УДК 004.056
doi:10.31799/1684-8853-2024-5-41-56
EDN: SWCOYY

Научные статьи
Articles

Обнаружение сетевых атак ботнетов на основе технологий машинного обучения и переноса знаний

Н. М. Башмаков[а], *аспирант, orcid.org/ 0000-0002-8647-6821*
В. И. Васильев[а], *доктор техн. наук, профессор, orcid.org/0000-0002-6105-5481*
А. М. Вульфин[а,б], *доктор техн. наук, профессор, orcid.org/0000-0001-5857-2413, vulfin.am@ugatu.su*
В. М. Картак[а], *доктор физ.-мат. наук, профессор, orcid.org/0000-0001-8167-8291*
А. Д. Кириллова[а], *канд. техн. наук, старший преподаватель, orcid.org/0009-0000-4164-2526*
[а] *Уфимский университет науки и технологий, Заки Валиди ул., 32, Уфа, 450076, РФ*
[б] *Омский государственный технический университет, Мира пр., 11, Омск, 644050, РФ*

Введение: *совершенствование сетевых средств защиты информации неразрывно связано с развитием инструментов интеллектуального мониторинга состояния и сетевого взаимодействия, повышающих наблюдаемость корпоративных информационных систем. Актуальной проблемой является оценка применимости предварительно обученных моделей машинного обучения к новым наборам данных сетевого трафика (с применением переноса обучения) и возможности их эксплуатации в реальных инфраструктурах для обнаружения узкого класса сетевых атак на примере взаимодействия скомпрометированных хостов с серверами управления ботнетов.* ***Цель:*** *совершенствование моделей и алгоритмов обнаружения сетевого трафика инфраструктур управления и контроля ботнетов в корпоративных информационных системах на основе технологий машинного обучения (в том числе глубокого обучения).* ***Результаты:*** *разработан прототип интеллектуальной системы обнаружения сетевых атак, позволяющей решать задачи сбора и предобработки данных сетевых сессий, обеспечивать взаимодействие с центром оперативного управления и мониторинга информационной безопасности, готовить данные для обучения локальных моделей анализа и управлять их жизненным циклом. Предложен алгоритм подготовки, предобработки трафика и оптимизации гиперпараметров бинарных классификаторов. Результаты экспериментов (F1-мера=0,71) подтверждают, что предлагаемые модели, обученные на одном наборе данных, могут успешно применяться на другом наборе узкоспециализированного домена трафика управления ботнетами. Отличительной особенностью является применение переноса обучения для глубоких нейросетевых моделей, что позволяет повысить эффективность обнаружения (величину F1-меры) специализированных сетевых атак на 16–21 %.* ***Практическая значимость:*** *применение переноса обучения обеспечивает возможность аккумулировать знания о проводимых атаках на различные информационные инфраструктуры в рамках единой нейросетевой модели, что позволяет повысить оперативность и достоверность обнаружения трафика управления ботнетами, тем самым усилить защищенность клиентских корпоративных информационных систем.* ***Обсуждение:*** *дальнейший подъем эффективности обнаружения специализированных сетевых атак возможен за счет применения более сложных нейросетевых моделей при использовании технологий федеративного трансферного обучения.*

Ключевые слова *— обнаружение сетевых атак, ботнеты, трафик управления, машинное обучение, глубокое обучение, трансферное обучение.*

Для цитирования: Башмаков Н. М., Васильев В. И., Вульфин А. М., Картак В. М., Кириллова А. Д. Обнаружение сетевых атак ботнетов на основе технологий машинного обучения и переноса знаний. *Информационно-управляющие системы*, 2024, № 5, с. 41–56. doi:10.31799/1684-8853-2024-5-41-56, EDN: SWCOYY
For citation: Bashmakov N. M., Vasilyev V. I., Vulfin A. M., Kartak V. M., Kirillova A. D. Detection of network botnet attacks based on machine learning and knowledge transfer technologies. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 41–56 (In Russian). doi:10.31799/1684-8853-2024-5-41-56, EDN: SWCOYY

Введение

В последние годы наблюдается устойчивая динамика роста количества сетевых атак на информационную инфраструктуру предприятий и организаций, возрастает сложность и разнообразие сценариев их реализации (https://ics-cert.kaspersky.ru/publications/reports/2024/03/19/threat-landscape-for-industrial-automation-systems-statistics-for-h2-2023/, https://www.ptsecurity.com/ru-ru/research/analytics/cybersecurity-threatscape-2023-q3/). Согласно прогнозам (https://www.crime-research.org/news/24.01.2024/4132), объем ущерба от киберпреступлений в мире к 2025 г. достигнет 12 трлн долл. Увеличилось число мультивекторных сетевых атак, направленных на выведение из строя сразу нескольких ключевых компонент информационной инфраструктуры корпоративных систем. Одним из основных инструментов реализации подобных атак являются ботнеты. Общее число устройств в ботнет-сетях, задействованных в атаках на IT-системы российских компаний в I квартале 2024 г., выросло в 1,6 раза и приближается к 5 млрд. Актуальной задачей является обнаружение сетевого трафика инфраструктур управления и контроля ботнетов (Command & Control, C&C) и пресечение их деятельности на ранних стадиях проникновения в корпоративные информационные инфраструктуры.

Наиболее современными средствами защиты корпоративных информационно-телекоммуникационных систем сегодня являются многоуровневые комплексы, включающие в себя средства обнаружения и предотвращения вторжений (Intrusion Detection System, IDS), расширенного обнаружения и реагирования (Endpoint Detection and Response, EDR/XDR) и анализа сетевого трафика (Network Traffic Analysis, NTA), интегрированные с системой управления событиями безопасности (Security Information and Event Management, SIEM), в составе центров оперативного управления и мониторинга информационной безопасности (ИБ). При этом классические сетевые IDS недостаточно хорошо справляются с обнаружением новых или модифицированных сценариев реализации специализированных типов сетевых атак (например, Advanced Persistent Threat, APT), они неспособны анализировать трафик защищенных соединений (TLS, VPN) и размещаются, как правило, на периметре сети. NTA-решения способны анализировать трафик внутри периметра сети, обладают сигнатурными и адаптивными моделями обнаружения сетевых атак, широко используются при расследовании инцидентов и активном поиске угроз (Threat Hunting).

Совершенствование сетевых средств защиты неразрывно связано с развитием инструментов интеллектуального мониторинга состояния и сетевого взаимодействия, повышающих наблюдаемость корпоративных информационных систем. Согласно [1], наибольшее число исследований применения технологий искусственного интеллекта в задачах обеспечения безопасности информационных систем посвящено именно вопросам создания средств обнаружения вторжений и анализа сетевого трафика.

Значительное количество исследователей использует для этих целей публичные наборы данных сетевого трафика с размеченными типами атак (например, KDD99, NSL-KDD, CIDDS, UNSW-NB15, BoT-IoT, ToN_IoT и т. п.). При этом многие популярные наборы данных сетевого трафика устарели [2] в связи с изменением ландшафта сетевых атак и эволюцией инфраструктуры корпоративных систем, при проведении исследований необходимо использовать актуальные данные. Еще одной особенностью является оценка эффективности предлагаемых решений: количественные и качественные характеристики моделей обнаружения атак получены, как правило, для данных из одного набора, примеры полноценного внедрения разработанных подходов для обнаружения сетевых атак в реальной инфраструктуре практически не освещаются. Таким образом, актуальной проблемой сегодня является оценка качества моделей обнаружения специализированных сетевых атак на различных наборах данных и возможности анализа данных в инфраструктуре реальной корпоративной сети.

Целью работы является совершенствование моделей и алгоритмов обнаружения сетевого трафика инфраструктур управления и контроля ботнетов в корпоративных информационных системах на основе технологий машинного обучения (в том числе глубокого обучения) и переноса знаний.

