УДК 004.056
Doi: 10.31772/2712-8970-2023-24-4-663-672

Для цитирования: Кононов Д. Д., Исаев С. В. Анализ киберугроз корпоративной сети на основе параллельной обработки данных Netflow // Сибирский аэрокосмический журнал. 2023. Т. 24, № 4. С. 663–672. Doi: 10.31772/2712-8970-2023-24-4-663-672.

For citation: Kononov D. D., Isaev S. V. [Analysis of corporate network cyber threats based on parallel processing of Netflow data]. *Siberian Aerospace Journal.* 2023, Vol. 24, No. 4, P. 663–672. Doi: 10.31772/2712-8970-2023-24-4-663-672.

Анализ киберугроз корпоративной сети на основе параллельной обработки данных Netflow

Д. Д. Кононов*, С. В. Исаев

Институт вычислительного моделирования СО РАН
Российская Федерация, 660036, г. Красноярск, ул. Академгородок, 50/44
*E-mail: ddk@icm.krasn.ru

Публичные сервисы различных организаций подвергаются постоянным кибератакам, что повышает риски информационной безопасности. Анализ сетевого трафика является важной задачей для обеспечения безопасного функционирования сетевой инфраструктуры, в том числе корпоративных сетей. В данной работе представлен обзор основных подходов для анализа сетевого трафика, приведены смежные работы, указаны недостатки существующих работ. Одним из методов является анализ данных сетевого трафика с использованием протокола Netflow, который позволяет сохранять данные о трафике на уровне L3 модели OSI. Особенностью исследования является использование длительных периодов наблюдения. При сохранении данных на длительных временных интервалах журналы имеют большой объем, что требует распараллеливания для первичной обработки данных. Авторами разработан кросс-платформенный программный комплекс распределенной обработки журналов сетевой активности, который использовался для анализа сетевой активности корпоративной сети Красноярского научного центра за 2021–2022 гг. Показана схема программного комплекса, описаны его возможности и особенности функционирования. Приведены источники данных для анализа и методика обработки. В работе были сформулированы и формализованы эвристические критерии аномальности сетевого трафика, которые сигнализируют о наличии возможных атак на сеть, также выделены датасеты по сетевой активности различных протоколов прикладного уровня. Для полученных наборов данных были рассчитаны статистические показатели, на основе которых получена информация об аномальной сетевой активности в течение двух лет. В работе проверен предложенный ранее авторами метод сравнения рисков киберугроз для различных временных интервалов, показавший существенное увеличение рисков для 50 % показателей в 2022 г. Сравнение месячных интервалов за различные годы показало аналогичное увеличение риска. Таким образом, метод доказал свою работоспособность и может применяться в других областях, в которых существуют группы критериев независимых показателей. Авторы привели планы по дальнейшему развитию методики анализа сетевой активности.

Ключевые слова: интернет, сетевая безопасность, анализ сетевого трафика, киберугрозы, корпоративная сеть.

Analysis of corporate network cyber threats based on parallel processing of Netflow data

D. D. Kononov*, S. V. Isaev

Institute of Computational Modelling of SB RAS
50/44, Akademgorodok St., Krasnoyarsk, 660036, Russian Federation
*E-mail: ddk@icm.krasn.ru

Public services of various organizations are subject to constant cyber attacks, which increases information security risks. Network traffic analysis is an important task to ensure the safe operation of network infrastructure, including corporate networks. This paper provides an overview of the main approaches for analyzing network traffic, provides related works, and points out the shortcomings of existing works. One method is to analyze network traffic data using the Netflow protocol, which allows traffic data to be stored at the L3 layer of the OSI model. A feature of the study is the use of long observation periods. When storing data over long time intervals, the logs become large, which requires parallelization for primary data processing. The authors developed a cross-platform software package for distributed processing of network activity logs, which was used to analyze the network activity of the corporate network of the Krasnoyarsk Scientific Center for 2021–2022. A diagram of the software package is shown, its capabilities and operating features are described. Data sources for analysis and processing methods are provided. In this paper the authors formulated and formalized heuristic criteria for the anomaly of network traffic, which identify the presence of possible network attacks, and extracted datasets on the network activity of various application-level protocols. For the obtained data sets, statistical indicators were calculated, information about anomalous network activity was obtained for two years. In this work we tested the previously proposed method for comparing the cyber threats risks for different time intervals, which showed a significant increase in risks for 50% of indicators in 2022. Comparisons of monthly intervals over different years showed similar increases in risk. Thus, the method has shown its efficiency and can be used in other areas in which there are groups of criteria for independent indicators. The authors proposed plans for further development of methods for analyzing network activity.

