UDC 004.738
Doi: 10.31772/2712-8970-2025-26-1-8-20

Для цитирования: Исаева О. С., Кулясов Н. В., Исаев С. В. Инфраструктура сбора данных и имитации угроз безопасности сети интернета вещей // Сибирский аэрокосмический журнал. 2025. Т. 26, № 1. С. 8–20. Doi: 10.31772/2712-8970-2025-26-1-8-20.

For citation: Isaeva O. S., Kulyasov N. V., Isaev S. V. [Infrastructure for collecting data and simulating security threats in the internet of things network]. *Siberian Aerospace Journal.* 2025, Vol. 26, No. 1, P. 8–20. Doi: 10.31772/2712-8970-2025-26-1-8-20.

Инфраструктура сбора данных и имитации угроз безопасности сети интернета вещей

О. С. Исаева*, Н. В. Кулясов, С. В. Исаев

Институт вычислительного моделирования СО РАН –
обособленное подразделение ФИЦ КНЦ СО РАН
Российская Федерация, 660036, г. Красноярск, ул. Академгородок, 50/44
*E-mail: isaeva@icm.krasn.ru

Аннотация. Внедрение технологии интернета вещей (internet of things, IoT) на предприятиях ракетно-космической отрасли требует обеспечения повышенных мер безопасности информационно-коммуникационных процессов. Существующие системы обнаружения вторжений не способны учитывать гетерогенность структуры сети и масштабность циркулирующей между устройствами информации. Для решения этой проблемы системы обнаружения вторжений используют метод аномалий, для применения которого требуется большое число репрезентативных данных. Авторами выполнен обзор публичных наборов данных, на основе которых может быть построена система выявления аномалий. Они содержат информацию из искусственных имитационных сред или изолированных окружений с имитацией устройств, включают примеры, которые напрямую не связаны с интернетом вещей, и не учитывают динамический характер изменения трафика.

В данной статье мы представляем новую инфраструктуру, которая позволит избежать указанных недостатков. Она собирает данные функционирования реальной сети интернета вещей и позволяет выполнять её тестирование на устойчивость к характерным атакам. Мы используем прикладной протокол MQTT (message queuing telemetry transport) и программные платформы, поддерживающие информационное взаимодействие на основе шаблона «издатель – подписчик». Инфраструктура содержит устройства, осуществляющие мониторинг технологических помещений с телекоммуникационным оборудованием, сервера с различными настройками политик безопасности, приложения для контроля и анализа данных, программные агенты сбора сетевого трафика и имитаторы угроз, выполняющие атаки на узлы сети с одиночных источников или в распределённой среде. Исследователи смогут, применяя собираемые в инфраструктуре данные для анализа кибербезопасности, создавать надёжные решения на базе интернета вещей, необходимые для внедрения этой технологии в наукоёмкие производства космических систем.

Ключевые слова: кибербезопасность, интернет вещей, протокол MQTT, брокер данных, базы данных вторжений, имитация угроз безопасности.

Infrastructure for collecting data and simulating security threats in the internet of things network

O. S. Isaeva*, N. V. Kulyasov, S. V. Isaev

Institute of Computational Modelling of the Siberian Branch of the SB RAS – subdivision Federal Research Center "Krasnoyarsk Scientific Center of the SB RAS"
50/44, Akademgorodok St., Krasnoyarsk, 660036, Russian Federation
*E-mail: isaeva@icm.krasn.ru

Abstract. The implementation of the internet of things technologies in the rocket-space industry requires increased security measures for information and communication processes. Existing intrusion detection systems are unable to take into account the heterogeneity of the network structure and the scale of information circulating between devices. To solve this problem, intrusion detection systems use an anomaly method, which requires a large number of representative data sets. The authors have reviewed public datasets that can be used to build an anomaly detection system. They contain information from artificial simulation medium or isolated environments with simulated devices, include examples that are not directly related to the internet of things, and do not take into account the dynamic nature of traffic changes.

In this paper, we present a new infrastructure that will avoid these drawbacks. It collects data on the functioning of a real Internet of Things network and allows testing its stability to typical attacks. We use the MQTT (message queuing telemetry transport) application protocol and software platforms that support information interaction based on the publisher-subscriber pattern. The infrastructure contains devices that monitor technological rooms with telecommunications equipment, brokers with various security policy settings, applications for data control and analysis, software agents for collecting network traffic and threat simulators that perform attacks on network nodes from single sources or in a distributed environment. Researchers will be able to use the data collected in the infrastructure for cybersecurity analysis to create reliable IoT-based solutions needed to implement this technology in knowledge-intensive space systems production.

Keywords: cybersecurity, internet of things, protocol MQTT, data broker, intrusion databases, simulated security threats.

Introduction

Innovations defined by the concept of the Internet of Things (IoT), according to the recommendations of the International Telecommunication Union [1], reflect modern trends aimed at building infrastructures that unite physical and virtual objects based on functionally compatible information and communication technologies. The need for enterprises in the rocket and space industry for such solutions is determined not only by modern requirements for the digitalization of production, but also by the need for their prompt transformation, meeting the needs of consumers of products, as well as resource and technological capabilities [2]. The introduction of digital technologies allows for continuous quality control of manufactured products, identifying operational risks, predicting failures of technical systems, ensuring increased efficiency of production processes [3]. However, the undeniable advantages of using Internet of Things technologies at enterprises in the rocket and space industry are offset by the need to ensure increased security and reliability of all information and communication processes.

