Advances in Intelligent Systems Research, volume 174
Proceedings of the 8th Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS 2020)

Providing Information Security on the Base of Artificial Immune System for Industrial Internet of Things

Vladimir Vasilyev
Department of Computing and Information Security
Ufa State Aviation Technical University
Ufa, Russia
vasilyev@ugatu.ac.ru

Rinat Shamsutdinov*
Department of Computing and Information Security
Ufa State Aviation Technical University
Ufa, Russia
shrr2019@yandex.ru

***Abstract*—The paper analyzes the issues of providing information security of Industrial Internet of Things (IIoT). The features of both known and unknown attack detection in IIoT systems, in particular in wireless sensor networks, are considered. The analysis is carried using WSN-DS database as an exemplary case. The possibility of reducing the dimension of informative parameter space is investigated. The architecture of network anomaly detection system on the base of distributed artificial immune system offered for attack detection in wireless sensor networks is considered. The results of computational experiments confirming the effectiveness of the proposed approach to providing wireless sensor network security are presented.**

Keywords—cybersecurity, industrial internet of things, ambient security, intrusion detection system, artificial immune system

I. INTRODUCTION

Industrial Internet of Things (IIoT) is a system comprising computer networks and connected industrial objects with integrated sensors and software for data collection and exchange. IIoT provides remote control and administration of industrial processes in an automated mode. Nowadays, this class of systems is increasingly being used. IIoT is used in such fields as jet engine maintenance, elevator service as well as in industrial monitoring systems [1].

IIoT systems are distributed systems including a lot of different types of sensors, actuating mechanisms, human-machine interfaces. IIoT systems contribute to implementation of complex industrial processes, characterized by the use of cloud technologies and other perspective information and telecommunication technologies.

One of the most important concerns in the field of IIoT is providing security. Each IIoT solution, in one way or another, includes collection of information, and such information can be quite sensitive. For example, the data cloud of a large medical center IIoT network used to monitor the patient status can hold detailed information about the status of each patient, their identity and other information, the confidentiality of which is guaranteed.

Cyberattacks can easily disable IIoT devices responsible for the critical infrastructure [2]. Moreover, an unauthorized intrusion into technological process management is a great threat, especially at critical and potentially dangerous objects. Such intrusions can cause enormous damage, which is why they are the subject of cybersecurity.

The Internet of Things (IoT) systems and IIoT systems are similar in their infrastructure and tasks to systems that comply with the concept of ambient intelligence [3]. The major difference between the two is that the concept of ambient intelligence involves solving problems using artificial intelligence methods, including solving issues of interaction in complex technical systems such as IIoT.

The ambient intelligence security issues are considered in the framework of ambient security concept. It should be noted that the application area of ambient security, along with IoT and IIoT, also covers urban safety systems, environmental monitoring systems, etc. [4]. The main problems of ambient security systems are similar to problems of Security Information and Event Management (SIEM) systems development. They include collection and aggregation of data, data normalization, correlation or intelligent data analysis, security incidents detection, logging and notification of possible threats.

IIoT can be considered in ambient security concept framework as:

- security system that ensures the safety of managed technological processes;
- security object, unauthorized access or impact on which may cause certain damage.

European Union Agency for Cybersecurity (ENISA) published recommendations for IoT devices security provision in context of critical infrastructure objects at the end of November 2017 [5]. In November 2018, ENISA also released a paper called “Good Practices for Security of Internet of Things in the Context of Smart Manufacturing” [6], combining IIoT best cybersecurity policies, organizational and technical practices. Also, in 2018, a Draft NISTIR 8200 “Interagency Report on Status of International Cybersecurity Standardization for the Internet of Things (IoT)” [7] was presented, it contained tasks, risks and threats for IoT, a review of cybersecurity standards.

In May 2019, Methodological Recommendations MP 26.4.001-2019 “Secure Exchange Protocol for Industrial Systems (CRISP 1.0)” [8] were approved by the Technical Committee for Standardization “Cryptography and Security

Mechanisms". Protocol discussed in the paper is designed to protect communications between Automated Control System devices, Machine-to-Machine and IIoT devices. CRISP is the Russian cryptographic protocol that ensures the integrity and authenticity of transmitted information without ensuring confidentiality.