Keywords: Internet, network security, network traffic analysis, risk assessment, cyber threats, corporate network.

Введение

В настоящее время информационные технологии используются повсеместно для организации работы различных сервисов, в том числе корпоративных. Доступные публично сервисы подвержены рискам информационной безопасности, что требует организации комплексных мер по защите информации. Одним из важных частей обеспечения информационной безопасности является мониторинг и анализ сетевой активности (NTA – Network traffic analysis), который позволяет обнаружить аномалии в работе сети, идентифицировать причину (внешнюю или внутреннюю) и принять соответствующие меры. Анализ сетевой активности является актуальной задачей и применяется в различных областях. Например, анализ сетевой активности устройств IoT позволяет идентифицировать их тип [1] и выявлять проблемы безопасности [2]. Часто разрозненность стандартов и методик сбора и анализа сетевого трафика не позволяют воспроизвести результаты. Некоторые авторы предпринимают попытки создания универсальных форматов и программного обеспечения для анализа трафика [3]. Несмотря на то, что автоматизация процессов обработки и анализа трафика позволяет снизить количество рутинных операций, выявление аномалий с помощью человека-оператора представляет проблему из-за большого объема данных. Представляют интерес методики машинного обучения для выявления аномалий сетевой активности [4; 5]. Применение различных методик анализа и выявления особенностей трафика используется для мобильных устройств [6], в частности, анализ зашифрованного трафика позволяет с высокой

точностью определять используемые мобильные приложения [7]. Таким образом, анализ сетевой активности позволяет получить множество полезной информации, которая может быть использована для улучшения производительности и защищённости информационных систем, в том числе сервисов корпоративных сетей.

Важной частью обеспечения информационной безопасности является оценка рисков киберугроз. Применение различных методик анализа сетевого трафика позволяет выявлять аномалии сетевой активности, что дает возможность оценить риски киберугроз, которым подвержена сетевая инфраструктура. Оценка рисков может проводиться на различных временных интервалах, поэтому необходимо иметь возможность оценить динамику изменения рисков киберугроз.

Смежные работы

Анализ сетевого трафика позволяет выявлять аномалии в работе сетевой инфраструктуры. Методы выявления аномалий сетевого трафика можно разделить на три категории: неконтролируемые, контролируемые, частично контролируемые. Неконтролируемые методы достаточно распространены и не требуют предварительной подготовки данных. Предполагается, что нормальные данные встречаются в датасетах чаще аномальных [8]. Контролируемые методы предполагают построение моделей с разделением данных на две категории: нормальные и аномальные. Анализируемые данные сравниваются с помощью двух моделей, делается вывод о принадлежности данных к определенной категории [9]. Частично контролируемые методы предполагают построение модели только для нормальных данных [10] и являются более простыми и распространенными, чем контролируемые модели.

Одними из самых часто используемых методов анализа является кластерный анализ и методы классификации. Кластерный анализ предполагает разбиение свойств и атрибутов данных на кластеры, из которых выделяют нормальное и аномальное поведение [11]. Как правило, большие кластеры являются нормой, а малые аномалией. Методы классификации используются для разделения данных на заранее известные категории, их можно разделить на следующие виды: метод опорных векторов (SVM), нейросети, дерево решений, статистические методы. Одним из видов классификации с использованием контролируемого обучения является метод опорных векторов (SVM – support vector machine), который используется для обучения и тестирования большого объема данных. В работе [12] метод SVM применен для анализа трафика, достигнута высокая точность сопоставления. Нейросетевые методы используют модель нейронов со связями, которые преобразуют входной сигнал в выходной. Метод является распространенным и используется, например, для выявления злонамеренного траффика [13] и построения систем обнаружения вторжения (IDS – intrusion detections systems) [14]. Методы дерева решений позволяют построить дерево, состоящее из листьев, узлов и ребер, которое может быть использовано для многоступенчатой классификации данных. В работе [15] показана его высокая эффективность при анализе вредоносного трафика устройств IoT. Статистические методы основаны на теореме Байеса: если известен класс, возможно предсказать атрибуты; если класс неизвестен, правила могут определить класс на основе имеющихся атрибутов. В работе [16] байесовская сеть используется для анализа трафика и обеспечения безопасности облачных сервисов.