IoT networks are of interest to cybercriminals due to the nature of the devices, the ease of their connection protocols, and the value and scale of the information that can be obtained by accessing them. The main goal of cyber attacks is to distort the generated data that determines the actions of users or automated systems, disrupt current processes (denial of service), and disclose information that can be used to gain competitive advantages [4]. Compromise of any single device can spread to all enterprise systems and disrupt its critical functions [5]. Reports of hacks of IoT devices that allow remote control interception and penetration into corporate networks, or the unification of such devices into botnets appear regularly [6].

Intrusion detection systems, according to research methods, are divided into signature-based (allow identifying attacks that exploit network vulnerabilities), rule-based (identify actions that are inconsistent with legitimate users), and anomaly-based (use machine learning methods to detect atypical behavior or statistical discrepancies) [7]. Signature-based and rule-based security solutions are not designed to support IoT networks, which are characterized by a heterogeneous structure, limited computing power of interconnected devices, multi-platform connection protocols used, large amounts of network traffic, heterogeneity of security events, and a lack of accurate data on attack behavior [8]. For the IoT, an anomaly-based approach is preferable. However, its use requires the creation of representative datasets and reliable assessment methods that take into account the properties of real networks.

Machine learning techniques can be used to generate patterns from malicious network traffic captured during an infection, which can be used to detect similar attacks in the future [9]. Unlike other areas where machine learning is widely used, the field of intrusion detection assumes that the characteristics of open-world traffic are constantly changing in terms of its content, service methods, attack scenarios, and their results, which change depending on the development of evasion tools [10]. As a result, the behavior of network traffic demonstrated in a training dataset cannot be expected to correspond to its behavior in a production environment to any extent over time [11]. Anomaly-based systems using machine learning techniques must have a significant number of real-world network traffic examples with all types of attacks and common user behaviors and payloads, which are relevant and cover long observation periods [12].

In this paper, we present a new IoT data collection infrastructure designed to accumulate information about network traffic and simulate security threats. The study was conducted in the corporate network of the Krasnoyarsk Scientific Center of the Siberian Branch of the Russian Academy of Sciences(SB RAS), which includes IoT devices and software that collects and analyzes data [13]. To identify parameters characterizing the state of the network and the processes occurring in it, the ontology constructed in [14] is used. A concept for organizing a data collection infrastructure is proposed that will allow identifying the features of the IoT operating environment and taking into account the dynamic nature of security events, which will expand the scope of application of machine learning methods in the intrusion detection system.

Overview of public data sets

To develop and test solutions for information security and network interaction of IoT devices, public data sets are created that describe various cyberspace infrastructures and scenarios that violate their reliability and performance. In order to determine the functions of the created infrastructure, the composition of indicators for security analysis, popular datasets widely used in the construction of intrusion detection systems were considered.

The KDD99 [15] and NSL-KDD [16] datasets are among the most frequently cited, containing normal and aggressive traffic scenarios obtained from a testbed set up at the MIT lab. Researchers using these datasets have encountered redundancy and insufficient balancing of training samples. These datasets are now outdated and do not reflect the dynamics of modern network systems.

Another widely used dataset is the CICIDS2017 Intrusion Detection Evaluation Dataset [17], developed by the Canadian Cybersecurity Institute (CIC), which represents network traffic generated in an isolated environment by simulating the actions of legitimate users and intruders. The set contains more than 50 GB of raw data and files with labeled sessions and extracted features on different observation days. The popularity of this set has allowed researchers to obtain comprehensive data on the errors it contains. In [18], the features of network sessions from the source files were analyzed and compared with the labeled data. Significant discrepancies in session allocation, errors in calculating feature values and their duplication, incorrect termination of sessions, assignment of boundary packets to the next session, incorrect calculation of packet lengths, which can have a significant impact on the results of machine learning, were revealed. Our experience in applying the CIC Flow Meter packet processing tool [19], on which CICIDS2017 is built, to our own raw data confirmed the above issues. In NTLF low Lyzer [20], these issues are fixed.

The listed data sets do not take into account the specifics of the networks and protocols of the Internet of Things. In our study, to identify anomalies, we combine the characteristics of the transport layer traffic with indicators describing the application protocols of the IoT. To identify them, we considered data sets obtained within the framework of the implementation of such networks.

The TON IoT 2021 comprehensive data set [21] includes information from IoT systems collected as log files from 10 telemetry sensors. Preprocessing of raw data was performed using the ZEEK (Bro) tool. For the Linux operating system, data was collected using its tracing tool, which monitors processor, memory, and network activity. For Windows, the operating system performance data collector was used. The set includes 9 different security threats.

The UQ-IOT-IDS-2021 dataset [22] is obtained from a real-world heterogeneous IoT network environment that includes various devices such as smartphones, smart TVs, IP cameras, smart speakers, etc.

The MQTT-IoT-IDS-2020 set [23] contains traffic of a simulated network consisting of 12 sensors and a broker interacting with each other via the MQTT (Message Queuing Telemetry Transport) protocol and threat sources (4 types). The data includes protocol flags describing the session state and the level of service quality. In our study, we included similar flags and indicators that summarize the characteristics of MQTT protocol sessions.

CIC EVSE2024 [24] contains benign traffic, network attacks, and attacks targeting an EV charger. The anomalous traffic represents data collection (Reconnaissance) and denial of service (DoS) attacks, targeted attacks such as Backdoor and Cryptojacking. These datasets demonstrate the potential of using payloads to detect anomalous behavior of IoT devices.

