 0,9876 | 0,9894 | 0,9921 |
| | Recall | 0,9893 | 0,994 | 0,9969 | 0,9986 | 0,99 | 0,9953 | 0,9961 |
| | F1-мера | 0,9881 | 0,9911 | 0,9955 | 0,9971 | 0,9888 | 0,9923 | 0,9941 |
| NF-UNSW-NB15 (весь набора) | Precision | 0,8547 | 0,8872 | 0,9416 | 0,9888 | 0,8636 | 0,9108 | 0,8872 |
| | Recall | 0,7484 | 0,7889 | 0,9149 | 0,988 | 0,7589 | 0,8059 | 0,8907 |
| | F1-мера | 0,798 | 0,8352 | 0,9281 | 0,9884 | 0,8079 | 0,8551 | 0,8889 |

ее обучения существенно возрастает. Однако можно прогнозировать стабильную работу модели при увеличении количества клиентских моделей и объемов доступных сетевых сессий, т.к. основным затруднением является именно обработка неидентичных (non-IID) данных. Построение эффективных векторов вложений позволит избежать дальнейшего повышения сложности классификатора.

## Заключение

На основании проведенного анализа источников литературы для повышения эффективности систем обнаружения сетевых атак и вторжений в корпоративных информационных системах предлагается использовать модели и алгоритмы федеративного трансферного обучения.

Разработан прототип интеллектуальной системы обнаружения сетевых атак и вторжений. Предложена архитектура СОА/СОВ в составе ЦМИБ, приведена структурная схема серверной и клиентской компонент системы, позволяющих решать задачи сбора и предобработки данных сетевых сессий, обеспечивать взаимодействие с ЦМИБ и управлять жизненным циклом моделей.

Предложена гибридная нейросетевая модель классификации сетевых сессий, включающая автоэнкодер, блоки построения вложений и классификатор. Отличительной особенностью является возможность эффективного кодирования разреженных категориальных и непрерывных признаков без использования размеченной обучающей выборки.

Результаты проведенных вычислительных экспериментов позволяют сделать выводы о высокой эффективности обнаружения специализированных сетевых атак на примере C&C трафика с помощью предложенного прототипа СОА/СОВ. Применение федеративного трансферного обучения обеспечивает при этом как сохранение конфиденциальности данных локальных клиентов, так и возможность переноса обучения – аккумулирования знаний о проводимых атаках на различные информационные инфраструктуры в рамках единой гибридной нейросетевой модели, что позволяет повысить оперативность и достоверность обнаружения вредоносного сетевого трафика, и тем самым, повысить защищенность клиентских корпоративных информационных систем.

*Благодарности.*
*Работа выполнена в ОмГТУ в рамках государственного задания Минобрнауки России на 2023–2025 годы № FSGF-2023-0004.*

***Литература***

1. *Wagle S. et al. Embedding alignment for unsupervised federated learning via smart data exchange // GLOBECOM 2022-2022 IEEE Global Communications Conference. IEEE, 2022, pp. 492–497.*
2. *McMaham H. B. et al. Communication-Efficient Learning of Deep Networks from Decentralized Data // arXiv preprint arXiv: 1602.05629 [cs.LG]. 2023. DOI: 10.48550/arXiv.1602.05629.*
3. *Wen J. et al. A Survey on Federated Learning: challenges and applications // International Journal of Machine Learning and Cybernetics. 2023, vol. 14, pp. 513–535.*
4. *Yang Q. et al. Federated Machine Learning: concept and applications // ACM Transactions on Intelligent Systems and Technology (TIST). 2019, vol. 10, no. 2, pp. 1–19. DOI: 10.1145/3298981.*
5. *Новикова Е. С. и др. Обнаружение вторжений на основе федеративного обучения: архитектура системы и эксперименты // Вопросы кибербезопасности. 2023, №6 (58). С. 50–66. DOI: 10.21681/2311-3456-2023-6-50-66.*
6. *Новикова Е. С. и др. Аналитический обзор подходов к обнаружению вторжений, основанных на федеративном обучении: преимущества использования и открытые задачи // Информатика и автоматизация. 2023. Т. 22, № 5. С. 1034–1082. DOI: 10.15622/ia.22.5.4.*
7. *Hernandez-Ramos J. L. et al. Intrusion Detection based on Federated Learning: a systematic review // arXiv preprint arXiv:2308.09522. 2023. DOI: 10.48550/arXiv.2308.09522.*
8. *Liu Y. et al. A secure federated transfer learning framework // IEEE Intelligent Systems. 2020 vol. 35, no. 4, pp. 70–82. DOI: 10.1109/MIS.2020.2988525.*
9. *Guo W. et al. A Comprehensive Survey of Federated Thansfer Learning: Challenges, Methods and Applications // arXiv preprint arXiv:2403.01387. 2024. DOI: 10.48550/arXiv.2403.01387.*
10. *Kholod I. et al. Open-Source Federated Learning Frameworks for IoT: A Comparative Review and Analysis // Sensors. 2020, no. 21, pp. 167. DOI: 10.3390/521010167.*
11. *Ефремов М. А., Холод И. И. Разработка архитектуры универсального фреймворка федеративного обучения // Программные продукты и системы. 2022. Т. 35, № 2. С. 263–273. DOI: 10.15827/0236-235X.138.263-272.*
12. *Otoum K., Yadlappali S. K., Nayk A. FTLIoT: A Federated Thansfer Learning Framework for Securing IoT // GLOBECOM 2022-2022 IEEE Global Communications Conference. IEEE, 2022, pp. 1146–1151. DOI: 10.1109/GLOBECOM48099.2022.10001461.*
13. *Otoum K., Chamola V., Nayak A. Federated and Transfer Learning – Empowered Intrusion Detection for IoT Applications // IEEE Internet of Things Magazine. 2022, vol. 5, no. 3, pp. 50–54. DOI: 10.1109/IOTM.001.2200048.*