Важным компонентом при защите информационных систем и сетей передачи данных является оценка рисков киберугроз. Методы оценки рисков бывают количественные и качественные [17]. Количественные методы оперируют числовыми показателями, в то время как качественные методы – фиксированными категориями «высокий риск», «средний риск», «низкий риск». Как правило, качественные методы являются субъективными и полагаются на экспертные оценки, в то время как количественные являются объективными и позволяют воспроизвести результат. Существуют различные методы оценки риска киберугроз [18]. В работе [19] авторы используют различные модели количественного метода оценки рисков кибербезопасности облачных вычислений. Также применяются гибридные модели оценки рисков, комбинирую различные подходы [20]. Оценка рисков информационной безопасности применяется в различных областях:

индустриальные сети [21], системы управления Supervisory Control and Data Acquisition (SCADA) [22], устройства «интернета вещей» (IoT) [23]. Некоторые исследователи рассматривают оценку рисков как комплексную задачу, которая включает функциональную безопасность, физическую безопасность и кибербезопасность [24].

Существующие работы используют различных подходы к анализу трафика и оценке рисков киберугроз. Например, авторы разрабатывают новую методику и используют стандартные тестовые датасеты, что не позволяет оценить эффективность метода на реальной сетевой инфраструктуре. Другие авторы используют реальные данные с короткими временными интервалами, что не позволяет сделать глубокий анализ и выявить динамику происходящих процессов. В данной работе проводится исследование безопасности корпоративной сети Красноярского научного центра (ФИЦ КНЦ СО РАН) на основе анализа сетевого трафика Netflow. В работе описано применение предложенного ранее авторами статистического метода сравнения рисков киберугроз [25]. В отличие от существующих работ, метод использует длительные временные интервалы при анализе реального сетевого трафика и позволяет сравнить риски киберугроз на произвольных интервалах.

Источник данных и методика обработки

Источником данных для анализа являются данные сетевой активности Netflow пограничного маршрутизатора и прокси-сервера корпоративной сети Красноярского научного центра за 2021–2022 гг. (объем 680 Гб, 19,5 млрд записей). Обработка данных производится в несколько этапов: 1) сбор статистики сетевых интерфейсов с помощью агентов nfcapd; 2) извлечение данных с помощью утилиты nfdump; 3) фильтрация и агрегирование необходимых полей на основе заданных правил; 4) сохранение полученных данных для анализа; 5) применение статистического анализа.

Обработка данных выполняется с помощью разработанного авторами программного комплекса распределенной обработки сетевой активности GNetProc на языке Go. Программный комплекс состоит из нескольких частей: клиент, брокер сообщений, вычислительный кластер (рис. 1).

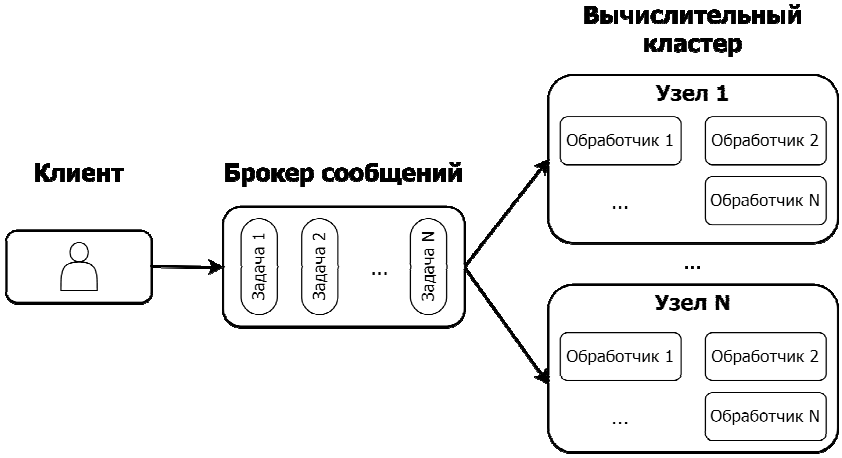

Рис. 1. Архитектура GNetProc

Fig. 1. GNetProc software architecture

Клиентская часть взаимодействует с брокером сообщений и позволяет отправлять задания на обработку в общую очередь заданий, серверная часть извлекает задание из очереди и запускает обработку с помощью обработчиков (workers). Серверная часть состоит из множества узлов вычислительного кластера, каждый узел может запускать несколько обработчиков (workers) параллельно, что обеспечивает параллельную обработку данных и позволяет умень-

шить время выполнения заданий. Обработчик обращается к брокеру сообщений, получает новое задание, запускает его на выполнение, сохраняет результаты обработки, затем получает новое задание. Задание состоит из списка элементов <имя, значение>, которые включают источник данных, приемник данных, тип задания, параметры агрегации, начало и окончание временных интервалов, фильтр и селектор данных. Фильтр посредством специального языка позволяет задавать набор правил для фильтрации данных. Селектор задает список полей, которые после фильтрации и агрегации будут извлечены и записаны как результат выполнения задания. Брокер сообщений использует СУБД Redis для обработки очереди заданий. Для описания заданий используется язык YAML. Программный комплекс является кросс-платформенным и поддерживает работу в гетерогенных конфигурациях с операционными системами Linux, *BSD, Windows.