The peculiarities of public datasets consist of the heterogeneity of the provided data, differences in the methods of simulating security events and approaches to collecting and pre-processing information, which can have a huge impact on the efficiency of anomaly detection. The purpose, functions and type of devices determine the permissible limits of measurements received from them as a payload. To detect anomalies in traffic, it is necessary to study the results of the operation of those devices whose security we plan to ensure. Most datasets include not only network traffic associated with the IoT, but also data from third-party services, systems and protocols, which complicates their use for localizing threats aimed at the operation of IoT devices. In addition, almost all datasets are obtained in an artificial simulation environment or isolated environments with device imitation. They consider threats that are typical not only for IoT systems, but also for any other information systems, thereby overloading data sources for machine learning with unnecessary examples not found in the networks under study. In such conditions, it is justified to develop our own infrastructure for collecting and accumulating data, taking into account the existing experience of creating public data and allowing us to update the data when the network operating conditions change.

Corporate IoT structure

In the Krasnoyarsk Scientific Center, the IoT technology is used to monitor the temperature and humidity of the premises where the servers supporting the network and mail services are located [25]. The network structure is shown in Fig. 1.

When constructing data sets for anomaly analysis, it is necessary to take into account not only the network structure, but also the properties of the protocols used at individual machine-to-machine communication levels. Table 1 provides examples of protocols distributed across network levels (in the context of the TCP/IP protocol stack and the OSI reference network model).

When building an IoT network, it is necessary to choose which protocols will operate at the data link layer and which will perform data exchange between devices and applications at the application layer. The requirements for these protocols include ensuring reliable and efficient communication between distributed devices, real-time data synchronization, and the ability to transmit asynchronously over unstable connections and in low network bandwidth conditions [26].

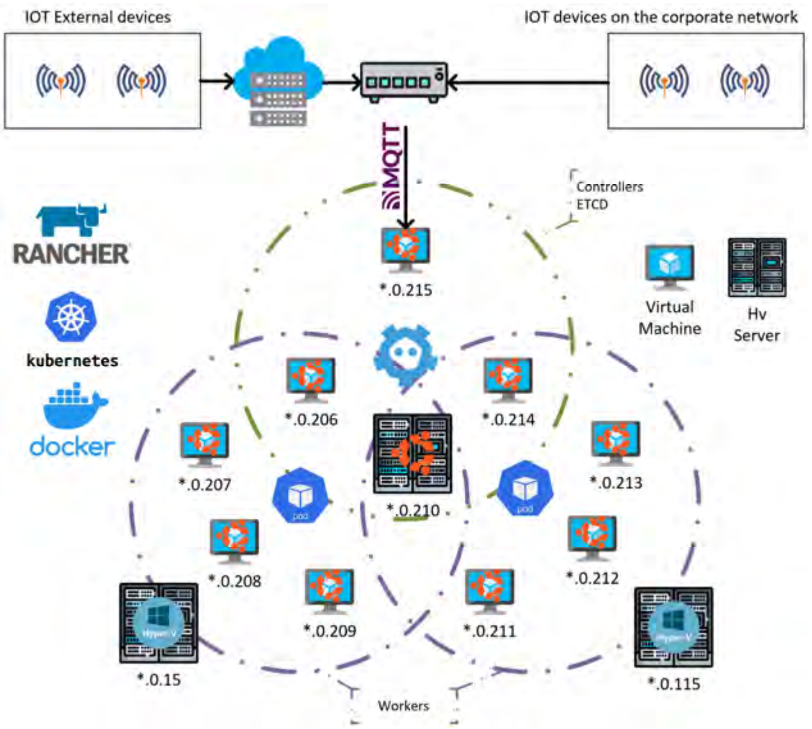

Рис. 1. Структура сети интернета вещей

Fig. 1. Structure of the internet of things network

For example, Ethernet, Modbus, Zigbee, WiFi, LoRaWAN are channel layer protocols. In our case, Ethernet and WiFi protocols are used, which is determined by the functionality of the corporate network into which the Internet of Things devices are built. Application layer protocols perform the functions of packaging, formatting and delivering data, ensuring control of their integrity and quality of service.

The application layer protocols include MQTT, CoAP, AMQP. In the constructed network, the MQTT protocol was selected for the application layer – an open standard developed specifically for small computing capabilities of devices [27] and is currently one of the most frequently used protocols for machine-to-machine interaction in the Internet of Things. Its operation is based on the “publisher – subscriber” pattern (Fig. 2).

Table 1

Distribution of network protocols by layers of the OSI and TCP/IP models

OSI model	*TCP/IP*	*Examples*
Applied	Applied	HTTP, HTTPS, FTP, MQTT, CoAP, AMQP
Representative		
Session		
Transport	Transport	TCP, UDP
Network	Inter-network	IPv4, IPv6, ICMP
Channel	Channel	Ethernet, Wi-Fi, BLE, Zigbee, Z-Wave LoRaWan
Physical		