Анализ проблемы обнаружения сетевых атак ботнетов на основе технологий машинного обучения и переноса знаний

Обзор существующих публичных наборов данных сетевого трафика

В работе [2] проведен анализ 52 наиболее широко известных на сегодня наборов данных, которые могут быть использованы для исследования возможностей применения методов интеллектуального анализа и машинного обучения в задачах обнаружения вредоносного сетевого трафика, в том числе трафика инфраструктур управления и контроля ботнетов. Отмечено, что наиболее часто используемыми наборами данных являются KDD99 и NSL-KDD, но эти наборы данных устарели, и следует отказаться от их использования в пользу более актуальных. Однако авторы не указывают форматы представления анализируемых наборов данных, способы выделения ключевых признаков и используемое для этой цели программное обеспечение, что существенно затрудняет анализ «дрейфа данных», «дрейфа концепции» и возможности переноса обучения для конструируемых моделей обнаружения сетевых атак.

Сделаем далее допущение, что актуальными можно назвать наборы данных, появившиеся в последнее десятилетие, и произведем анализ форматов представления данных, перечня признаков и инструментальных средств их выделения (табл. 1).

Анализ актуальных наборов данных сетевого трафика показал, что для этапов предобработки, выделения сетевых сессий между конечными системами и последующего извлечения признаков сетевого взаимодействия конечных систем в составе вычислительной сети используется несколько основных подходов:

— для выделения сетевых сессий из дампа трафика канального уровня применяются как разнообразные сетевые инструменты, так и специализированные модульные инструменты обнаружения вторжений (например, Zeek, Bro IDS и пр.);

■ ***Таблица 1.*** Наборы данных сетевого трафика за 2015–2024 гг. [3–6]
■ ***Table 1.*** Network traffic datasets for the period 2015–2024 [3–6]

Название (год)	Используемые форматы	Формат набора данных	Количество признаков	Инструментальные средства обработки и извлечения признаков	Примечание
UNSW-NB15 (2015)	pcap, bro, argus, CSV (Netflow), отчеты	CSV	49	IXIA Perfect-Storm, Tcpdump, Argus, Bro-IDS	–
NDSec (2016)	pcap, CSV (журналы работы ОС)	pcap, CSV	–	YAF	Дамп pcap без извлечения сетевых сессий
DDoS 2016 (2016)	arff	CSV	28	Не указано	Недоступен
NGIDS-DS (2016)	pcap, CSV (журналы работы ОС)	pcap, CSV	–	IXIA Perfect Storm	Дамп pcap без извлечения сетевых сессий
UGR'16 (2016)	pcap, CSV (Netflow)	CSV	12	nfdump	–
Witty Worm (2016)	pcap	pcap	–	Не указано	Дамп pcap без извлечения сетевых сессий; исходный трафик собран в 2004 г.
Unified Host and Network (2016)	Netflow, JSON (журналы работы ОС)	CSV	11	Не указано	Нет разметки
CIDDS-001 (2017)	CSV (Netflow)	CSV	16	OpenStack	–
CIDDS-002 (2017)	CSV (Netflow)	CSV	16	OpenStack	–
CICIDS 2017 (2017)	pcap, CSV (Netflow)	CSV	71	CICFlowMeter	–
SUEE 2017 (2017)	pcap	pcap	–	Не указано	Дамп pcap без извлечения сетевых сессий
ISOT HTTP Botnet (2017)	pcap	pcap	–	Не указано	Дамп pcap без извлечения сетевых сессий; недоступен, создан в 2010 г.
PUF (2018)	–	–	–	–	Недоступен
ISOT CID (2018)	–	–	–	OpenStack, Tcpdump	Недоступен; исходный трафик собран в 2004 г.
CICDDoS 2019 (2019)	pcap, CSV (Netflow)	CSV	87	CICFlowMeter-V3	–
BoT-IoT (2019)	pcap, argus, csv (Netflow), отчеты	CSV	47	Не указано	–
MTA-KDD19 (2019)	CSV	CSV	33	–	–
IoT-23 (2020)	pcap, capinfos, dnstop, passivedns, tcpdstat, weblogng, zeek, bro	Zeek conn. log file	23	Zeek	2018–2019 гг.
InSDN (2020)	pcap, CSV (Netflow)	CSV	84	CICFlowMeter	–
CIRA-CIC-DoHBrw 2020 (2020)	pcap, CSV (Netflow)	CSV	28	DoHMeter	–
OPCUA (2020)	CSV (Netflow)	CSV	32	Python Scapy, Pyshark (Tshark)	–
TON_IoT (2021)	pcap, CSV	Zeek conn. log file	44/46	Zeek	–
VHS22 (2022)	CSV	–	48	–	–

■ *Окончание табл. 1*
■ *End of Table 1*

Название (год)	Используемые форматы	Формат набора данных	Количество признаков	Инструментальные средства обработки и извлечения признаков	Примечание
BH-KSU23 (2023)	CSV	–	78	CICFlowMeter	–
Trojan Detection (2023)	CSV	–	86	CICFlowMeter	–

– для описания сетевых сессий не используется единый набор признаков, например, предлагаемый семейством протоколов Netflow (xFlow), следовательно, количество и состав признаков существенно варьируются;

– достаточно часто применяются дополнительные алгоритмы конструирования признаков описания сетевого взаимодействия конечных систем (например, извлекается информация о времени доставки пакетов (information about interarrival time, IAT), определяются статистические параметры IAT, рассчитывается общее количество установленных флагов TCP, на сетевом уровне определяется общее количество подключений к (от) заданного хоста или соотношение подключений к (от) заданного хоста по отношению к общему количеству активных потоков [4]).

Следовательно, прямое сравнение эффективности работы моделей обнаружения сетевых атак на различных наборах данных без предварительного этапа преобразования признаков затруднительно.

В ряде случаев авторы наборов данных также публикуют исходные pcap-файлы (дампы трафика канального уровня), которые можно преобразовать в необходимый формат с использованием, например, таких сетевых инструментов, как tshark или cicflowmeter. Исследователи из австралийского университета Квинсленда опубликовали в едином формате Netflow такие актуальные наборы данных сетевого трафика, как UNSW-NB15, ToN-IoT, BoT-IoT, CSE-CIC-IDS2018. Таким образом, становится возможным прямое сравнение наборов данных и обученных на них моделей после унификации формата представления и набора признаков. Оценка эффективности моделей на новых наборах данных позволяет имитировать как условия появления новых атак (zero-days), так и «дрейф данных», связанный с изменениями сетевой инфраструктуры и характера взаимодействия конечных систем.

Анализ публикаций подтверждает широкое использование методов и алгоритмов машинного обучения для построения моделей обнаружения вредоносного сетевого трафика с чрезвычайно высокими показателями качества на исходных наборах данных [7, 8]. Среди рассматриваемых методов часто встречаются классические методы и модели машинного обучения: метод k-средних (K-means), метод опорных векторов (SVM), случайный лес (Random Forest), деревья решений C4.5 и т. п.

В последние несколько лет создано множество новых наборов данных, результаты анализа которых в публикациях представлены гораздо меньше: MTA-KDD19, Trojan Detection, CTU-IoT, VHS22 (является комбинацией наборов данных ISOT, CICIDS-17, CTU-13 и трафика с сайта Malware Traffic Analysis), BH-KSU23 и др.

Однако почти не встречаются публикации, в которых производится оценка классических моделей при их переносе на иные возможные наборы данных со сходными признаками или в условиях их применения для реальной инфраструктуры вычислительных сетей, а также не проводится сравнение с сигнатурными системами обнаружения сетевых атак.

Анализ возможностей машинного обучения и переноса знаний в задаче обнаружения сетевого трафика инфраструктур управления и контроля ботнетов

При использовании моделей на основе глубокого обучения, напротив, становится возможным применение трансферного обучения, или «переноса обучения» (Transfer Learning, TL) [9] – возможности дообучить модель на подмножестве новых данных, сохранив накопленные ранее знания и обобщающую способность модели [10–13] для обнаружения модификаций сетевых атак. Исследование эффективности трансферного обучения раскрывается в первую очередь на примере сверточных глубоких нейросетевых моделей, а также моделей глубоких автоэнкодеров.

На основе анализа публикаций [9, 11–13] могут быть выделены следующие сценарии трансферного обучения для глубоких нейросетевых моделей (табл. 2).