Анализ данных

Для оценки уровня риска киберугроз на основании принципов нормального функционирования различных протоколов были сформулированы эвристические критерии аномальности трафика:

1. Входящие потоки TCP, имеющие менее 4-х пакетов, свидетельствует о том, что TCP-соединение не было завершено регламентированным способом. Подозрение на DoS-атаку.
2. Входящие потоки TCP длительностью 0 свидетельствует о том, что TCP-соединение не было установлено. Подозрение на DoS-атаку.
3. Большое количество входящих UDP потоков – подозрение на DDoS атаку.
4. Превышение количества входящих UDP потоков над исходящими – подозрение на DoS-атаку.
5. Превышение количества входящих TCP потоков над исходящими – подозрение на DoS-атаку.
6. Входящие потоки TCP и UDP на неиспользуемые, но маршрутизируемые адреса, – подозрение на сканирование уязвимостей (Unused).
7. Входящие потоки TCP на распространенные сервисы (MSSQL, MySQL, RDP, SMB, SMTP, SSH, Telnet) – попытки несанкционированного входа.

Эти критерии были формализованы в виде правил системы обработки первичного трафика согласно заданному синтаксису. После обработки первичных данных с помощью созданного ПО были получены агрегированные данные для дальнейшего анализа. Каждый набор данных представляет собой совокупность пар <метка времени, интегральная характеристика>. Типичный набор данных за год состоит из 105 тыс. записей (количество пятиминутных интервалов в году) и может быть обработан средствами настольных систем, таких как MS Excel. На рис. 2 представлена диаграмма количества потоков нулевой длительности за 2022 г. Аномальные значения не имеют выраженной периодичности, на несколько порядков превышают среднее значение (35256), за счет чего можно автоматизировать их выявление.

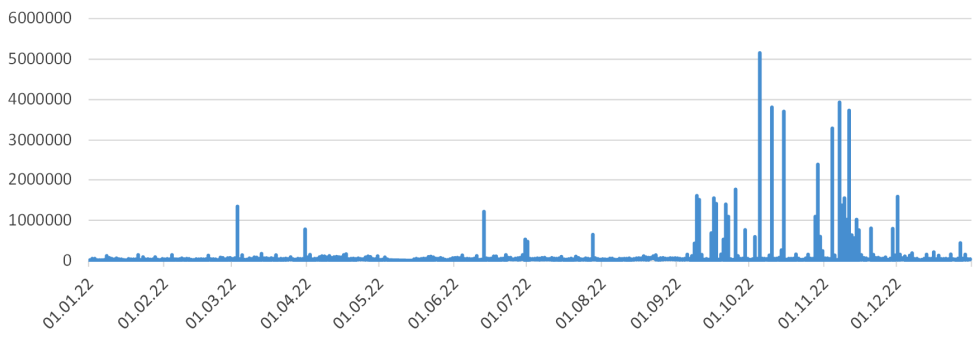

Рис. 2. Количество TCP-потоков нулевой длительностью

Fig. 2. Number of TCP flows with zero duration

Полученное распределение имеет явно выраженную правую асимметрию, удовлетворяется следующее условие:

$$Mode < Median < Mean.$$

На рис. 3 представлена гистограмма распределения частот TCP-потоков длительностью 0 с. Более 99 % выборки содержатся в интервале [μ – 3*σ, μ + 3*σ], что позволяет идентифицировать киберугрозы на основе данных выбросов.