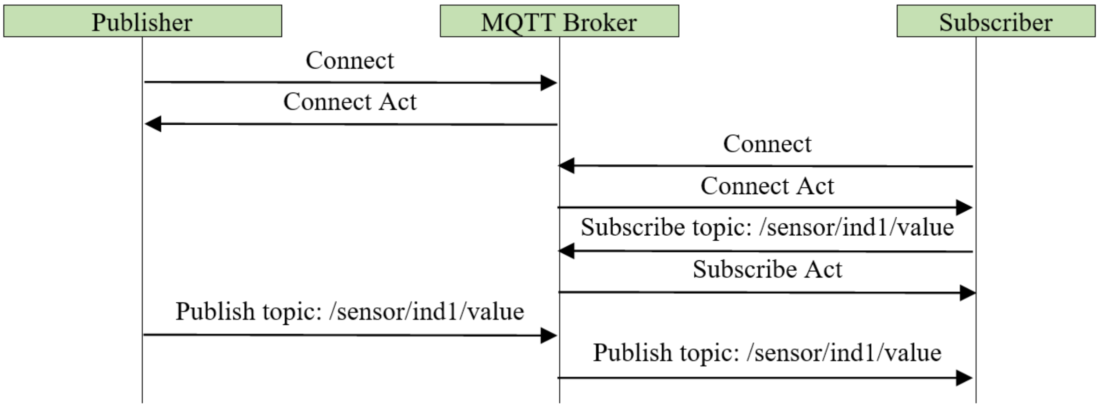
Publisher
MQTT Broker
Subscriber
Connect
Connect Act
Connect
Connect Act
Subscribe topic: /sensor/ind1/value
Subscribe Act
Publish topic: /sensor/ind1/value
Publish topic: /sensor/ind1/value

Рис. 2. Шаблон «издатель – подписчик»

Fig. 2. Template “publish – subscribe”

The following entities participate in data exchange: Publisher is a device that collects information or performs measurements; Subscriber is a client that receives and uses this information in the course of its work; Broker is an intermediary that receives data from the publisher and distributes it by subscription to topics that determine the type and source of information. The mechanisms of information interaction in the developed Internet of Things network take into account the cyclical nature of the analyzed processes, which allows reducing the load on information consumers and the networks used [28].

Despite the fact that the connection of IoT devices to the corporate network has not been announced in the public space, there is a steady increase in illegitimate connection requests. There is a constant scanning of various services on the ports of the MQTT protocol.

Fig. 3 shows an example of the distribution of requests by source countries. Indicators characterizing the dynamics of growth of requests to devices of the corporate IoT network are considered in [29].

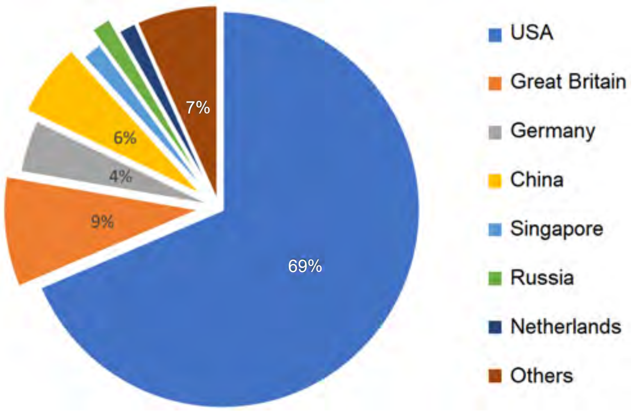

Рис. 3. Распределение запросов по странам (статистика за 6 месяцев)

Fig. 3. Distribution of requests by country (6 months statistics)

The growing interest in IoT devices from illegitimate users and scanning of MQTT protocol network ports show the relevance of monitoring the security of the constructed structure and creating tools for detecting network anomalies. An infrastructure has been developed for the study, ensuring the collection and updating of data.

Building a data collection infrastructure

The infrastructure for collecting data and simulating security threats is placed within the existing network of the organization, without additional optimization and isolation. The previously shown structure of the IoT network is supplemented by various implementations of data brokers, software agents that collect information at key points of the corporate network, and threat simulators specific to the IoT.

The choice of a software platform for implementing brokers is determined by the requirements for ensuring performance, reducing data transfer delays, supporting clustering, limiting resource consumption, etc. As a rule, public data sets do not focus on the brokers used and their configuration, which does not allow identifying possible features of the functioning of the software platforms implementing their functions. To take into account the features of various brokers, EMQX, NanoMQ, VerneMQ brokers [30] are installed in addition to Eclipse Mosquitto. Broker configurations are configured that differ in the software platforms used and settings for the authorization method, encryption (TLS – Transport Layer Security protocol or without encryption) and access (from an internal or external network). Examples of broker symbols are given in Table 2.

Table 2

Configuring security settings

Designation on the diagram	Authorization	Encryption
auth_priv	Login/Password	No encryption
anon_priv	Without authorization	No encryption
auth_priv_tls	Login/Password	TLS
anon_priv_tls	Without authorization	TLS

The conceptual diagram of the data collection and security threat simulation infrastructure is shown in Figure 4. The infrastructure includes 16 independent brokers and a replicating broker for distributing data from publishers, as well as software agents that collect information about the interaction of IoT devices with the outside world and within the network.

Software agents running on brokers collect both external and internal network traffic coming to standard and encrypted ports of the MQTT protocol. In addition, server metrics and indicators provided by the broker software are collected. From servers, we receive percentage characteristics of system resources (processor, memory, network channel load) required to perform work, the number of input/output operations performed, and others. From brokers, we receive the number of active clients, received and sent messages, their volume, static indicators obtained after analyzing network traffic. Data analysis will allow you to select a broker and its parameters in such a way as to maintain a balance between service availability and information integrity under the same usage scenarios without loss of functionality and under conditions of limited resource consumption.

Network traffic collection agents are also located on the Proxy server (a device that services information flows between users and web resources). The data collected there contains information about all external threat sources.