Другой значимой проблемой оценки возможности переноса моделей является разнообразный состав типов сетевых атак и способов их

■ ***Таблица 2.*** Сценарии применения трансферного обучения в задаче обнаружения сетевых атак
■ ***Table 2.*** Scenarios for applying transfer learning to network attack detection

№	Исследование	Модель для переноса обучения	Исходный набор данных	Набор данных, на котором протестирована модель	Доля правильных ответов, %	Прочие метрики
1	Wu P., Guo H., Buckland R. [9]	CNN-CNN	UNSW-NB15	NSL-KDD	81	FPR = 98,65 (сбалансирован)
2	Masum M., Shahriar H. [14]	DNN-DNN	VGG-16	NSL-KDD	70	Precision = 82,82 Recall = 82,15
3	Singla A., Bertino E., Verma D. [15]	DNN-DNN	UNSW-NB15 (175 341 запись)	UNSW-NB15 (девять выборок из исходного набора для каждого из девяти типов атак, сбалансированных по количеству примеров)	98	–
4	Fan Y., Li Y., Zhan M., Cui H., Zhang Y. [16]	CNN-CNN	CICIDS 2017	Собственный набор данных	91	FPR = 0,034 TPR = 0,99 TNR = 0,96
5	Idrissi I., Azizi M., Moussaoui O. [17]	CNN-CNN	BoT-IoT	TON-IoT	99	Precision = 0,99 Recall = 0,99
6	Rodríguez E., et al. [18]	CNN	BoT-IoT	UNSW-NB15	99	Precision = 0,996 Recall = 0,991 FPR = 0,05 F1 = 0,992
7	Gebresilassie S. K., Rafferty J., Chen L., Cui Z., Abu-Tair M. [19]	CNN	Собственный набор данных	Собственный набор данных	99	Precision = 0,981 Recall = 0,981 F1 = 0,981
8	Yehezkel A., Elyashiv E., Soffer O. [20]	DAE	Собственный набор данных	Собственный набор данных	98	–

реализации в различных наборах данных. Для проанализированных ранее наборов данных типы сетевых атак приведены в табл. 3.

Большинство типов сетевых атак (сетевое сканирование, брутфорс, DDoS) либо с высокой долей вероятности обнаруживаются и блокируются классическими средствами анализа сетевого трафика (IDS, NGFW), либо не представлены сразу в нескольких вышеупомянутых наборах данных в достаточном объеме, что затрудняет их сравнительный анализ. Следовательно, особое внимание необходимо обратить на анализ возможностей обнаружения специализированных сетевых атак, что является наиболее сложной задачей.

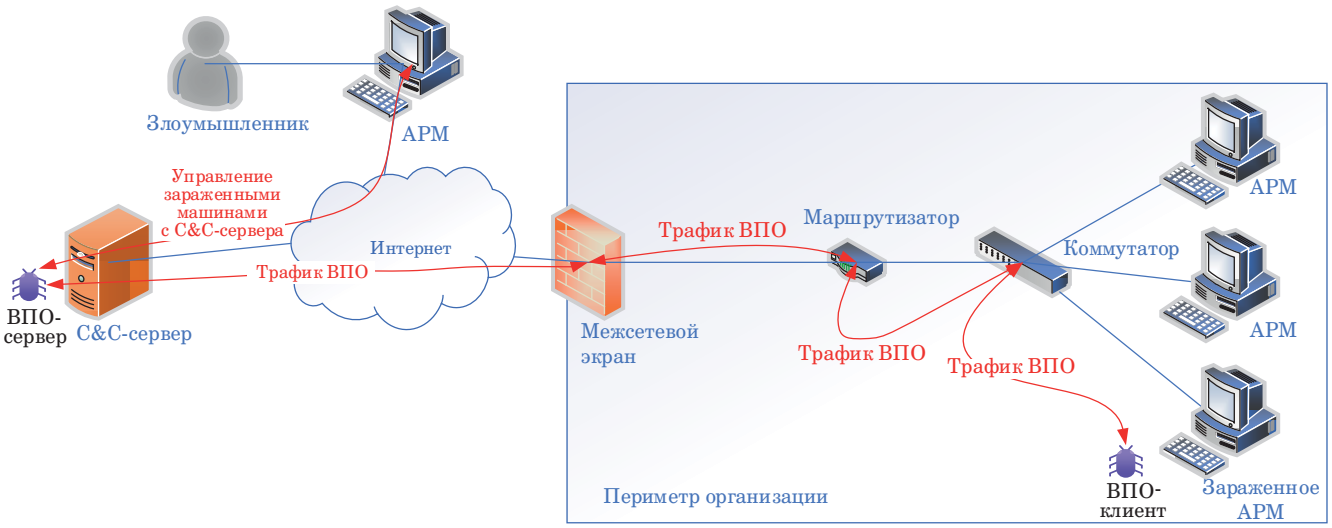
Анализ особенностей сетевого трафика инфраструктур управления и контроля ботнетов

Среди специализированных типов атак следует выделить сетевой трафик инфраструктур управления и контроля ботнетов, который характеризует взаимодействие серверов управления под контролем злоумышленника и скомпрометированных устройств. Взаимодействие со скомпрометированными устройствами заключается в отправке им команд и получении от них данных. Согласно базе знаний ATT & CK (Adversarial Tactics, Techniques & Common Knowledge) компании MITRE, сетевой трафик C&C является одной из тактик реализации сетевых атак. Злоумышленники могут использовать зараженные машины для проведения разведки внутри периметра, распространения вредоносного программного обеспечения или осуществления DDoS-атак.

Выделяют несколько топологий серверов C&C: звезда, звезда с несколькими серверами для обеспечения отказоустойчивости, иерархическая, случайная и P2P. Упрощенная (не приведены подробности внутренней сетевой инфраструктуры, C&C-сервер представлен в единственном экземпляре) схема взаимодействия C&C-сервера и зараженных устройств показана на рис. 1.

■ ***Таблица 3.*** Типы сетевых атак в актуальных наборах данных сетевого трафика
■ ***Table 3.*** Network attack types in current network traffic datasets

Набор данных	Количество типов сетевых атак	Типы сетевых атак в наборе данных
UNSW-NB15	9	Normal, Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, Worms
UGR'16	8	DoS, scan11, scan44, nerisbotnet, blacklist, anomaly-udpscan, anomaly-sshscan, anomaly-spam
Unified Host and Network	–	Нет разметки
CIDDS-001	4	Normal, portScan, DoS, pingScan, bruteForce
CIDDS-002	4	Normal, portScan, DoS, pingScan, bruteForce
CICIDS 2017	4	Benign, DDoS, PortScan, WebAttacks, Infilteration
CSE-CIC-IDS2018	6	Bruteforce attack, DoS attack, Web attack, Infiltration attack, Botnet attack, DDoS+PortScan
CICDDoS 2019	13	PortMap DDoS, NetBIOS DDoS, LDAP DDoS, MSSQL DDoS, UDP DDoS, UDP-Lag DDoS, SYN DDoS, NTP DDoS, DNS DDoS, SNMP DDoS, SSDP DDoS, WebDDoS, TFTP DDoS
Bot-IoT	4	Normal, DDoS, DoS, Reconnaisence, Theft
IoT-23 (CTU-IoT)	8	Benign, C&C, DDoS, FileDownload, HeartBeat, Mirai, Okiru, PartOfAHorizontalPortScan, Torii
InSDN	7	Normal, botnet, brute-force-attack, DoS, DDoS, Web_attack, Probe attack, Exploitation (R2L)
CIRA-CIC-DoHBrw	2	Non-DoH, Benign-DoH, Malicious-DoH
OPCUA	3	Normal, DoS, MITM, Impersonation
ToN-IoT	9	Normal, Backdoor, DDoS, DoS, Injection, Mitm, Password, Ransomware, Scanning, XSS
MTA-KDD19	1	Normal, Malware
Trojan Detection	1	Normal, Malware
VHS22	1	Normal, Malware
BH-KSU23	1	Normal, Malware



■ ***Рис. 1.*** Упрощенная схема C&C-взаимодействия: *АРМ* – автоматизированное рабочее место; *ВПО* – вредоносное программное обеспечение
■ ***Fig. 1.*** Simplified diagram of C&C interaction: *АРМ* – automated workstation; *ВПО* – malicious software

Особенности сетевого трафика C&C между зараженными устройствами и командным центром:

— сетевой трафик практически соответствует обычному использованию протоколов и аналогичен обычному трафику;

— объем трафика передачи управляющих команд небольшой;

— в анализируемой сети может быть очень мало зараженных устройств;

— использование шифрования сетевых C&C-сессий.