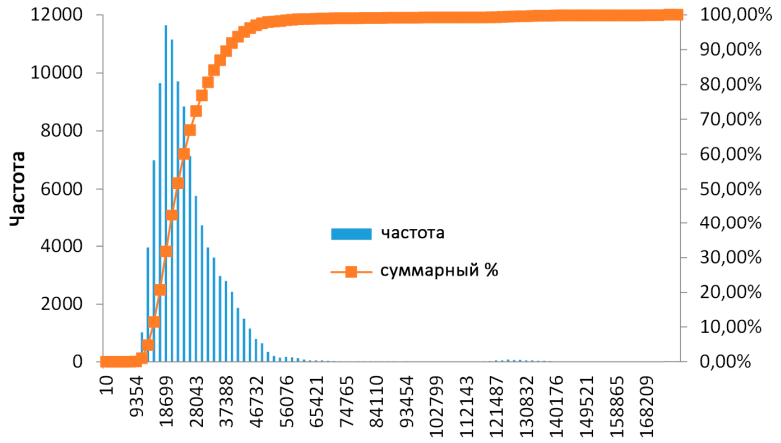

Рис. 3. Гистограмма распределения частот TCP-потоков длительностью 0

Fig. 3. Histogram of frequency distribution of TCP flows with zero duration

Для всех полученных наборов данных были рассчитаны показатели: среднее, мода, медианы, стандартное отклонение и процент выбросов со значениями больше 3 стандартных отклонений от среднего. Дополнительно было рассчитано относительное увеличение стандартного отклонения показателей в 2022 г. по сравнению с 2021 г. Результаты представлены в таблице: Mean – среднее по выборке μ; S – стандартное отклонение σ; Me – медиана выборки; Mo – мода выборки; P – процент данных, лежащих за пределами диапазона [μ – 3*σ, μ + 3*σ]; ΔS – относительный прирост S в 2022 г. относительно 2021 г.

Результаты показателей, рассчитанные по критериям аномальности трафика

Название	2021 г.					2022 г.					ΔS, %
	Mean	S	Me	Mo	P, %	Mean	S	Me	Mo	P, %	
TCP-потоки длит. 0	25205	8343	22079	18975	2,15	35257	14301	29059	22951	1,13	71
TCP-потоки < 4 пакетов	34906	11608	30188	27535	2,35	45287	17732	37461	35331	1,02	53
UDP-потоки	15071	5809	13541	13217	2,62	17221	6069	15589	13965	2,95	4
UDP (вход–выход)	1878	2114	1678	1532	2,08	3511	2428	3175	1325	1,90	15
TCP (вход–выход)	16759	8751	14846	12540	2,06	26087	11707	21363	16602	1,42	34
Неисп. TCP	20615	6357	18662	15287	2,30	22035	6736	20457	15148	0,98	6
Неисп. UDP	965	500	762	376	1,82	1519	1184	862	454	4,24	137
MSSQL	128	63	93	87	7,01	96	30	82	82	3,93	–52
MySQL	21	19	10	3	2,93	32	31	13	5	4,25	63
RDP	105	66	78	39	3,97	120	80	84	40	4,27	21
SMB	367	267	190	128	3,33	224	128	154	129	7,84	–52
SMTP	190	80	166	139	3,88	195	95	162	104	2,98	19
SSH	281	160	197	141	4,19	494	389	270	226	4,02	143
Telnet	405	70	384	363	4,10	922	259	953	1151	0,17	270

Все распределения, кроме Telnet 2022, обладают правосторонней асимметрией. Набор данных Telnet 2022 имеет два максимума гистограммы, что связано с перенастройкой оборудования сбора данных. На рис. 4 показаны гистограммы за 2021 и 2022 гг.

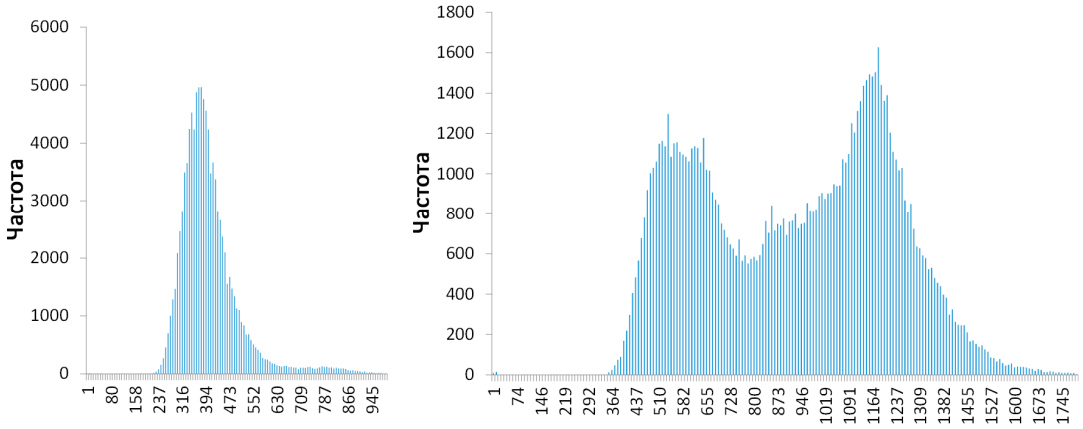

Рис. 4. Гистограмма наборов Telnet за 2021 и 2022 гг.