To collect data that occurs during network attacks on Internet of Things resources, security threat simulators are configured (the eMQTT-Bench performance testing tool is used). Currently, traffic is collected with the following threats simulated:

– DOS con, a large number of requests to the broker for connection are made in a short period of time;

– DOS sub, a large number of requests to the broker for subscriptions are made in a short period of time;

– DOS pub, a large number of requests to the broker to publish messages in a short period of time are made.

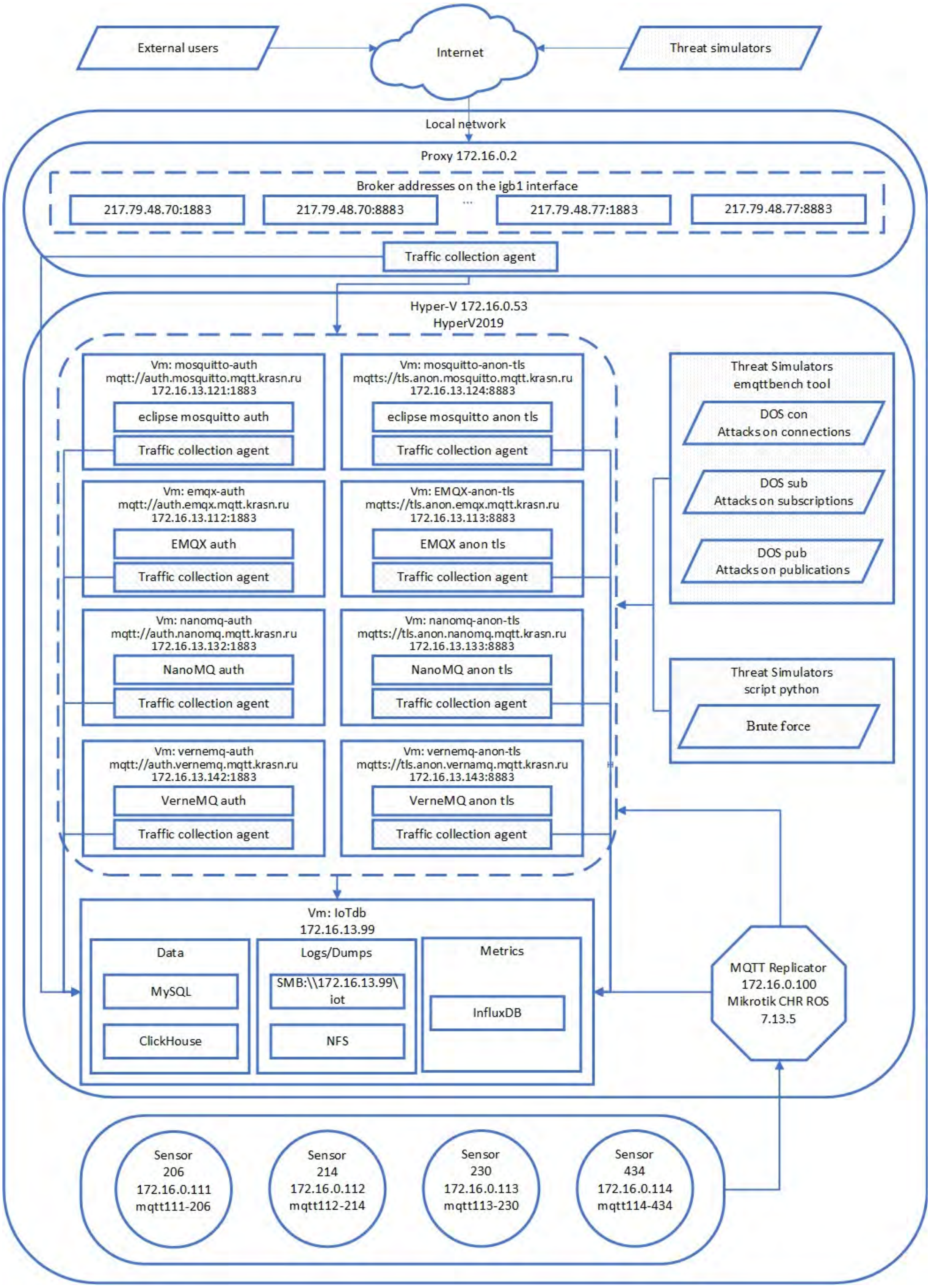

Рис. 4. Схема инфраструктуры сбора данных

Fig. 4. Scheme of the data collection infrastructure

The listed attacks are performed both from a single source and simulate distributed threat variants from many sources. Additionally, a password attack is implemented, which performs a login/password pair selection for brokers with configured authentication. The data collected by agents is largely unstructured and their use requires preprocessing, which divides traffic into sessions and calculates the parameters of the received sessions.

The selection of the composition of the analyzed indicators is based on the conducted review of popular data sets. We combined indicators similar to those used in public sources, characterizing different levels of the network. By analogy with CICIDS2017, we calculate the characteristics of sessions: date - time, flow identifier, source and receiver IP, protocol, flow duration, its speed, average value of the interpacket interval, total length of packets transmitted in the forward and reverse directions, and others. We supplement them with indicators generated in CIC EVSE2024 and not included in the first data source, for example, the number of SYN flags (ACK, FIN, etc.) in the forward or reverse direction. To study the features of the MQTT protocol, we calculate characteristics that summarize the flags of its sessions, similar to those used in the MQTT-IoT-IDS-2020 set, for example, the flag of the current session, connection, hold, session clear, quality of service level (requested and provided), message type, its length, etc. Thus, we use the structural elements of public data sources as dictionaries, which ensures the comparability of our metadata with them and allows us to compare the results of machine learning methods configured on our data with other works using public sets.