Злоумышленники часто используют различные техники для скрытия C&C-трафика, что усложняет его обнаружение классическим сигнатурным способом. Применение методов машинного обучения к обнаружению подобного рода сетевых атак, напротив, является перспективным ввиду их адаптивности и возможности обнаружения модифицированных сценариев их реализации за счет возможности дообучения на малых объемах данных.

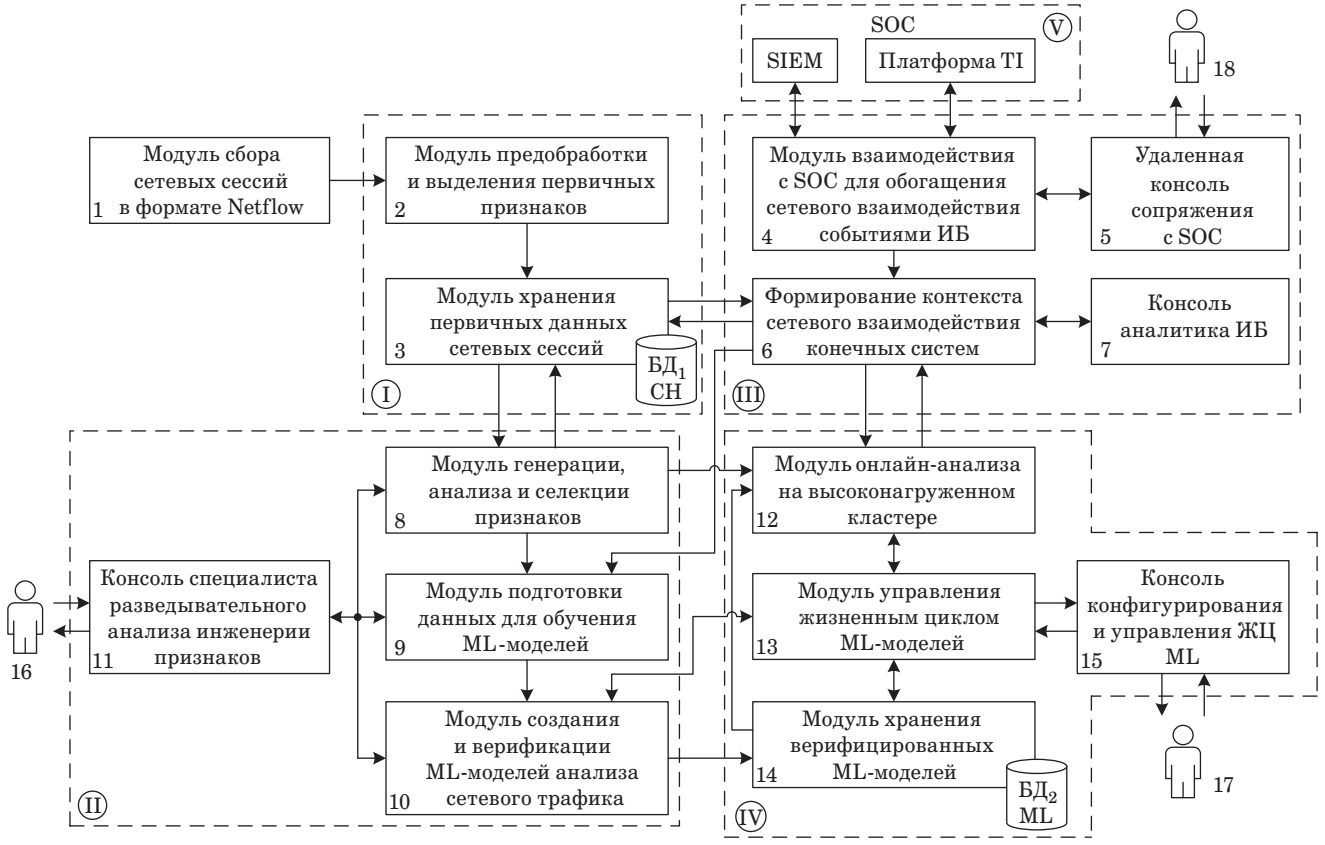
Следовательно, актуальной является как оценка применимости предварительно обученных моделей машинного обучения к новым наборам данных сетевого трафика (с применением переноса обучения), так и возможность их эксплуатации в реальных инфраструктурах для обнаружения узкого класса сетевых атак на примере взаимодействия скомпрометированных хостов с C&C-серверами.

Разработка системы обнаружения сетевого трафика инфраструктур управления и контроля ботнетов

Структурная схема предлагаемой системы обнаружения сетевого трафика инфраструктур управления и контроля ботнетов представлена на рис. 2. Данная структурная схема состоит из четырех подсистем (I, II, III и IV, символом V на схеме обозначен центр мониторинга ИБ, который получает оповещения от системы обнаружения).

Основные подсистемы:

— подсистема I получает данные от внешнего модуля *1* сбора сетевых сессий в формате Netflow, далее в модуле *2* осуществляется предобработка поступивших данных, выделяются ключевые первичные признаки (например, сетевые адреса и порты источника и назначения, количество переданных и полученных пакетов и байт, про-



■ ***Рис. 2.*** Структурная схема системы обнаружения сетевого трафика инфраструктур управления и контроля ботнетов

■ ***Fig. 2.*** Structural diagram of network traffic detection system for botnet management and checking infrastructures

токолы сетевого и вышестоящих уровней, длительность и т. д.) и производится запись сетевых сессий в хранилище на основе колоночной базы данных Yandex ClickHouse (модуль *3*). Специфика подсистемы в задаче обнаружения управляющего сетевого трафика C&C заключается в подборе узлов-источников Netflow (модуль *1*), а также составлении экспертами перечня наблюдаемых узлов сети исходя из их роли;

— в подсистеме II осуществляется генерация, анализ и селекция признаков (модуль *8*), выполняется подготовка данных для обучения моделей — создаются обучающие, проверочные и тестовые выборки по схеме k-fold перекрестной проверки (модуль *9*), создаются и верифицируются модели машинного обучения (модуль *10*), управление осуществляется специалистом разведывательного анализа данных и инженерии признаков (модуль *11*);

— подсистема III предназначена для создания контекста анализа сетевых сессий (совокупность сведений о принадлежности взаимодействующих узлов к сегментам сети, типам запрашиваемых ресурсов, задействованном программном обеспечении, пользовательских учетных записях и т. д., характеризующая событие ИБ, с которым ассоциирована данная сетевая сессия) с учетом событий ИБ в ходе взаимодействия с центром оперативного управления и мониторинга ИБ (согласно базе знаний и набору эвристических правил проводится дополнительная оценка сетевых сессий на предмет отношения к C&C-трафику);

— подсистема IV предназначена для управления жизненным циклом созданных моделей машинного обучения (модуль *13*) и обеспечением их работы для анализа поступающих событий в режиме, близком к реальному времени (модуль *12*). Обновление моделей производится либо по заданному сценарию, либо по достижении ими срока устаревания. Реализована подсистема с использованием фреймворков Apache Airflow и Apache Spark.

Вычислительный эксперимент по обнаружению сетевого трафика инфраструктур управления и контроля ботнетов

Для проведения серии экспериментов были выбраны специализированные наборы данных BH-KSU23 и Trojan Detection, содержащие размеченный трафик C&C-сессий и сессии нормальной работы (табл. 4) виртуальной сетевой инфраструктуры модельных стендов.

Схема серии экспериментов (рис. 3) включает следующие основные шаги.

1. Наборы данных BH-KSU (B) и Trojan Detection (T) в бинарном колоночном формате parquet размещаются в хранилище, доступном для последующей обработки.