Fig. 4. Histogram of Telnet datasets for 2021 and 2022

Среднеквадратическое отклонение для большинства наборов данных за 2022 г. имеет существенное увеличение по сравнению с 2021 г., что свидетельствует об увеличении меры неопределенности появления различных компьютерных атак. Применяя метод сравнения рисков, описанный в [25] для выборок 2021 (V_1) и 2022 (V_2) гг., вычислим функцию R оценки изменения рисков:

$$R(V_1, V_2) = \frac{1}{N} * \sum_{i=1}^{N} K_i ,$$

где $K_i = \begin{cases} 1, \text{ если } \mu_{2i} > \mu_{1i} + 0,6745 * \sigma_{1i}, \\ 0, \text{ если } \mu_{1i} - 0,6745 * \sigma_{1i} \le \mu_{2i} \le \mu_{1i} + 0,6745 * \sigma_{1i}, \\ -1, \text{ если } \mu_{2i} < \mu_{1i} - 0,6745 * \sigma_{1i}; \end{cases}$ μ_{ji} – среднее значение выборки i-го признака выборки V_j; σ_{ji} – среднеквадратическое отклонение выборки i-го признака выборки V_j.

Половина из исследуемых показателей дает 1 в сумму, остальные дают 0, в результате $R(2021, 2022) = 0,5$. Это можно проинтерпретировать таким образом: риск в рамках исследуемых критериев в целом в 2022 г. по сравнению с 2021 г. существенно вырос – на 50 % показателей. При сравнении интервалов за март 2021 г. и март 2022 г. мы получаем значение $R = 0,43$. Таким образом, предложенный метод позволяет сравнивать риски как на больших временных интервалах (год), так и на средних, при условии наличия достаточной выборки.

Заключение

В работе исследованы риски кибератак на основе данных интернет-трафика сети Красноярского научного центра СО РАН за 2021 и 2022 гг. Разработано кросс-платформенное программное обеспечение для обработки больших объемов данных, которое обладает возможностями масштабирования с помощью параллельной обработки и может выполнять анализ данных в режиме реального времени. Сформулированы и формализованы эвристические критерии аномальности интернет-трафика, сигнализирующие о возможных атаках на сеть. Рассчитаны статистические показатели для полученных наборов данных, на основе которых сделаны выводы о форме распределений и динамике атак. Проверен предложенный авторами метод сравнения рисков киберугроз для годичных и месячных интервалов, который показал схожее увеличение

рисков. Так как метод не зависит от сравниваемых временных интервалов и объема выборок, он может использоваться в других областях, в которых существуют группы критериев независимых показателей. Следующей задачей является расширение критериев анализа с учётом флагов TCP-соединения и протоколов прикладного уровня с учётом логики их функционирования. Планируется использование разработанного программного обеспечения при создании наборов данных для методов машинного обучения при решении задач идентификации киберугроз.

Библиографические ссылки

1. Shahid M. R., Blanc G., Zhang Z., and Debar H. IoT Devices Recognition Through Network Traffic Analysis // 2018 IEEE International Conference on Big Data (Big Data). 2018. P. 5187–5192.

2. Sairam R., Bhunia S. S., Thangavelu V., Gurusamy M. NETRA: Enhancing IoT security using NFV-based edge traffic analysis // IEEE Sensors Journal. 2019. No. 19(12). P. 4660–4671.

3. Towards Reproducible Network Traffic Analysis / J. Holland, P. Schmitt, P. Mittal, N. Feamster // arXiv preprint arXiv:2203. 2022. 12410 p.

4. Alqudah N., Yaseen Q. Machine learning for traffic analysis: a review // Procedia Computer Science. 2020. Vol. 170. P. 911–916.

5. Abbasi M., Shahraki A., Taherkordi A. Deep learning for network traffic monitoring and analysis (NTMA): A survey // Computer Communications. 2021. Vol. 170. P. 19–41.

6. The dark side (-channel) of mobile devices: A survey on network traffic analysis / M. Conti, Q. Q. Li, A. Maragno, R. Spolaor // IEEE communications surveys & tutorials. 2018. No. 20(4). P. 2658–2713.

7. Robust smartphone app identification via encrypted network traffic analysis / V. F. Taylor, R. Spolaor, M. Conti, I. Martinovic // IEEE Transactions on Information Forensics and Security. 2017. Vol. 13(1). P. 63–78.

8. Goldstein M., Uchida S. A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data // PloS one. 2016. Vol. 11(4). P.e0152173.

9. Garg R., Mukherjee S. A comparative study using supervised learning for anomaly detection in network traffic // Journal of Physics: Conference Series. 2022. Vol. 2161, No. 1. P. 012030.

10. Comparison of supervised, semi-supervised and unsupervised learning methods in network intrusion detection system (NIDS) application / N. S. Arunraj, R. Hable, M. Fernandes et al. // Anwendungen und Konzepte der Wirtschaftsinformatik. 2017. No. 6. P. 10–19.