The built infrastructure allows collecting IoT traffic and testing the network for resistance to several types of attacks typical for such networks. Simulating attacks using tools included in the data broker will ensure their compliance with accepted standards.

Conclusion

The paper describes the infrastructure for collecting network activity data and simulating threats to the security of the IoT deployed within the corporate network of the Krasnoyarsk Scientific Center. It includes: sensors measuring temperature, humidity, etc.; brokers operating via the MQTT protocol; subscribers located in the local network or accessing data via the Internet; agents collecting information; threat simulators that allow attacks to be carried out that are typical for the IoT and the network protocols used.

Information is collected in a distributed environment that takes into account the specifics of the protocols used for interaction with all devices in various parts of the network. The collected data is processed and transferred to information systems within the network in an orderly manner, and access is provided to users located outside the local network.

The infrastructure allows for the prompt receipt of new data sets with updated attack scenarios, taking into account the impact of the drift of attack concepts. Its use allows for the formation of a data set for the creation and verification of new methods and tools for detecting information security threats aimed at working in real IoT networks. Further development of the research topic consists in building a system for identifying and blocking cyber threats based on the analysis of network anomalies.

Библиографические ссылки

1. Recommendation ITU-T Y.2060. Series Y: Global information infrastructure, internet protocol aspects and next-generation networks. Next generation networks – Frameworks and functional architecture models. Overview of the Internet of things – Switzerland, Geneva: International telecommunication union, 2013. 22 p. [Электронный ресурс]. URL: https://handle.itu.int/11.1002/ 1000/11559 (дата обращения: 10.01.2025).

2. Абрашкин М. С., Афанасьев В. Я., Бускин Н. С. Цифровизация предприятий ракетно-космической промышленности в условиях новой промышленной революции // Russian journal of management. 2024. № 12(2). С. 369–389.

3. Internet of things: Vision, applications and research challenges / D. Miorandi, S. Sicari, F. Pellegrini, I. Chlamtac // Ad Hoc Networks, 2012. Vol. 10(7). P. 1497–1516.

4. A survey of network-based intrusion detection data sets / M. Ring, S. Wunderlich, D. Scheuring, D. Landes, A. Hotho // Computers & Security, 2019. Vol. 86. P. 147–167.

5. Al-Hawawreh M., Sitnikova E., Aboutorab N. X-IIoTID: A Connectivity- and Device-agnostic Intrusion Dataset for Industrial Internet of Things // IEEE Internet of Things Journal. 2021. No. 99. P. 1-1. DOI: 10.1109/JIOT.2021.3102056.

6. Шмелев Я., Моргунов В. Обзор угроз для IoT-устройств в 2023 году [Электронный ресурс]. URL: https://securelist.ru/iot-threat-report-2023/108088/ (дата обращения: 30.09.2024).

7. Jabez J., Muthukumar B. Intrusion Detection System (IDS): Anomaly Detection Using Outlier Detection Approach // Procedia Computer Science, 2015. Vol. 48. P. 338–346.

8. Security, privacy and trust in Internet of Things: The road ahead / S. Sicari, A. Rizzardi, L.A. Grieco, A. Coen-Porisini // Computer Networks, 2015. Vol. 76. P. 146–164.

9. Botnet attacks detection in IoT environment using machine learning techniques / M. AL-Akhrasa, A. Alshunaybirb, H. Omarc, S. Alhazmib // International Journal of Data and Network Science, 2023. Vol. 7. P. 1683–1706.

10. Paxson V., Floyd S. Wide area traffic: the failure of Poisson modeling // IEEE/ACM Transactions on Networking, 1995. Vol. 3(3). P. 226–244.

11. Viegas E. K., Santin A. O., Oliveira L. S. Toward a reliable anomaly-based intrusion detection in real-world environments // Computer Networks, 2017. Vol. 127. P. 200–216.

12. Generating realistic intrusion detection system dataset based on fuzzy qualitative modeling / W. Haider, J. Hu, J. Slay, B. P. Turnbull, Y. Xie // Journal of Network and Computer Applications, 2017. Vol. 87. P. 185–192.

13. Исаева О. С. Построение цифрового профиля устройств Интернета вещей // Информационные и математические технологии в науке и управлении, 2023. № 2(30). С. 36–44.

14. Исаева О. С. Построение онтологии для систематизации характеристик сети Интернета вещей // Онтология проектирования, 2024. Т. 14, № 2(52). С. 243–255.

15. KDD Cup 1999 Data [Электронный ресурс]. URL: https://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html (дата обращения: 20.01.2025).

16. Mohi-ud-din G. NSL-KDD // IEEE Dataport, 2018. doi: 10.21227/425a-3e55.

17. Xinpeng C. CICIDS2017 and UNBSW-NB15 // IEEE Dataport, 2023. DOI: 10.21227/ykpn-jx78.

18. 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. P. 102994.

19. Lashkari A. H. CICFlowmeter-V4.0 (formerly known as ISCXFlowMeter) is a network traffic Bi-flow generator and analyser for anomaly detection. [Электронный ресурс]. URL: https://github.com/ISCX/CICFlowMeter.10.13140/RG.2.2.13827.20003 (дата обращения: 20.01.2025).