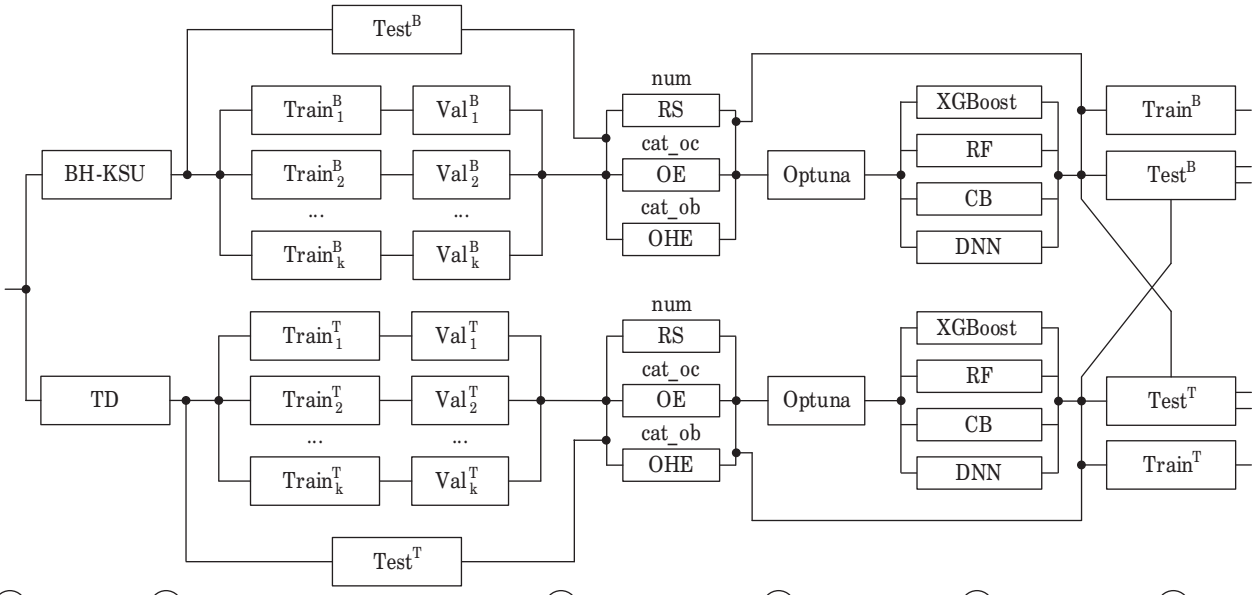
2. Из каждого набора данных выделяются 25 % примеров для тестовых выборок ($Test^B$ и $Test^T$) с сохранением пропорций между классами. На основе оставшихся данных готовятся выборки обучающих ($Train_i^B$ и $Train_j^T$) и проверочных (Val_i^B и Val_j^T) данных по схеме перекрестной проверки с пятью заходами с сохранением пропорций между классами (k-fold cross-validation в режиме стратификации).

3. Строится конвейер кодирования признаков. Все признаки разделяются на три группы: количественные признаки (используется RobustScaler, который при масштабировании вычитает из данных медиану и делит результат на интерквартильный размах), порядковые категориальные признаки (для кодирования используется OrdinarEncoder), номинальные признаки (для кодирования данных без внутренней иерархии используется OneHotEncoder) и категориальные порядковые признаки (для целевой переменной использован LabelEncoder). Параметры каждого типа кодировщиков настраиваются на обучающей выборке и применяются для преобразования проверочных и тестовых выборок во избежание «утечки данных».

■ ***Таблица 4.*** Характеристики наборов данных C&C

■ ***Table 4.*** Characteristics of C&C datasets

Набор данных	Источник	Количество признаков	Утилита извлечения признаков	Количество сетевых сессий	Особенности
BH-KSU23 (B)	Университет имени Короля Сауда (KSU), Саудовская Аравия	79	CICFlowmeter	400 000	Семь различных C&C-систем
Trojan Detection (T)	Университет Дрексела (Drexel), США	79	CICFlowmeter	180 000	1041 вариация серверной части C&C, 960 типов клиент-серверного взаимодействия C&C



■ ***Рис. 3.*** Схема проведения вычислительного эксперимента
■ ***Fig. 3.*** Scheme of conducting computational experiment

4. С помощью фреймворка Optuna выполняется оптимизация гиперпараметров четырех бинарных классификаторов на основе вероятностной оптимизации согласно алгоритму TPE.

5. Строятся бинарные классификаторы (табл. 5) с подобранными гиперпараметрами.

6. Выполняются прямая и перекрестная проверки обученных классификаторов на тестовых выборках из наборов данных T и B.

Бинарный классификатор на основе полносвязной глубокой нейронной сети (Deep Neural Network, DNN) используется в схеме трансферного обучения.

В схеме трансферного обучения модели DNN использованы 25 % данных из целевого набора (из подмножества Train) по следующим сценариям:

— полное дообучение модели (без «заморозки» весовых коэффициентов нейронов всех слоев);

— «заморозка» весовых коэффициентов первых двух слоев;

— «заморозка» весовых коэффициентов первых трех слоев.

Серия экспериментов показала, что наилучшие результаты достигаются в случае дообучения всей модели DNN, но применение «заморозки» первых двух слоев, выполняющих функцию извлечения признаков, снижает качество классификации на 2–5 %, сокращая время на дообучение на 30–40 % (в зависимости от размера целевого набора). Таким образом, эффективность применения переноса обучения заключается:

— в сокращении времени на дообучение модели на 30–40 % в случае «заморозки» слоев извлечения признаков;

— в возможности использовать предварительно обученную на большом объеме данных модель в сценариях обнаружения модификаций специализированных сетевых атак (в том числе при наличии ограниченного набора размеченных данных для дообучения, собранных для конкретной сети).

При оценке качества моделей машинного обучения, используемых для классификации, применяются следующие метрики:

— True Positive (TP) — количество правильно предсказанных положительных случаев;

— True Negative (TN) — количество правильно предсказанных отрицательных случаев;

— False Positive (FP) — количество неправильно предсказанных положительных случаев;

— False Negative (FN) — количество неправильно предсказанных отрицательных случаев;

■ ***Таблица 5.*** Используемые модели машинного обучения
■ ***Table 5.*** Machine learning models used

Модель	Описание	Основные параметры модели, подобранные в результате оптимизации гиперпараметров			Особенности использования
Random Forest	В случайном лесе объединяется множество деревьев решений, каждое обучается на разных подвыборках данных. Окончательные предсказания делаются путем усреднения предсказаний каждого дерева	n_estimators	Количество деревьев	100	Алгоритм случайного леса обладает более высокой предсказательной точностью по сравнению с одним деревом решений; возможность работы с категориальными переменными
		max_features	Количество параметров, которые следует учитывать	sqrt(n), n — количество примеров	
		max_depth	Максимальная глубина дерева	50	
		min_samples_split	Минимальное количество выборок, необходимое для разделения внутреннего узла	6	
		min_samples_leaf	Минимальное количество выборок, которое должно находиться в листовом узле	1	
XGBoost	Ансамблевый метод, который объединяет несколько слабых моделей-классификаторов на основе деревьев решений для создания одной сильной модели	max_depth	Максимальная глубина дерева	16	Обладает высокой точностью предсказаний благодаря своей способности последовательно уменьшать ошибки, часто достигая более высокой точности по сравнению с другими алгоритмами; возможность работы с категориальными переменными
		learning_rate	Скорость обучения	0,029	
		n_estimators	Количество слабых классификаторов в ансамбле	316	
		min_child_weight	Минимальная сумма веса экземпляра, необходимая для дочернего элемента	2	
		lambda	Коэффициент регуляризации L2 по весам	2,840	
		alpha	Коэффициент регуляризации L1 по весам	1,391	
		eta	Уменьшение размера шага для предотвращения переобучения	0,051	
		gamma	Минимальное сокращение потерь, необходимое для создания дальнейшего разделения на конечном узле дерева	0,383	
		subsample	Соотношение подвыборок обучающих экземпляров	0,997	
CatBoost	Осуществление выбора случайных образцов из набора данных, создание дерева решений для каждого выбранного образца, получение предсказаний от каждого дерева, проведение голосования для каждого предсказанного результата и выбор наиболее часто встречающегося предсказания как окончательного результата	iterations	Максимальное количество деревьев, которое можно построить при решении задач машинного обучения	8550	Отличается высоким качеством предсказаний без необходимости тонкой настройки параметров, поддержкой категориальных признаков, быстрой и масштабируемой версией для GPU, улучшенной точностью за счет уменьшения переобучения, быстрыми предсказаниями и хорошей работой с малым
		learning_rate	Скорость обучения	0,025	
		l2_leaf_reg	Коэффициент при члене регуляризации L2 функции стоимости	2,681	
		random_strength	Степень случайности, используемая для оценки разделений при выборе древовидной структуры	0,0003	
		depth	Глубина деревьев	9	
		bagging_temperature	Определяет настройки байесовского бутстрапа	0,211	