14. Fan Y. et al. IoTDefender: A Federated Transfer Learning Intrusion Detection Framework for 5G IoT // 2020 IEEE 14th International Conference on Big Data Science and Engineering (BigDataSE). IEEE, 2020, pp. 88–95. DOI:10.1109/BigDataSE50710.2020.00020.
15. Rajesh L. T. et al. Give and Take: Federated Transfer Learning for Industrial IoT Network Intrusion Detection // 2023 IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom). IEEE, 2023, pp. 2365–2371. DOI: 10.1109/TrustCom60117.2023.00333.
16. Cheng Y. et al. Federated Transfer Learning with Client Selection for Intrusion Detection in Mobile Edge Computing // IEEE Communications Letters. 2022, vol. 26, no. 3, pp. 552–556. DOI:10.1109/LCOMM.2022.3140273.
17. Wang K., Li J., Wu W. An efficient intrusion detection method based on federated transfer learning and an extreme learning machine with privacy preservation // Security and Communication Networks. 2022, vol. 2022, no. 1, pp. 2913293. DOI:10.1155/2022/291329.
18. Guo W. et al. Federated transfer learning for auxiliary classifier generative adversarial networks: framework and industrial application // Journal of intelligent manufacturing. 2024, vol. 35, no. 4, pp. 1439–1454.
19. Metwaly A. A., Elhenawy I. Protecting IoT Devices from BotNet threats: a federated machine learning solution // Sustainable Machine Intelligence Journal. 2023, vol. 2, pp. 1–12. DOI:10.61185/SMIJ.2023.22105.
20. Azizjon M., Jumabek A., Kim W. 1D CNN based network intrusion detection with normalization on imbalanced data // 2020 international conference on artificial intelligence in information and communication (ICAIIC). IEEE, 2020, pp. 218-224. DOI:10.1109/ICAIIC48513.2020.9064976.
21. Новикова Е. С., Чен Я., Мелешко А. В. Методы оценки уровня разнородности данных в федеративном обучении // XXVII Международная конференция по мягким вычислениям и измерениям (SCM'2024) (Санкт-Петербург, 22-24 мая 2024). 2024. С. 446–450.
22. Yang Z. et al. A systematic literature review of methods and datasets for anomaly-based network intrusion detection // Computers & Security. 2022, vol. 116, pp. 102675. DOI:10.1016/j.cose.2022.102675.
23. Lee G. et al. Network Intrusion Detection with Improved Feature Representation // 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). IEEE, 2021, pp. 2049–2054.
24. He Y., Yan D., Chen F. Hierarchical federated learning with local model embedding // Engineering Applications of Artificial Intelligence. 2023, vol. 123, pp. 106148. DOI:10.1016/j.engappai.2023.106148.
25. Sivasubramanian A., Devisetty M., Bhavukam P. Feature Extraction and Anomaly Detection Using Different Autoencoders for Modeling Intrusion Detection Systems // Arabian Journal for Science and Engineering. 2024, pp. 1–13.
26. Wang Z. X. et al. Network traffic classification based on federated semi-supervised learning // Journal of Systems Architecture. 2024, vol. 149, pp. 103091. DOI: 10.1016/j.sysarc.2024.103091.