11. Zhang P., Ma W., Qian S. Cluster analysis of day-to-day traffic data in networks // Transportation Research Part C: Emerging Technologies. 2022. Vol. 144. P. 103882.

12. Retracted: Traffic identification and traffic analysis based on support vector machine / W. Zhongsheng, W. Jianguo, Y. Sen, G. Jiaqiong // Concurrency and Computation: Practice and Experience. 2020. Vol. 32(2). P. e5292.

13. Malicious network traffic detection based on deep neural networks and association analysis / M. Gao, L. Ma, H. Liu et al. // Sensors. 2020. Vol. 20(5). P. 1452.

14. Vinayakumar R., Soman K. P., Poornachandran P. Evaluation of recurrent neural network and its variants for intrusion detection system (IDS) // International Journal of Information System Modeling and Design (IJISMD). 2017. Vol. 8(3). P. 43–63.

15. Comparing Malware Attack Detection using Machine Learning Techniques in IoT Network Traffic / Y. Z. Wei, M. Md-Arshad, A. A. Samad, N. Ithnin // International Journal of Innovative Computing. 2023. Vol. 13(1). P. 21–27.

16. Nie L., Jiang D., Lv Z. Modeling network traffic for traffic matrix estimation and anomaly detection based on Bayesian network in cloud computing networks // Annals of Telecommunications. 2017. Vol. 72. P. 297–305.

17. Landoll D. The security risk assessment handbook: A complete guide for performing security risk assessments // CRC Press. 2021.

18. Macek D., Magdalenic I., Redep N. B. A systematic literature review on the application of multicriteria decision making methods for information security risk assessment // International Journal of Safety and Security Engineering. 2020. Vol. 10, No. 2. P. 161–174.

19. Jouini M., Rabai L. B. A. Comparative study of information security risk assessment models for cloud computing systems // Procedia Computer Science. 2016. Vol. 83. P. 1084–1089.

20. Haji S., Tan Q., Costa R. S. A hybrid model for information security risk assessment // Int. j. adv. trends comput. sci. eng. 2019. ART-2019-111611.

21. Summary of research on IT network and industrial control network security assessment / L. Hu, H. Li, Z. Wei et al. // 2019 IEEE 3rd information technology, networking, electronic and automation control conference (ITNEC). 2019. P. 1203–1210.

22. A review of cyber security risk assessment methods for SCADA systems / Y. Cherdantseva, P. Burnap, A. Blyth et al. // Computers & security. 2016. Vol. 56. P. 1–27.

23. Mahak M., Singh Y. Threat modelling and risk assessment in internet of things: A review // Proceedings of Second International Conference on Computing, Communications, and Cyber-Security: IC4S 2020, 2021. Springer Singapore. P. 293–305.

24. Lyu X., Ding Y., Yang S. H. Safety and security risk assessment in cyber-physical systems // IET Cyber-Physical Systems: Theory & Applications, 2019. Vol. 4(3). P. 221–232.

25. Исаев С. В., Кононов Д. Д. Исследование динамики и классификация атак на веб-сервисы корпоративной сети // Сибирский аэрокосмический журнал. 2022. Т. 23, № 4. С. 593–600.

References

1. Shahid M. R., Blanc G., Zhang Z., Debar H. IoT Devices Recognition Through Network Traffic Analysis. *2018 IEEE International Conference on Big Data (Big Data)*. 2018, P. 5187–5192.

2. Sairam R., Bhunia S. S., Thangavelu V., Gurusamy M. NETRA: Enhancing IoT security using NFV-based edge traffic analysis. *IEEE Sensors Journal*. 2019, Vol. 19(12), P. 4660–4671.

3. Holland J., Schmitt P., Mittal P., Feamster N. Towards Reproducible Network Traffic Analysis. *arXiv preprint* arXiv:2203. 2022, P. 12410.

4. Alqudah N., Yaseen Q. Machine learning for traffic analysis: a review. *Procedia Computer Science*. 2020, Vol. 170, P. 911–916.

5. Abbasi M., Shahraki A., and Taherkordi A. Deep learning for network traffic monitoring and analysis (NTMA): A survey. *Computer Communications*. 2021, Vol. 170, P. 19–41.

6. Conti M., Li Q. Q., Maragno A., Spolaor R. The dark side (-channel) of mobile devices: A survey on network traffic analysis. *IEEE communications surveys & tutorials*. 2018, Vol. 20(4), P. 2658–2713.

7. Taylor V. F., Spolaor R., Conti M., Martinovic I. Robust smartphone app identification via encrypted network traffic analysis. *IEEE Transactions on Information Forensics and Security*. 2017, Vol. 13(1), P. 63–78.