20. Методика сбора обучающего набора данных для модели обнаружения компьютерных атак / А. И. Гетьман, М. Н. Горюнов, А. Г. Мацкевич, Д. А. Рыболовлев // Труды ИСП РАН, 2021. № 33(5). С. 83–104.

21. Shafi M. M., Lashkari A. H., Roudsari A. H. NLFlowLyzer: Toward generating an intrusion detection dataset and intruders behavior profiling through network layer traffic analysis and pattern extraction // Computers & Security. 2024. Vol. 148, No. 1. P. 104160. DOI:10.1016/j.cose.2024.104160.

22. UQ IoT IDS dataset / H. Ke, K. Dan, Z. Zhien, G. Mengmeng, L. Ulysses, Y. Jiaqi // The University of Queensland. Data Collection, 2022.

23. MQTT-IoT-IDS2020: MQTT Internet of Things Intrusion Detection Dataset / H. Hindy, C. Tachtatzis, R. Atkinson, E. Bayne, X. Bellekens // IEEE Dataport, 2020.

24. CIC EV charger attack dataset 2024 (CICEVSE2024) [Электронный ресурс]. URL: https://www.unb.ca/cic/datasets/evse-dataset-2024.html (дата обращения: 15.10.2024).

25. Isaeva O. S., Kulyasov N. V., Isaev S. V. Creation of a simulation stand for studying of the internet of things' technologies // AIP Conference Proceedings, 2022. № 2647. P. 040030-1–040030-5.

26. A survey of intrusion detection in Internet of Things / B. B. Zarpelão, R. S. Miani, C. T. Kawakani, S. C. de Alvarenga // Journal of Network and Computer Applications, 2017. Vol. 84. P. 25–37.

27. MQTT: The Standard for IoT Messaging [Электронный ресурс]. URL: https://mqtt.org/ (дата обращения: 14.08.2024).

28. Исаева О. С., Исаев С. В., Кулясов Н. В. Формирование адаптивных рассылок брокера данных интернета вещей // Информационно-управляющие системы, 2022. № 5(120). С. 23–31.

29. Исаева О. С., Кулясов Н. В., Исаев С. В. Создание инструментов сбора данных для анализа аспектов безопасности Интернета вещей // Информационные и математические технологии в науке и управлении, 2022. № 3(27). С.113–125.

30. A scalable and low-cost MQTT broker clustering system / P. Jutadhamakorn, T. Pillavas, V. Visoottiviseth, R. Takano, J. Haga, D. Kobayashi // 2nd International Conference on Information Technology (November 2017, Thailand). IEEE. 2017. P. 1–5.

References

1. Recommendation ITU-T Y.2060. Series Y: Global information infrastructure, internet protocol aspects and next-generation networks. Next generation networks – Frameworks and functional architecture models. Overview of the Internet of things. *International telecommunication union*. 2013. 22 p. Available at: https://handle.itu.int/11.1002/1000/11559 (accessed: 10.01.2025).

2. Abrashkin M. S., Afanas'ev V. Ya., Buskin N. S. [Digitalization of rocket-space industry enterprises in the context of the new industrial revolution]. *Russian journal of management.* 2024, No. 12(2), P. 369–389 (In Russ.).

3. Miorandi D., Sicari S., Pellegrini F., Chlamtac I. Internet of things: Vision, applications and research challenges. *Ad Hoc Networks*, 2012, Vol. 10(7), P. 1497–1516.

4. Ring M., Wunderlich S., Scheuring D., Landes D., Hotho A. A survey of network-based intrusion detection data sets. *Computers & Security*, 2019, Vol. 86, P. 147–167.

5. Al-Hawawreh M., Sitnikova E., Aboutorab N. X-IIoTID: A connectivity and device agnostic intrusion dataset for industrial Internet of Things. *IEEE Internet of Things Journal*. 2021, No. 99, P. 1-1. DOI:10.1109/JIOT.2021.3102056.

6. Shmelev Ya., Morgunov V. [IoT threat landscape in 2023]. (In Russ.). Available at: https://securelist.ru/iot-threat-report-2023/108088/ (accessed 30.09.2024).

7. Jabez J., Muthukumar B. Intrusion detection system (IDS): anomaly detection using outlier detection approach. *Procedia Computer Science*, 2015, Vol. 48, P. 338–346.

8. Sicari S., Rizzardi A., Grieco L.A., Coen-Porisini A. Security, privacy and trust in Internet of Things: the road ahead. *Computer Networks*, 2015, Vol. 76, P. 146–164.

9. AL-Akhrasa M., Alshunaybirb A., Omarc H., Alhazmib S. Botnet attacks detection in IoT environment using machine learning techniques. *International Journal of Data and Network Science*, 2023, Vol. 7, P. 1683–1706.

10. Paxson V., Floyd S. Wide area traffic: the failure of Poisson modeling. *IEEE ACM Transactions on Networking*, 1995, Vol. 3(3), P. 226–244.

11. Viegas E. K., Santin A. O., Oliveira L. S. Toward a reliable anomaly-based intrusion detection in real-world environments. *Computer Networks*, 2017, Vol. 127, P. 200–216.

12. Haider W., Hu J., Slay J., Turnbull B. P., Xie Y. Generating realistic intrusion detection system dataset based on fuzzy qualitative modeling. *Journal of Network and Computer Applications*, 2017, Vol. 87, P. 185–192.

13. Isaeva O. S. [Building a digital profile of IoT devices]. *Informatsionnye i matematicheskie tekhnologii v nauke i upravlenii*, 2023, No. 2(30), P. 36–44 (In Russ.).