The paper is organized as follows. Section 2 deals with general information security issues of wireless sensor networks. Section 3 describes Artificial Immune System. The results of conducted computational experiments are presented in Section 4 followed by Conclusion.

II. Information Security of Wireless Sensor Network

Wireless Sensor Networks (WSNs) are frequently used in the framework of IIoT. WSN consists of a large number of autonomous sensor nodes that collect important data in various fields and share them wirelessly to a more powerful node, called the receiver node or Base Station (BS). WSNs are very vulnerable to attacks because of their distributed nature, openness and limited resources of sensor nodes. Moreover, WSN often broadcasts packets with sensor nodes being deployed in the environment randomly, which greatly simplifies the attacker's intrusion into WSN [9].

Attackers can compromise a sensor node, sniff, edit and imitate messages. They also can break data integrity and consume network resources. Denial of Service (DoS) attacks are considered as one of the most common and dangerous attacks that threaten the security of WSN. They can take several forms, their main purpose is to interrupt or suspend services provided by WSN [10, 11].

Different network connection databases are used for development of network security systems. The KDD Cup 99 dataset [12] is one of the most famous and common one. It was generated by emulating a military network environment in 1999. The dataset was created in a military environment in which a local air force network was subjected to simulated attacks. It contains about 4 million lines of data about network connections, each line contains 41 parameters.

NSL-KDD [13] is an improved version of KDD Cup 99. In [14], the NSL-KDD database was used to detect attacks in wireless networks at the network level and higher, and data generated by the author himself was used to detect attacks at the lower levels.

There is Aegean Wi-Fi Intrusion Dataset (AWID) presented in [15]. It is an open database of network connections in a Wi-Fi network, contains two different versions: full and reduced. According to [16], each AWID dataset record contains 155 attributes, including an attribute that determines whether the record matches the normal or abnormal state.

However, KDD Cup 99 and, accordingly, NSL-KDD do not contain data specific to Wi-Fi networks. Moreover, in WSN, according to [17], routing attacks are of great interest. Such attacks are contained in the wireless sensor network interaction dataset called WSN-DS and presented in [18]. In this regard, the dataset WSN-DS was selected, containing many lines of LEACH protocol parameter values in the normal state of the system and during the simulation of four types of DoS attacks: Blackhole, Grayhole, Flooding and Scheduling.

A. Attacks Description

- The Blackhole attack. The attacker declares himself the Cluster Head (CH). After that, any node joining this channel will send data packets to the attacker to forward them to the BS. The attacker will receive packets, but will not forward them.
- The Grayhole attack is similar to Blackhole attack, but still some of the packets intended for transmission to the BS will be sent to the addressee.
- The Flooding attack here is to send many different messages, including to nodes that are far away, with high transmit power about the announcement of the CH.
- Scheduling attack. The attacker configures the LEACH protocol so that collisions occur [18].

B. WSN-DS Dataset Attributes

A short description of the analyzed WSN-DS attributes presented in [18] is given in Table 1.

TABLE I. WSN-DS Parameters Description

Parameter		
Number	***Name***	***Description***
1	ID	Node ID: a unique ID to distinguish the sensor node in any round and at any stage
2	Time	the current simulation time of the node
3	Is_CH	Is Cluster Head? A flag to distinguish whether the node is Cluster Head with value 1 or normal node with value 0
4	Who_CH	Who is Cluster Head? The ID of the Cluster Head in the current round
5	Dist_ To_CH	Distance to Cluster Head: the distance between the node and its Cluster Head in the current round
6	ADV_S	ADV_CH send: the number of advertise Cluster Head's broad-cast messages sent to the nodes
7	ADV_R	ADV_CH receives: the number of advertise Cluster Head messages received from Cluster Heads
8	JOIN_S	Join_REQ_send: the number of join request messages sent by the nodes to the Cluster Head
9	JOIN_R	Join_REQ_receive: the number of join request messages received by the Cluster Head from the nodes
10	SCH_S	ADV_SCH_send: the number of advertise TDMA schedule broadcast messages sent to the nodes
11	SCH_R	ADV_SCH_receives: the number of TDMA schedule messages received from Cluster Heads
12	Rank	Rank: the order of this node within the TDMA schedule
13	DATA_S	Data sent: the number of data packets sent from a sensor to its Cluster Head
14	DATA_R	Data received: the number of data packets received from Cluster Head
15	Data_Sent_ To_BS	Data sent to Base Station: the number of data packets sent to the Base Station
16	dist_CH_ To_BS	Distance Cluster Head to Base Station: the distance between the Cluster Head and the Base Station
17	send_code	Send Code: the cluster sending code
18	Consumed Energy	Energy consumption: the amount of energy consumed in the previous round

