Method of hiding the architecture and configuration of the sensor network based on the dynamic topology

E A Kushko and N Y Parotkin
Institute of Informatics and Telecommunications, Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russian Federation

E-mail: evgeny.kushko@gmail.com

Abstract. Sensor network usually refers to a wireless distributed self-configuring network that consists of sensors and execution units and is designed for machine-to-machine communication, automation, machine diagnostics, telemetry task solving. Sensor networks differ from other networks by the ease of use and deployment, low maintenance requirements, fault tolerance, and reliability. Every year, sensor networks, which are usually used in industrial systems, are increasingly exposed to malicious attacks. Although some sensor network communication protocols support encrypted connections, the nature of node data exchange allows attackers to identify key nodes and possible attack vectors. In this paper we present a method of hiding the architecture and configuration of the sensor network, which allows to protect against compromise not only the nodes involved in data exchange, but also the transmitted data. This method does not interfere with the attacker's actions, but significantly complicates network traffic analysis, so that the intruder cannot carry out an attack because of the lack of data on network structure and configuration.

1. Introduction
Every year, more and more industrial enterprises around the world are exposed to hacker attacks. Attackers have different goals, including:
- disrupting the production process;
- disabling company's equipment and infrastructure;
- compromising confidential data;
- damaging company's reputation;
- others.

Although there is a trend of increasing information security budget, the majority of industrial enterprises cannot resist an attacker who penetrates the network infrastructure of an enterprise [1]. Once the intruder has penetrated the enterprise network, he analyzes it to identify data flows and key nodes in order not to waste time and resources on an unprofitable attack [2]. After he has identified the key nodes, an attacker investigates how these nodes interact with others in order to subsequently make an effective attack. Even if the data are encrypted, this does not guarantee that an attacker cannot identify the type of data being transmitted [3], and encryption does not hide the fact of interaction between two nodes.

One of the requirements of information security regulators of Russian Federation for automated production and process control systems is to hide the architecture and configuration of the information
(automated) system. In practice, this requirement is met by organizational measures in combination with the prevention of unauthorized access to the network by technical means. At the moment, there are no certified technical tools allowing to hide the architecture and configuration of the system. In order to address the above problems, the authors propose a new solution for the sensor network systems based on the secure data communication protocols stack based on the dynamic network topology, which is designed to counteract the analysis and interception of network traffic.

The MQTT (message queuing telemetry transport) protocol is selected as the application layer protocol, as it continues to be actively developed and applied in systems of various sizes and character, ranging from smart home systems to systems used in the aerospace industry. The choice of application protocol is due to the breadth of its use.

2. Interaction organization
The main idea of the method of hiding the architecture and configuration is the interval reconfiguration of the system and data transfer in such a way that the source and recipient of the data are not explicitly indicated. We use the secure data communication protocols stack based on the dynamic network topology as a session and transport protocol and MQTT as an application layer protocol, which operates on the publish-subscribe pattern and is constructed on the following principles [4]:

- ease of use and deployment;
- maintenance of communication in the conditions of low bandwidth and non-guaranteed communication quality;
- exchange parties have limited computing resources;
- high autonomy.

Refer to Reference [5] for a detailed description of the algorithms underlying the secure data communication protocols stack. Reference [6] provides information on comparing the protocols stack with similar existing solutions and its performance characteristics. However, the secure data communication protocols stack based on the dynamic network topology used to hide the architecture and configuration of the sensor network contains several differences.

![Figure 1. System structure.](image)

The machine-to-machine communication system based on the MQTT Protocol contains three types of participating nodes: client-subscriber, server-broker and server-publisher. All participants of the secure data communication decentralized form a multicast group address pool, through which data are transmitted. Decentralized configuration also increases system security [7]. This address pool changes at regular intervals, in addition, participating nodes are periodically randomly connect to two new multicast groups and disconnect from the old ones. The algorithm for selecting and generating a pool of multicast groups and the fact that each participating node relays all incoming data to its active multicast groups ensure that each node can communicate with any other node through any multicast.
This reconfiguration of multicast groups is necessary to ensure high complexity of network traffic analysis [8]. It should be noted that all data transmission is carried out via multicast groups of secure data communication.

At the initialization stage, the client-subscriber sends a request to open a secure connection to all currently connected multicast groups of the secure data communication. The server-broker, when receiving such a message, initiates the generation of a common session key for the client-subscriber and the server-broker under the Diffie-Hellman scheme. Using a session key, the server-broker and the client-subscriber exchange their public keys and identifiers. Since the network can contain several servers-brokers, the interaction is based on the principle of the first response. Further, servers-brokers store data on all nodes participating in secure data communication, which are synchronized with each other. Servers-publishers have predefined encryption keys, so they do not require additional procedure of authentication, keys and identifiers exchange. All MQTT data are encapsulated in the data packet of the secure data communication protocols stack and transmitted over the network.