■ *Окончание табл. 5*

■ *End of Table 5*

Модель	Описание	Основные параметры модели, подобранные в результате оптимизации гиперпараметров			Особенности использования
		od_type	Тип используемого детектора переобучения	IncToDec	объемом данных; специально разработан для работы с категориальными данными
		od_wait	Количество итераций для продолжения обучения после итерации с оптимальным значением метрики	45	
		min_data_in_leaf	Минимальное количество обучающих выборок в листе	21	
		leaf_estimation_iterations	Параметр регулирует количество шагов, выполняемых в каждом дереве при вычислении значений листьев	3	
		max_ctr_complexity	Максимальное количество функций, которые можно объединить	3	
DNN	Полносвязная глубокая нейронная сеть прямого распространения	learning_rate_init	Коэффициент скорости обучения	0,085	Проблемой является представление категориальных переменных в векторе входных признаков
		1 layer	Количество нейронов	98	
		2 layer	Количество нейронов	128	
		3 layer	3-й слой + дропаут	64	
		4 layer	4-й слой + дропаут	32	
		5 layer	5-й слой	4	
		output layer	Один нейрон + логистическая функция активации	–	
		–	Функция потерь – Binary Cross-Entropy With Logits	–	
		learning_rate	Скорость обучения	constant	
		activation	Тип функции активации	relu	

– False Positive rate (FPR) – доля отрицательных объектов, неправильно предсказанных положительными:

$$FPR = FP/(FP + TN);$$

– Precision (точность) – показывает долю правильно предсказанных положительных случаев среди всех предсказанных положительных случаев:

$$Precision = \frac{TP}{TP + FP};$$

– Recall (полнота) – показывает долю правильно предсказанных положительных случаев среди всех реальных положительных случаев:

$$Recall = \frac{TP}{TP + FN};$$

– F1-мера – является гармоническим средним точности и полноты:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}.$$

Результаты оценки качества моделей приведены в табл. 6 и на рис. 4.

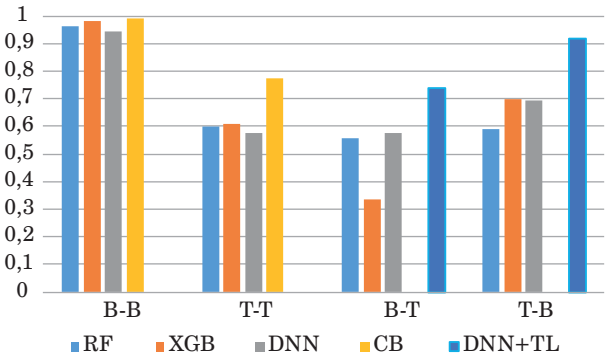
Высокие показатели F1-меры в экспериментах 1 и 2 (F1 > 0,96), а также ее относительно высокие показатели в экспериментах 6 и 7 (F1 > 0,6) указывают на то, что наборы данных BH-KSU23 и Trojan Detection имеют характеристики, позволяющие моделям машинного обучения успешно распознавать C&C-трафик.

Результаты экспериментов 9, 10 и 12, 13 (особенно эксперимент 13, где F1-мера достигла значения 0,70) подтверждают, что модели, обученные на одном наборе данных, могут успешно применяться на другом наборе в узкоспециализированном домене.

■ ***Таблица 6.*** Результаты серии экспериментов на наборах данных
■ ***Table 6.*** Results of a series of experiments on data sets

№	Модель	Исходный набор данных	Набор данных, на котором протестирована модель	Оценки качества модели на тестовом наборе			
				FPR	Precision	Recall	F1
1	Random Forest	BH-KSU23	BH-KSU23	0,042	0,949	0,978	0,964
2	Gradient Boosting	BH-KSU23	BH-KSU23	0,021	0,974	0,989	0,982
3	CatBoost	BH-KSU23	BH-KSU23	0,000	0,990	0,990	0,990
4	DNN	BH-KSU23	BH-KSU23	0,057	0,933	0,952	0,942
5	CatBoost	Trojan Detection	Trojan Detection	0,042	0,750	0,791	0,772
6	Random Forest	Trojan Detection	Trojan Detection	0,406	0,606	0,594	0,601
7	Gradient Boosting	Trojan Detection	Trojan Detection	0,322	0,652	0,574	0,611
8	DNN	Trojan Detection	Trojan Detection	0,342	0,670	0,656	0,663
9	Random Forest	BH-KSU23	Trojan Detection	0,647	0,504	0,629	0,559
10	Gradient Boosting	BH-KSU23	Trojan Detection	0,203	0,554	0,241	0,336
11	DNN	BH-KSU23	Trojan Detection	0,593	0,495	0,690	0,577
12	Random Forest	Trojan Detection	BH-KSU23	0,851	0,451	0,861	0,592
13	Gradient Boosting	Trojan Detection	BH-KSU23	0,517	0,581	0,883	0,701
14	DNN	Trojan Detection	BH-KSU23	0,523	0,572	0,879	0,693
15	DNN + TL	BH-KSU23	Trojan Detection	0,445	0,712	0,773	**0,741**
16	DNN + TL	Trojan Detection	BH-KSU23	0,221	0,876	0,966	**0,912**

Примечание: В моделях 1–14 TL не применяется, в моделях 15, 16 — применяется.



■ ***Рис. 4.*** Оценка F1-меры для серии экспериментов
■ ***Fig. 4.*** Estimation of F1-measure for a series of experiments

Результаты экспериментов с оптимизацией гиперпараметров с помощью фреймворка Optuna (эксперименты 3 и 5) демонстрируют, что модели CatBoost на наборах данных BH-KSU23 и Trojan Detection способны достичь высоких показателей точности, полноты и F1-меры после тщательной настройки гиперпараметров. Модель CatBoost на наборе данных Trojan Detection в эксперименте 5 показала высокие результаты, а на наборе данных BH-KSU23 — наилучшие результаты.

После завершения оптимизации гиперпараметров в экспериментах 3 и 5 были построены диаграммы оценки значимости признаков для классификаторов (результаты сведены в табл. 7).

Наиболее значимыми оказались признаки, характеризующие сетевой и транспортный уровни.

Результаты работы классификаторов с использованием глубокой нейронной сети сопоставимы с результатами моделей на основе комитетов деревьев решений. Однако при дообучении модели на 25 % данных из целевого набора результаты (значение F1-меры) на тестовой выборке значительно улучшаются (на 16,4 и 21,9 % соответственно) без существенного ухудшения результата на тестовом подмножестве исходного набора, что свидетельствует об эффективности трансферного обучения даже на сравнительно простых нейросетевых моделях.

Заключение

Для повышения эффективности систем обнаружения сетевого трафика инфраструктур управления и контроля ботнетов в корпоративных информационных системах предложено использовать модели и алгоритмы машинного

■ ***Таблица 7.*** Оценка значимости признаков для бинарной классификации
■ ***Table 7.*** Evaluation of the significance of features for binary classification

Эксперимент и целевой набор данных	Алгоритм оценки значимости признаков	Описание	Отобранные признаки	Описание отобранных признаков
3 BHKSU-23	Feature Importance (значимость параметра)	Алгоритм основан на оценке количества разбиений, которые использует признак в деревьях решений, и на уменьшении ошибки после использования признака	ECE_Flag_Cnt	Количество пакетов с флагом ECE
			Src_Port	Порт источника
			Pkt_Len_Var	Разница в длине пакета
5 Trojan Detection			Source_Port	Порт источника
			Bwd_URG_ Flags	Количество раз, когда флаг URG устанавливался для пакетов, отправляемых в обратном направлении
			Fwd_Header_ Length	Общее количество байт, используемых для заголовков в прямом направлении
3 BHKSU-23	Permutation Importance (важность перестановки)	Алгоритм оценивает, как перетасовка значений признака влияет на точность модели. Если перетасовка сильно снижает точность, значит, признак важен	Src_Port	Порт источника
			Dst_Port	Порт назначения
			Init_Bwd_Win Byts	Общее количество байт, отправленных в исходном окне в обратном направлении
5 Trojan Detection			Init_Win_ bytes_forward	Общее количество байт, отправленных в исходном окне в обратном направлении
			Source_Port	Порт источника
			Init_Win_ bytes_backward	Общее количество байт, отправленных в исходном окне в обратном направлении

обучения, в том числе глубокого обучения и переноса знаний (трансферного обучения).