C. Data Preprocessing

The values of various parameters in the original WSN-DS dataset are presented in both integer and fractional forms, the minimum and maximum values also differ greatly between the parameters. The values were preprocessed to a single presentation format. The values of all parameters after preprocessing are integer and, except for flag values, are in the range from 0 to 64.

Preprocessing was carried out as follows. Flag values were left unchanged. Parameters 1 and 4 included ID, round number and stage number. For example, node number 25 in the third round and in the first stage is to be symbolized as 001003025. We used only the ID value. For integer and fractional attributes, a threshold value was chosen such that, the initial values greater than the threshold became equal to 64 after preprocessing. If the initial value was strictly equal to zero, then the preprocessing result was equal to zero. All other values were distributed in the range from 1 to 63, it was calculated as:

$$y = 1 + \lfloor x(62/P) \rceil \quad (1)$$

where y_i is the parameter value after preprocessing, x_i is the parameter value before preprocessing, P is the selected threshold value.

The resulting dataset was divided into data on normal and malicious activities. A range from 0 to 64 was chosen, since greater compression resulted in a large number of identical lines between sets of normal and malicious activity. Less compression would degrade performance.

Next, the significant informative parameters were selected. To do this for each line of attack data, the most similar line of normal activity data was found, the numbers of matching parameters were written out. Based on the data obtained, the percentage of matches between the sets of normal and abnormal activity for each parameter was calculated. The parameters were ranked by the lowest percentage of matches, they are presented in Table 2.

The calculation of the minimum sufficient number of informative parameters is necessary in order to improve performance. We solved this problem as follows. We selected the first 15 ranked parameters, and then we calculated the number of identical lines between the sets of normal and malicious activities, gradually reducing the number of parameters. When choosing 7 parameters, the number of matches between these sets increased sharply. Further computational experiments were conducted according to the first ranked 8 parameters.

TABLE II. PARAMETER MATCHING PERCENTAGE

Parameter Number	Parameter Name	Parameter Matching Percentage
2	Time	20,65%
18	Consumed Energy	25,69%
15	Data_Sent_ To_BS	46,04%
7	ADV_R	50,11%
16	dist_CH_ To_BS	60,76%
1	ID	61,15%
9	JOIN_R	78,06%
4	Who_CH	81,82%
10	SCH_S	83,14%
14	DATA_R	86,59%
6	ADV_S	90,46%
11	SCH_R	98,96%
12	Rank	99,18%
5	Dist_ To_CH	99,34%
13	DATA_S	99,82%
17	send_code	99,94%
3	Is_CH	100,00%
8	JOIN_S	100,00%

III. ARTIFICIAL IMMUNE SYSTEM

Artificial Immune Systems (AISs) should be identified as perspective systems implementing artificial intelligence