14. Isaeva O. S. [Building an ontology to systematize the characteristics of the Internet of Things network], 2024, Vol. 14, No. 2(52), P. 243–255 (In Russ.).

15. KDD Cup 1999 Data. Available at: https://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html (accessed 20.01.2025).

16. Mohi-ud-din G. NSL-KDD. *IEEE Dataport*, 2018. doi: 10.21227/425a-3e55.

17. Xinpeng C. CICIDS2017 and UNBSW-NB15. *IEEE Dataport*, 2023. doi: 10.21227/ykpn-jx78.

18. 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, P. 102994.

19. Lashkari A. H. CICFlowmeter-V4.0 (formerly known as ISCXFlowMeter) is a network traffic Bi-flow generator and analyser for anomaly detection. Available at: https://github.com/ISCX/CICFlowMeter.10.13140/RG.2.2.13827.20003 (accessed 20.01.2025).

20. Getman A. I., Goryunov M. N., Matskevich A. G., Rybolovlev D. A. [Methodology for collecting training dataset for computer intrusion detection model]. *Trudy ISP RAN*, 2021, No. 33(5), P. 83–104 (In Russ.).

21. Shafi M. M., Lashkari A. H., Roudsari A. H. NLFlowLyzer: Toward generating an intrusion detection dataset and intruders behavior profiling through network layer traffic analysis and pattern extraction. *Computers & Security*. 2024, Vol. 148, No. 1, P. 104160. DOI:10.1016/j.cose.2024.104160.

22. Ke H., Dan K., Zhien Z., Mengmeng G., Ulysses L., Jiaqi Y. UQ IoT IDS dataset. *The University of Queensland. Data Collection*, 2022.

23. Hindy H., Tachtatzis C., Atkinson R., Bayne E., Bellekens X. MQTT-IoT-IDS2020: MQTT Internet of Things intrusion detection dataset. *IEEE Dataport*, 2020.

24. CIC EV charger attack dataset 2024 (CICEVSE2024) Available at: https://www.unb.ca/cic/ datasets/evse-dataset-2024.html (accessed 15.10.2024).

25. Isaeva O. S., Kulyasov N. V., Isaev S. V. Creation of a simulation stand for studying of the internet of things' technologies. *AIP Conference Proceedings*, 2022, No. 2647, P. 040030-1–040030-5.

26. Zarpelão B. B., Miani R. S., Kawakani C. T., Alvarenga S. C. A survey of intrusion detection in Internet of Things. *Journal of Network and Computer Applications*, 2017, Vol. 84, P. 25–37.

27. MQTT: The Standard for IoT Messaging [Электронный ресурс]. URL: https://mqtt.org/ (дата обращения: 14.08.2024).

28. Isaeva O. S., Isaev S. V., Kulyasov N. V. [Formation of adaptive publications from the Internet of Things data broker]. *Informatsionno-upravliaiushchie sistemy*, 2022, No. 5(120), P. 23–31 (In Russ.).

29. Isaeva O. S., Kulyasov N. V., Isaev S. V. [Creating data collection tools to analyze security aspects Internet of Things]. *Informatsionnye i matematicheskie tekhnologii v nauke i upravlenii*, 2022. No. 3(27). P. 113–125 (In Russ.).

30. Jutadhamakorn P., Pillavas T., Visoottiviseth V., Takano R., Haga J., Kobayashi D. A scalable and low-cost MQTT broker clustering system. *2nd International Conference on Information Technology,* IEEE, 2017, P. 1–5.

Исаева Ольга Сергеевна – доктор технических наук, старший научный сотрудник; Институт вычислительного моделирования СО РАН – обособленное подразделение ФИЦ КНЦ СО РАН. E-mail: isaeva@icm.krasn.ru. https://orcid.org/0000-0002-5061-6765.

Кулясов Никита Владимирович – программист; Институт вычислительного моделирования СО РАН – обособленное подразделение ФИЦ КНЦ СО РАН. E-mail: razor@icm.krasn.ru.

Исаев Сергей Владиславович – кандидат технических наук, доцент, заместитель директора по научной работе; Институт вычислительного моделирования СО РАН – обособленное подразделение ФИЦ КНЦ СО РАН. E-mail: si@icm.krasn.ru. https://orcid.org/0000-0002-6678-0084.

Isaeva Olga Sergeevna – Doct. Sc., Senior Researcher; Institute of Computational Modelling of the Siberian Branch of the SB RAS – subdivision Federal Research Center "Krasnoyarsk Scientific Center of the SB RAS". E-mail: isaeva@icm.krasn.ru. https://orcid.org/0000-0002-5061-6765.

Kulyasov Nikita Vladimirovich – programmer; Institute of Computational Modelling of the Siberian Branch of the SB RAS – subdivision Federal Research Center "Krasnoyarsk Scientific Center of the SB RAS". E-mail: razor@icm.krasn.ru.

Isaev Sergey Vladislavovich – Cand. Sc., Associate Professor, Deputy Director for Research; Institute of Computational Modelling of the Siberian Branch of the SB RAS – subdivision Federal Research Center "Krasnoyarsk Scientific Center of the SB RAS". E-mail: si@icm.krasn.ru. https://orcid.org/0000-0002-6678-0084.

Статья поступила в редакцию 09.12.2024; принята к публикации 16.01.2025; опубликована 11.04.2025
The article was submitted 09.12.2024; accepted for publication 16.01.2025; published 11.04.2025