# DISTRIBUTED NETWORK ATTACK DETECTION SYSTEM BASED ON FEDERATE TRANSFER LEARNING

*Vasilyev V. I.[6], Vulfin A. M.[7], Kartak V. M.[8], Bashmakov N. M.[9], Kirillova A. D.[10]*

***Purpose:*** *Improving the efficiency of detecting botnet network attacks through the use of federated transfer learning. This makes it possible to accumulate knowledge about network attacks on various client corporate information infrastructures within the framework of a hybrid neural network model, ensuring the confidentiality of client network traffic.*

***Methods:*** *Machine learning methods were used for operational processing and analysis of network traffic. Methods for constructing embedding models and autoencoders for feature extraction, methods for constructing binary classifiers based on deep neural networks, including convolutional neural networks and fully connected feedforward networks, are applied. Federated transfer learning methods were used.*

***Research results:*** *A prototype of an intelligent system for detecting network attacks and intrusions based on federated transfer learning was developed. The architecture of the system as part of the information security monitoring center is proposed. The structural diagram of the server and client components of the system is given. The components allow solving the problems of collecting and preprocessing network session data and managing the life cycle of analysis models. The results of a comparative assessment of the effectiveness of detecting specialized network attacks are presented using the example of botnet control traffic. Binary classifiers based on fully connected deep feedforward neural networks, convolutional neural networks with a one-dimensional input layer, ensemble models based on decision trees, hybrid autoencoders with an embedding layer and a convolutional classifier are compared in centralized and federated learning scenarios. The hybrid neural network model in the federated learning mode demonstrates the best performance (F1-measure = 0.91) due to the effective feature representation scheme, but its training time increases significantly (by 1.5–2 times).*

***The scientific novelty:*** *A hybrid neural network model for classifying network sessions is proposed, based on neural network embedding models and neural network convolutional autoencoder models. The neural network model is distinguished by an algorithm for encoding sparse categorical and continuous features without using a labeled training sample and by the use of federated transfer learning. This ensures the confidentiality of local client data and the ability to transfer training, as well as increases the speed and reliability of detecting malicious network traffic by specialists at information security monitoring centers.*

6 Vladimir I. Vasilyev, Dr.Sc. (of Tech.), Professor, Ufa University of Science and Technology, Ufa, Russia. E-mail: vas0015@yandex.ru
7 Alexey M. Vulfin, Dr.Sc. (of Tech.), Professor, Ufa University of Science and Technology, Ufa, Omsk State Technical University, Omsk, Russia. E-mail: vulfin.am@ugatu.su
8 Vadim M. Kartak, Dr.Sc. (in Physics and Math.), Professor, Ufa University of Science and Technology, Ufa, Russia. E-mail: kartak.vm@ugatu.su
9 Nail M. Bashmakov, Post-Graduate Student, Ufa University of Science and Technology, Ufa, Russia. E-mail: nail.bashmakov@gmail.com
10 Anastasia D. Kirillova, Ph.D. (of Tech.), Senior Lecturer, Ufa University of Science and Technology, Ufa, Russia. E-mail: kirillova.andm@gmail.com

***Authors' contributions:*** *Vasilyev V. I. – planning research in the field of building attack detection systems using machine learning methods, conducting a comparative analysis of modeling results, preparing an analytical review. Vulfin A. M. – conducting an experimental study based on the developed software. Kartak V. M. – preparation of analytical review, experimental planning, software design. Bashmakov N. M. – preparation of data for modeling, interpretation of research results; generalization of research results; formulation of conclusions. Kirillova A. D. – software development, article manuscript design; work with graphic material.*

***Keywords:*** *deep learning, botnet control traffic, convolutional neural network classifiers, autoencoders, neural network models of embeddings.*