Разработан прототип интеллектуальной системы обнаружения сетевых атак, позволяющей решать задачи сбора и предобработки данных сетевых сессий, обеспечивать взаимодействие с центром оперативного управления и мониторинга ИБ, готовить данные для обучения локальных моделей анализа и управлять их жизненным циклом. Проведенные вычислительные эксперименты позволяют сделать вывод о высокой эффективности обнаружения C&C-трафика с помощью предлагаемого подхода.

В ходе эксперимента был рассмотрен один конкретный тип атак — трафик C&C, на котором производилось обучение бинарных классификаторов на наборах данных Trojan Detection и BHKSU23. В этом случае эффективность обнаружения вредоносной активности при перекрестном применении моделей оказалась достаточно высокой.

Это в первую очередь связано с тем, что сходные типы атак имеют похожее отражение в параметрах сетевого трафика, и, соответственно, в наборах данных, что позволяет с осторожным оптимизмом предположить, что модель, обученная для обнаружения специализированных атак, сможет обнаруживать сходные типы атак и в реальном трафике. Дальнейшее повышение эффективности обнаружения сетевых атак C&C возможно за счет:

— более тщательного выбора параметров для анализа (с применением для группы моделей алгоритма оценки значимости с перекрестной проверкой);

— использования еще одного набора данных с трафиком целевой атаки для перекрестной проверки эффективности моделей (например, MTA-KDD-19) и применения технологий трансферного обучения;

— применения глубоких сверточных нейросетевых моделей и моделей с долгой-краткосрочной памятью;

— применения технологий федеративного трансферного обучения.

Финансовая поддержка

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

Литература

1. **Kaur R., Gabrijelčič D., Klobučar T.** Artificial intelligence for cybersecurity: Literature review and future research directions. *Information Fusion*, 2023, vol. 97, pp. 101804. doi:10.1016/j.inffus.2023.101804
2. **Yang Z., Liu X., Li T., Wu D., Wang J., Zhao Y., Han H.** 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
3. **Moustafa N.** A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets. *Sustainable Cities and Society*, 2021, vol. 72, pp. 102994. doi:10.1016/j.scs.2021.102994
4. **Szumelda P., Orzechowski N., Rawski M., Janicki A.** VHS-22 – a very heterogeneous set of network traffic data for threat detection. *2022 European Interdisciplinary Cybersecurity Conf.*, Barcelona, Spain, 15–16 June 2022, pp. 72–78. doi:10.1145/3528580.3532843
5. **Binsaeed K., Alaa-aldeen H.** *BH-KSU23: A Novel Dataset for Evaluating and Enhancing Intrusion Detection Systems Targeting Command-and-Control Traffic*. Mendeley Data, 2023, ver. 1. https://data.mendeley.com/datasets/wjxc69xj3n/1 (дата обращения: 03.10.2023).
6. **Subhadeep Ch.** *Trojan Detection [Data set],* 2021. doi:10.34740/KAGGLE/DSV/2625272. https://www.kaggle.com/datasets/subhajournal/trojan-detection (дата обращения: 03.10.2023).
7. **Vinayakumar R., Alazab M., Soman K. P., Poornachandran P., Al-Nemrat A., Venkatraman S.** Deep learning approach for intelligent intrusion detection system. *IEEE Access*, 2019, vol. 7, pp. 41525–41550. doi:10.1109/ACCESS.2019.2895334
8. **Jan S. U., Ahmed S., Shakhov V., Koo I.** Toward a lightweight intrusion detection system for the Internet of Things. *IEEE Access*, 2019, vol. 7, pp. 42450–42471. doi:10.1109/ACCESS.2019.2907965
9. **Wu P., Guo H., Buckland R.** A transfer learning approach for network intrusion detection. *2019 IEEE 4th Intern. Conf. on Big Data Analytics (ICBDA)*, *IEEE*, Suzhou, China, 15–18 March 2019, pp. 281–285. doi:10.1109/ICBDA.2019.8713213
10. **Беликов В. В.** Использование методов глубокого обучения с подкреплением для отбора признаков сетевого трафика при обнаружении компьютерных атак. *Программирование*, 2022, № 6, с. 3–13. doi:10.31857/S0132347422060024, EDN: KWOQRH
11. **Иогансон И.** Обзор методов федеративного обучения. *CEUR Workshop Proceeding*, 2023. https://www.researchgate.net/publication/376083291_ Obzor_ metodov_federativnogo_obucenia (дата обращения: 14.07.2024).
12. **Новикова Е. С., Котенко И. В., Мелешко А. В., Израилов К. Е.** 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
13. **Новикова Е. С., Федорченко Е. В., Котенко И. В., Холод И. И.** 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
14. **Masum M., Shahriar H.** TL-NID: Deep neural network with transfer learning for network intrusion detection. *2020 15th Intern. Conf. for Internet Technology and Secured Transactions (ICITST)*, *IEEE*, London, UK, 8 December 2020, pp. 1–7. doi:10.23919/ICITST51030.2020.9351317
15. **Singla A., Bertino E., Verma D.** Overcoming the lack of labeled data: Training intrusion detection models using transfer learning. *2019 IEEE Intern. Conf. on Smart Computing, IEEE*, Washington, USA, 12 June 2019, pp. 69–74. doi:10.1109/SMARTCOMP.2019.00031
16. **Fan Y., Li Y., Zhan M., Cui H., Zhang Y.** IoTDefender: A federated transfer learning intrusion detection framework for 5G IoT. *2020 IEEE 14th Intern. Conf. on Big Data Science and Engineering (BigDataSE)*, *IEEE*, Guangzhou, China, 31 December 2020, pp. 88–95. doi:10.1109/BigDataSE50710.2020.00020
17. **Idrissi I., Azizi M., Moussaoui O.** Accelerating the update of a DL-based IDS for IoT using deep transfer learning. *Indonesian Journal of Electrical Engineering and Computer Science*, 2021, vol. 23, no. 2, pp. 1059–1067. doi:10.11591/ijeecs.v23.i2.pp1059-1067
18. **Rodríguez E., Valls P., Otero B., Costa J. J., Verdú J., Pajuelo M. A., Canal R.** Transfer-learning-based intrusion detection framework in IoT networks. *Sensors*, 2022, vol. 22, no. 15, pp. 5621. doi:10.3390/s22155621
19. **Gebresilassie S. K., Rafferty J., Chen L., Cui Z., Abu-Tair M.** Transfer and CNN-based de-authentication (disassociation) DoS attack detection in IoT Wi-Fi networks. *Electronics*, 2023, vol. 12, no. 17, pp. 3731. doi:10.3390/electronics12173731
20. **Yehezkel A., Elyashiv E., Soffer O.** Network anomaly detection using transfer learning based on auto-encoders loss normalization. *Proceedings of the 14th ACM Workshop on Artificial Intelligence and Security*, Korea, 15 November 2021, pp. 61–71. doi:10.1145/3474369.3486869

UDC 004.056
doi:10.31799/1684-8853-2024-5-41-56
EDN: SWCOYY

Detection of network botnet attacks based on machine learning and knowledge transfer technologies

N. M. Bashmakov[a], Post-Graduate Student, orcid.org/ 0000-0002-8647-6821
V. I. Vasilyev[a], Dr. Sc., Tech., Professor, orcid.org/0000-0002-6105-5481
A. M. Vulfin[a,b], Dr. Sc., Tech., Professor, orcid.org/0000-0001-5857-2413, vulfin.am@ugatu.su
V. M. Kartak[a], Dr. Sc., Phys.-Math., Professor, orcid.org/0000-0001-8167-8291
A. D. Kirillova[a], PhD, Tech., Senior Lecturer, orcid.org/0009-0000-4164-2526
[a]Ufa University of Science and Technology, 32, Z. Validi St., Ufa, 450076, Russian Federation
[b]Omsk State Technical University, 11, Mira Pr., 644050, Omsk, Russian Federation