8. Goldstein M., and Uchida S. A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. *PloS one*. 2016, Vol. 11(4), P. e0152173.

9. Garg R., Mukherjee S. A comparative study using supervised learning for anomaly detection in network traffic. *Journal of Physics: Conference Series*. 2022, Vol. 2161, No. 1, P. 012030.

10. Arunraj N. S., Hable R., Fernandes M., Leidl K., Heigl M. Comparison of supervised, semi-supervised and unsupervised learning methods in network intrusion detection system (NIDS) application. *Anwendungen und Konzepte der Wirtschaftsinformatik*. 2017, No. 6, P. 10–19.

11. Zhang P., Ma W., Qian S. Cluster analysis of day-to-day traffic data in networks. *Transportation Research Part C: Emerging Technologies*. 2022, Vol. 144, P. 103882.

12. Zhongsheng W., Jianguo W., Sen Y., Jiaqiong G. Retracted: Traffic identification and traffic analysis based on support vector machine. *Concurrency and Computation: Practice and Experience*. 2020, Vol. 32(2), P. e5292.

13. Gao M., Ma L., Liu H., Zhang Z., Ning Z., Xu J. Malicious network traffic detection based on deep neural networks and association analysis. *Sensors*. 2020, Vol. 20(5), P. 1452.

14. Vinayakumar R., Soman K. P., Poornachandran P. Evaluation of recurrent neural network and its variants for intrusion detection system (IDS). *International Journal of Information System Modeling and Design (IJISMD)*. 2017, Vol. 8(3), P. 43–63.

15. Wei Y. Z., Md-Arshad M., Samad A. A., Ithnin N. Comparing Malware Attack Detection using Machine Learning Techniques in IoT Network Traffic. *International Journal of Innovative Computing*. 2023, Vol. 13(1), P. 21–27.

16. Nie L., Jiang D., Lv Z. Modeling network traffic for traffic matrix estimation and anomaly detection based on Bayesian network in cloud computing networks. *Annals of Telecommunications*. 2017, Vol. 72, P. 297–305.

17. Landoll D. The security risk assessment handbook: A complete guide for performing security risk assessments. *CRC Press*, 2021.

18. Macek D., Magdalenic I., Redep N. B. A systematic literature review on the application of multicriteria decision making methods for information security risk assessment. *International Journal of Safety and Security Engineering*. 2020, Vol. 10, No. 2, P. 161–174.

19. Jouini M., Rabai L. B. A. Comparative study of information security risk assessment models for cloud computing systems. *Procedia Computer Science*. 2016, 83, P. 1084–1089.

20. Haji S., Tan Q., Costa R.S. A hybrid model for information security risk assessment. *Int. j. adv. trends comput. sci. eng.* 2019, ART-2019-111611.

21. Hu L., Li H., Wei Z., Dong S., Zhang Z. Summary of research on IT network and industrial control network security assessment. *2019 IEEE 3rd information technology, networking, electronic and automation control conference (ITNEC)*. 2019, P. 1203–1210.

22. Cherdantseva Y., Burnap P., Blyth A., Eden P., Jones K., Soulsby H., Stoddart K. A review of cyber security risk assessment methods for SCADA systems. *Computers & security*. 2016, Vol. 56, P. 1–27.

23. Mahak M., Singh Y. Threat modelling and risk assessment in internet of things: A review. *Proceedings of Second International Conference on Computing, Communications, and Cyber-Security: IC4S 2020*. 2021, Springer Singapore, P. 293–305.

24. Lyu X., Ding Y., Yang S. H. Safety and security risk assessment in cyber-physical systems. *IET Cyber-Physical Systems: Theory & Applications*. 2019, Vol. 4(3), P. 221–232.

25. Isaev S. V., Kononov D. D. A study of dynamics and classification of attacks on corporate network web services. *Siberian Aerospace Journal*. 2022, Vol. 23, No. 4, P. 593–601.

Кононов Дмитрий Дмитриевич – научный сотрудник; Институт вычислительного моделирования СО РАН. E-mail: ddk@icm.krasn.ru.

Исаев Сергей Владиславович – кандидат технических наук, доцент, заведующий отделом информационно-телекоммуникационных технологий; Институт вычислительного моделирования СО РАН. E-mail: si@icm.krasn.ru.

Kononov Dmitry Dmitrievich – scientific researcher; Institute of Computational Modelling SB RAS. E-mail: ddk@icm.krasn.ru.

Isaev Sergey Vladislavovich – Cand. Sc., associate professor, head of the Department of Information and Telecommunication Technologies; Institute of Computational Modelling SB RAS. E-mail: si@icm.krasn.ru.