## References

1. *Wagle S. et al. Embedding alignment for unsupervised federated learning via smart data exchange // GLOBECOM 2022-2022 IEEE Global Communications Conference. IEEE, 2022, pp. 492–497.*
2. *McMaham H. B. et al. Communication-Efficient Learning of Deep Networks from Decentralized Data // arXiv preprint arXiv: 1602.05629 [cs.LG]. 2023. DOI: 10.48550/arXiv.1602.05629.*
3. *Wen J. et al. A Survey on Federated Learning: challenges and applications // International Journal of Machine Learning and Cybernetics. 2023, vol. 14, pp. 513–535.*
4. *Yang Q. et al. Federated Machine Learning: concept and applications // ACM Transactions on Intelligent Systems and Technology (TIST). 2019, vol. 10, no. 2, pp. 1–19. DOI: 10.1145/3298981.*
5. *Novikova E. S. et al. Federated Learning Based Intrusion Detection: System Architecture and Experiments // Voprosy kiberbezopasnosti. 2023, no. 6 (58), pp. 50–66. DOI: 10.21681/2311-3456-2023-6-50-66.*
6. *Novikova E. S. et al. Analytical review of intelligent intrusion detection systems based on federated learning: advantages and open challenges // Informatics and Automation. 2023, vol. 22, no. 5, pp. 1034–1082. DOI: 10.15622/ia.22.5.4.*
7. *Hernandez-Ramos J. L. et al. Intrusion Detection based on Federated Learning: a systematic review // arXiv preprint arXiv:2308.09522. 2023. DOI: 10.48550/arXiv.2308.09522.*
8. *Liu Y. et al. A secure federated transfer learning framework // IEEE Intelligent Systems. 2020 vol. 35, no. 4, pp. 70–82. DOI: 10.1109/MIS.2020.2988525.*
9. *Guo W. et al. A Comprehensive Survey of Federated Thansfer Learning: Challenges, Methods and Applications // arXiv preprint arXiv:2403.01387. 2024. DOI: 10.48550/arXiv.2403.01387.*
10. *Kholod I. et al. Open-Source Federated Learning Frameworks for IoT: A Comparative Review and Analysis // Sensors. 2020, no. 21, pp. 167. DOI: 10.3390/521010167.*
11. *Efremov M. A., Kholod I. I., Developing universal framework design for federated learning // Software & systems. 2022, vol. 35, no. 2, pp. 263–273. DOI: 10.15827/0236-235X.138.263-272.*
12. *Otoum K., Yadlappali S. K., Nayk A. FTLIoT: A Federated Thansfer Learning Framework for Securing IoT // GLOBECOM 2022-2022 IEEE Global Communications Conference. IEEE, 2022, pp. 1146–1151. DOI: 10.1109/GLOBECOM48099.2022.10001461.*
13. *Otoum K., Chamola V., Nayak A. Federated and Transfer Learning – Empowered Intrusion Detection for IoT Applications // IEEE Internet of Things Magazine. 2022, vol. 5, no. 3, pp. 50–54. DOI: 10.1109/IOTM.001.2200048.*
14. *Fan Y. et al. IoTDefender: A Federated Transfer Learning Intrusion Detection Framework for 5G IoT // 2020 IEEE 14th International Conference on Big Data Science and Engineering (BigDataSE). IEEE, 2020, pp. 88–95. DOI:10.1109/BigDataSE50710.2020.00020.*
15. *Rajesh L. T. et al. Give and Take: Federated Transfer Learning for Industrial IoT Network Intrusion Detection // 2023 IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom). IEEE, 2023, pp. 2365–2371. DOI: 10.1109/TrustCom60117.2023.00333.*
16. *Cheng Y. et al. Federated Transfer Learning with Client Selection for Intrusion Detection in Mobile Edge Computing // IEEE Communications Letters. 2022, vol. 26, no. 3, pp. 552–556. DOI: 10.1109/LCOMM.2022.3140273.*
17. *Wang K., Li J., Wu W. An efficient intrusion detection method based on federated transfer learning and an extreme learning machine with privacy preservation // Security and Communication Networks. 2022, vol. 2022, no. 1, pp. 2913293. DOI: 10.1155/2022/291329.*
18. *Guo W. et al. Federated transfer learning for auxiliary classifier generative adversarial networks: framework and industrial application // Journal of intelligent manufacturing. 2024, vol. 35, no. 4, pp. 1439–1454.*
19. *Metwaly A. A., Elhenawy I. Protecting IoT Devices from BotNet threats: a federated machine learning solution // Sustainable Machine Intelligence Journal. 2023, vol. 2, pp. 1–12. DOI: 10.61185/SMIJ.2023.22105.*
20. *Azizjon M., Jumabek A., Kim W. 1D CNN based network intrusion detection with normalization on imbalanced data // 2020 international conference on artificial intelligence in information and communication (ICAIIC). IEEE, 2020, pp. 218–224. DOI: 10.1109/ICAIIC48513.2020.9064976.*
21. *Novikova E. S., Chen Ya., Meleshko A. V. Methods for Assessing the Level of Data Heterogeneity in Federated Learning // XXVII International Conference on Soft Computing and Measurements (SCM'2024) (Saint Petersburg, May 22–24, 2024). 2024, pp. 446–450.*
22. *Yang Z. et al. A systematic literature review of methods and datasets for anomaly-based network intrusion detection // Computers & Security. 2022, vol. 116, pp. 102675. DOI: 10.1016/j.cose.2022.102675.*
23. *Lee G. et al. Network Intrusion Detection with Improved Feature Representation // 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). IEEE, 2021, pp. 2049–2054.*
24. *He Y., Yan D., Chen F. Hierarchical federated learning with local model embedding // Engineering Applications of Artificial Intelligence. 2023, vol. 123, pp. 106148. DOI: 10.1016/j.engappai.2023.106148.*
25. *Sivasubramanian A., Devisetty M., Bhavukam P. Feature Extraction and Anomaly Detection Using Different Autoencoders for Modeling Intrusion Detection Systems // Arabian Journal for Science and Engineering. 2024, pp. 1–13.*
26. *Wang Z. X. et al. Network traffic classification based on federated semi-supervised learning // Journal of Systems Architecture. 2024, vol. 149, pp. 103091. DOI: 10.1016/j.sysarc.2024.103091.*