Introduction: The improvement of network information protection tools is inextricably linked to the development of tools for intelligent monitoring of the state and network interaction, increasing the observability of corporate information systems. A pressing issue is to assess the applicability of pre-trained machine learning models to new network traffic datasets (using transfer learning) and the possibility of their exploitation in real infrastructures to detect a narrow class of network attacks using the example of interactions between compromised hosts and botnet control servers. **Purpose:** To improve models and algorithms for detecting network traffic of botnet management and control infrastructures in corporate information systems based on machine learning technologies (including deep learning). **Results:** We develop a prototype of an intelligent network attack detection system, which makes it possible to solve the problems of collecting and pre-processing network session data, ensuring interaction with the operational control and information security monitoring center, preparing data for training local analysis models and managing their life cycle. We propose an algorithm for preparation, preprocessing of traffic and optimization of hyperparameters of binary classifiers. The experimental results (F1-measure = =0.71) confirm that the proposed models trained on one dataset can be successfully applied to another dataset of a highly specialized botnet control traffic domain. A distinctive feature is the use of transfer learning for deep neural network models, which makes it possible to increase the efficiency of detecting specialized network attacks by 16–21%. **Practical relevance:** The use of transfer learning makes it possible to accumulate knowledge about attacks on various information infrastructures within a single neural network model, which allows one to increase efficiency and reliability of detecting botnet control traffic, as well as to increase the security of client corporate information systems. **Discussion:** Further improvement of the efficiency of detection of specialized network attacks is possible through the use of more complex neural network models involving federated transfer learning technologies.

Keywords — network attack detection, botnets, control traffic, machine learning, deep learning, transfer learning.

For citation: Bashmakov N. M., Vasilyev V. I., Vulfin A. M., Kartak V. M., Kirillova A. D. Detection of network botnet attacks based on machine learning and knowledge transfer technologies. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 5, pp. 41–56 (In Russian). doi:10.31799/1684-8853-2024-5-41-56, EDN: SWCOYY

Financial support

The research was carried out in OmSU within the State assignment of the Ministry of Science and Higher Education of Russian Federation (theme No. FSGF-2023-0004).

References

1. Kaur R., Gabrijelčič D., Klobučar T. Artificial intelligence for cybersecurity: Literature review and future research directions. *Information Fusion*, 2023, vol. 97, pp. 101804. doi:10.1016/j.inffus.2023.101804
2. Yang Z., Liu X., Li T., Wu D., Wang J., Zhao Y., Han H. 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
3. Moustafa N. A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets. *Sustainable Cities and Society*, 2021, vol. 72, pp. 102994. doi:10.1016/j.scs.2021.102994
4. Szumelda P., Orzechowski N., Rawski M., Janicki A. VHS-22 – a very heterogeneous set of network traffic data for threat detection. *2022 European Interdisciplinary Cybersecurity Conf.*, Barcelona, 2022, pp. 72–78. doi:10.1145/3528580.3532843
5. Binsaeed K., Alaa-aldeen H. *BH-KSU23: A Novel Dataset for Evaluating and Enhancing Intrusion Detection Systems Targeting Command-and-Control Traffic*. Mendeley Data, 2023, ver. 1. Available at: https://data.mendeley.com/datasets/wjxc69xj3n/1 (accessed 3 October 2023).
6. Subhadeep Ch. *Trojan Detection [Data set]*, 2021. doi:10.34740/KAGGLE/DSV/2625272. Available at: https://www.kaggle.com/datasets/subhajournal/trojan-detection (accessed 3 October 2023).
7. Vinayakumar R., Alazab M., Soman K. P., Poornachandran P., Al-Nemrat A., Venkatraman S. Deep learning approach for intelligent intrusion detection system. *IEEE Access*, 2019, vol. 7, pp. 41525–41550. doi:10.1109/ACCESS.2019.2895334
8. Jan S. U., Ahmed S., Shakhov V., Koo I. Toward a lightweight intrusion detection system for the Internet of Things. *IEEE Access*, 2019, vol. 7, pp. 42450–42471. doi:10.1109/ACCESS.2019.2907965
9. Wu P., Guo H., Buckland R. A transfer learning approach for network intrusion detection. *2019 IEEE 4th Intern. Conf. on Big Data Analytics (ICBDA)*, Suzhou, IEEE, 2019, pp. 281–285. doi:10.1109/ICBDA.2019.8713213
10. Belikov V. V. Using deep reinforcement learning for selecting network traffic features in intrusion detection systems. *Programming and Computer Software*, 2022, vol. 48, no. 6, pp. 359–368. doi:10.31857/S0132347422060024, EDN: KWOQRH
11. Ioganson I. Review of federated learning methods. *CEUR Workshop Proceeding*, 2023. Available at: https://www.researchgate.net/publication/376083291_ Obzor_metodov_federativnogo_obucenia (accessed 14 July 2024) (In Russian).
12. Novikova E. S., Kotenko I. V., Meleshko A. V., Izrailov K. E. Federated learning based intrusion detection: system architecture and experiments. *Cybersecurity Issues*, 2023, no. 6(58), pp. 50–66 (In Russian). doi:10.21681/2311-3456-2023-6-50-66
13. Novikova E. S., Fedorchenko E. V., Kotenko I. V., Kholod I. I. 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 (In Russian). doi:10.15622/ia.22.5.4
14. Masum M., Shahriar H. TL-NID: Deep neural network with transfer learning for network intrusion detection. *2020 15th Intern. Conf. for Internet Technology and Secured Transactions (ICITST)*, London, IEEE, 2020, pp. 1–7. doi:10.23919/ICITST51030.2020.9351317
15. Singla A., Bertino E., Verma D. Overcoming the lack of labeled data: Training intrusion detection models using transfer learning. *2019 IEEE Intern. Conf. on Smart Computing*,

Washington, IEEE, 2019, pp. 69–74. doi:10.1109/SMARTCOMP.2019.00031
16. Fan Y., Li Y., Zhan M., Cui H., Zhang Y. IoTDefender: A federated transfer learning intrusion detection framework for 5G IoT. *2020 IEEE 14th Intern. Conf. on Big Data Science and Engineering (BigDataSE)*, Guangzhou, IEEE, 2020, pp. 88–95. doi:10.1109/BigDataSE50710.2020.00020
17. Idrissi I., Azizi M., Moussaoui O. Accelerating the update of a DL-based IDS for IoT using deep transfer learning. *Indonesian Journal of Electrical Engineering and Computer Science*, 2021, vol. 23, no. 2, pp. 1059–1067. doi:10.11591/ijeecs.v23.i2.pp1059-1067
18. Rodríguez E., Valls P., Otero B., Costa J. J., Verdú J., Pajuelo M. A., Canal R. Transfer-learning-based intrusion detection framework in IoT networks. *Sensors*, 2022, vol. 22, no. 15, pp. 5621. doi:10.3390/s22155621
19. Gebresilassie S. K., Rafferty J., Chen L., Cui Z., Abu-Tair M. Transfer and CNN-based de-authentication (disassociation) DoS attack detection in IoT Wi-Fi networks. *Electronics*, 2023, vol. 12, no. 17, pp. 3731. doi:10.3390/electronics12173731
20. Yehezkel A., Elyashiv E., Soffer O. Network anomaly detection using transfer learning based on auto-encoders loss normalization. *Proc. of the 14th ACM Workshop on Artificial Intelligence and Security*, Korea, 2021, pp. 61–71. doi:10.1145/3474369.3486869

УВАЖАЕМЫЕ АВТОРЫ!

Научные базы данных, включая Scopus и Web of Science, обрабатывают данные автоматически. С одной стороны, это ускоряет процесс обработки данных, с другой — различия в транслитерации ФИО, неточные данные о месте работы, области научного знания и т. д. приводят к тому, что в базах оказывается несколько авторских страниц для одного и того же человека. В результате для всех по отдельности считаются индексы цитирования, что снижает рейтинг ученого.

Для идентификации авторов в сетях Thomson Reuters проводит регистрацию с присвоением уникального индекса (ID) для каждого из авторов научных публикаций.

Процедура получения ID бесплатна и очень проста, есть возможность провести регистрацию на 12 языках, включая русский (чтобы выбрать язык, кликните на зеленое поле вверху справа на стартовой странице): https://orcid.org