The node sends data packet to all multicast groups of the secure data communication it is currently connected to. Connected nodes relay all data packets passing through multicast groups. To ensure that the relay is not infinite, each packet has a lifetime, i.e. the number of times a packet can be relayed. All data packets are of the same size so that the packet size does not depend on the type of transmitted data, this is necessary to complicate the analysis of network traffic. If the data size exceeds the packet size, then the data is divided into the required number of packets. When a data packet is received, the node retransmits it and then tries to decrypt the packet header with its private key. If the node is able
to extract its identifier from the packet header when decrypting, then the node considers that the packet is intended for it, otherwise the packet is discarded. Furthermore, the data packet is discarded if the maximum packet size is exceeded or the packet has been previously processed. Figure 2 shows the general packet processing scheme.

The main advantages of this method of data communication are:

- there are no point-to-point connections;
- the data packet does not explicitly contain information about the source and the recipient;
- the system configuration is decentralized;
- traffic analysis requires significant time and computational resources due to the use of encryption, reconfiguration of the address pool of multicast groups of secure data communication and use of the packets of the same size.

3. Performance results

To evaluate the effectiveness of the secure data communication protocols stack based on the dynamic network topology, we assembled a bench test consisting of three servers-publishers, two server-brokers and three clients-subscribers.

Every 15 seconds the server-publisher updates the data and the server-broker notifies the subscribers. Two servers-publishers are connected to one server-broker, and the remaining server-publisher is connected to the second server-broker. Each client-subscriber receives data from three different topics corresponding to three servers-publishers. Every 2 minutes, the pool of available multicast groups of secure data communication is reconfigured. Every minute, secure data communication participants reconnect to available multicast groups. The packet lifetime is 5. Figure 3 shows the initialization of the address pool of multicast groups and connection to them.

![Figure 3. Initialization of the address pool of multicast groups of secure data communication and connection to them.](image-url)
The data were collected within three minutes. The analysis of the transmitted data was carried out in three points: server-publisher with the address 192.168.1.1, client-subscriber with the address 192.168.1.4 and server-broker with the address 192.168.1.6. Figure 4 shows the histogram of the amount of traffic received and transmitted by each of three nodes. The data are evenly distributed, that means that it is extremely difficult to specify connections between two specific nodes.

Figure 4. Amount of the relayed and transmitted traffic.

4. Conclusion
The result of using the secure data communication protocols stack based on the dynamic network topology in a sensor network is:

- data are encrypted;
- connections between network nodes are hidden;
- node roles are hidden.

We achieve this result by means of decentralized system configuration, absence of point-to-point connections, absence of explicit indication of the data source and recipient IP addresses and generating identical traffic regardless of its type and content. The even distribution of all traffic indicates that the connection between two nodes is hidden among many identical data streams. It is extremely difficult for an attacker to identify these connections, as well as the amount of data transferred between specific nodes and their role in the system.

References
[1] Positive Research 2019 Retrieved from: https://www.ptsecurity.com/upload/corporate/ww-en/analytics/Positive-Research-2019-eng.pdf
[2] Husak M, Komarkova J, Bou-Harb E and Celeda P 2018 Survey of attack projection, prediction, and forecasting in cyber security IEEE Communications Surveys & Tutorials 21 640–60
[3] Bacquet C, Gumus K, Tizer D, Zincir-Heywood A and Heywood M 2009 A comparison of unsupervised learning techniques for encrypted traffic identification Journal of Information Assurance and Security 5 464–72
[4] Hunkeler U, Truong H L and Stanford-Clark A 2008 MQTT-S – a publish/subscribe protocol for wireless sensor networks Proc. of the Int. Conf. Communication Systems Software and Middleware and Workshops (Bangalore: IEEE)
[5] Kushko E A and Parotkin N Y 2019 Efficiency evaluation of secure data communication protocols stack based on dynamic network topology Proc. of the Int. Russian Automation
Conf. (Sochi: IEEE)

[6] Kushko E A and Parotkin N Y 2019 Software implementation details of the secure data communication protocols stack based on the dynamic network topology J. Phys.: Conf. Ser. 1399 033106

[7] Shirazi F, Simeonovski M, Asgharand M R, Backes M and Diaz C 2018 A survey on routing in anonymous communication protocols ACM Computing Surveys 51(3)

[8] Xu J, Guo P, Zhao M, Erbacher R, Zhu M and Liu P 2014 Comparing different moving target defense techniques Proc. of the 1st ACM Workshop on Moving Target Defense (Scottsdale: ACM) pp 97–107