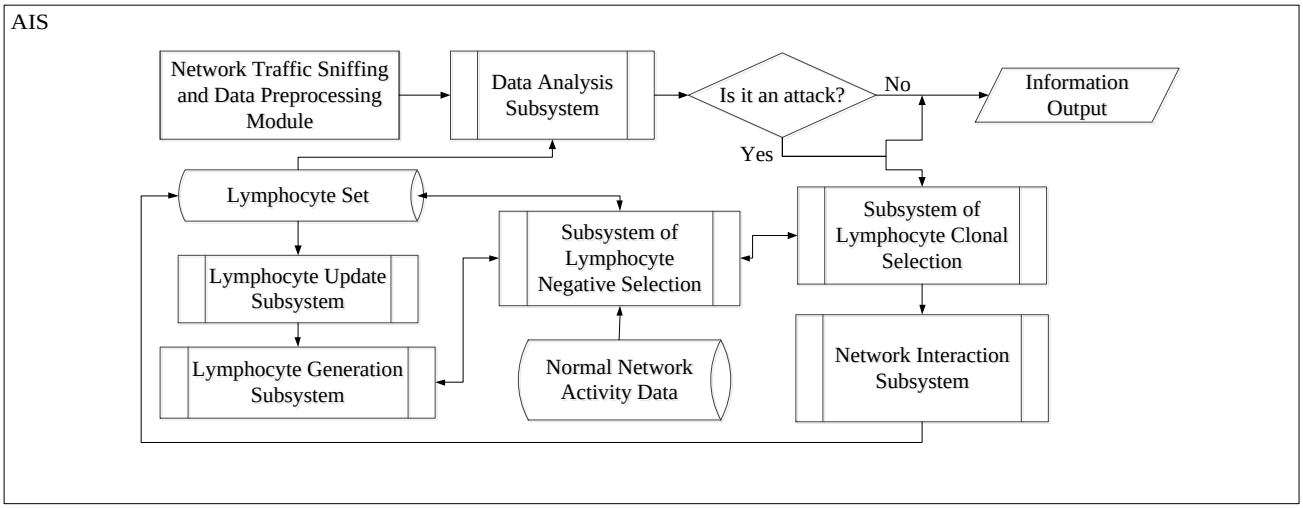

Fig. 1. The modules interaction scheme of the developed system [19]

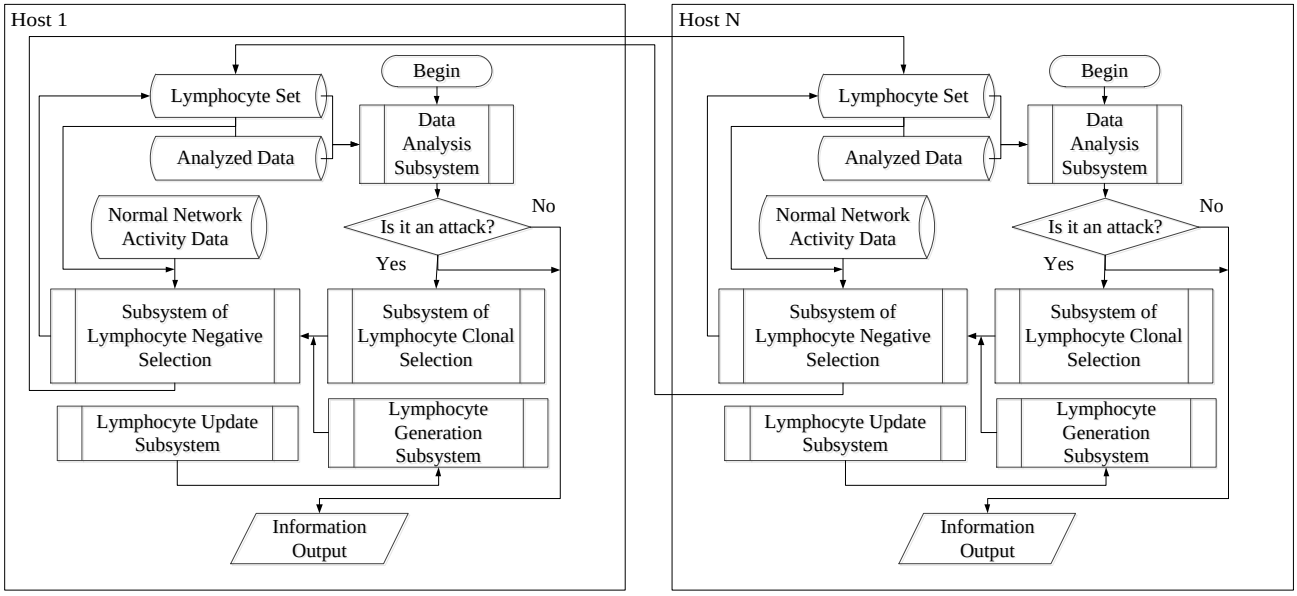

Fig. 2. Distributed AIS Hosts Iteraction Scheme

methods for solving ambient security issues in IIoT. The block diagram of AIS is shown in Fig. 1.

AIS imitates the functioning of human natural immune system, it doesn't have a single manage center. AISs are a distributed multi-agent system capable of detecting anomalies in a controlled system. Previously, studies of AISs using effectiveness were conducted relative to information systems, where they demonstrated a high level of unknown network attacks detection. Now we use the AIS developed and described in [19] for detecting attacks in wireless multisensor networks.

The choice of AIS as a classifier is due to its following advantages:

- an ability to detect unknown attacks;
- high tolerance to the normal state of the system;
- an ability to automatically learn during its functioning;
- high performance, as it was presented in [20, 21], AIS exceeds its main competitors (artificial neural networks and genetic algorithms) at least 40 times in speed and 2 times in recognition error-free.

This AIS was improved and distributed on several computers (hosts). The scheme of their interaction is presented in Fig. 2. AIS analyzes the strings of integer data, recognize arrays corresponding to attacks among them, sends alerts, generates training information, trains itself on it, transfers it to the other hosts of the distributed system.

AIS contains such important elements as:

- database of specialized agents also called artificial lymphocytes;
- training database of normal activity;
- agent generation subsystem;
- agent negative selection subsystem;
- data analysis subsystem;
- agent clonal selection subsystem;
- subsystem of logging and information output.

As can be seen in Fig. 2, each host after detecting an anomaly not only automatically learns to detect similar anomalies, but also trains the other hosts. AIS have been adapted for analysis of the WSN-DS dataset, which was divided into 3 parts:

- 50% of the normal state data intended for teaching AIS;
- 50% of the data on the normal state, intended for detect False Positives;
- 100% attack data designed to measure detection efficiency.

Initial training of AIS was carried out on the basis of half of data on normal activity, 500,000 agents were generated that have to be tolerant to the normal state of the system. AIS carries out subsequent training itself independently on the base of analysis results. Thus, all attacks presented in WSN-DS are unknown to the system. On three hosts, an instance of the program was launched. Host *A* performed the analysis procedure, it recognized only 3 attacks, generated training data, which was used to train itself and sent to hosts *B* and *C*. Then host *B* conducted the analysis procedure, trained itself, sent training data to the others. Then the analysis was carried out by host *C*, which already identified 10 attacks. Thus, the hosts trained themselves and trained the entire system, each time increasing the detection efficiency.

IV. Results of Computational Experiments

Since the AIS must additionally learn based on the results of data analysis, a series of experiments to detect attacks was conducted. At the first iteration of the analysis, only 3 attacks were detected. However, this turned out to be enough for system self-learning based on them, and the rapid growth of the detection efficiency of similar attacks.

Computational experiments were started based on eight selected parameters. With each subsequent iteration of the analysis and further training, the number of False Negative decreased until it reached a certain limit of 5%. This is a fairly large number of errors. Eight parameters were not enough. Therefore, we began to increase the number of parameters and repeat computational experiments until the absence of a limit of error reduction was noted.

The False Negative Rate (*FNR*) was calculated as:

$$FNR = FN / (TP + FN) \times 100\%, \quad (2)$$

where, *FN* is the number of False Negative errors, *TP* is the number of True Positive.

The False Positive Rate (*FPR*) was calculated as:

$$FPR = FP / (TN + FP) \times 100\%, \quad (3)$$

where *FP* is the number of False Positive errors, *TN* is the number of True Negative.

The False Positive Rate at each iteration did not exceed 0.1%, due to the features of the AIS. The False Negative Rate depending on the number of parameters and the number of iterations of the analysis and further training are presented in Table 3.

TABLE III. THE FALSE NEGATIVE ERRORS RATE

Iteration Number	Number of Parameters						
	8	*9*	*10*	*11*	*12*	*13*	*14*
1	99.9%	99.8%	99.9%	99.9%	99,9%	98.6%	99.2%
10	66.4%	55.7%	59.2%	38.2%	45.6%	66.5%	52.3%
20	45.9%	40.0%	43.1%	29.9%	35.2%	55.9%	28.0%
30	31.7%	29.0%	31.7%	25.5%	19.0%	41.5%	22.1%
40	21.5%	23.0%	24.8%	19.8%	16.6%	25.0%	17.2%
50	15.1%	17.0%	15.2%	14.0%	9.8%	21.3%	11.1%
60	8.8%	7.2%	4.5%	9.9%	7.5%	13.1%	7.9%
70	5.0%	4.1%	3.9%	4.7%	5.2%	7.3%	6.9%
80	5.0%	4.1%	3.9%	1.8%	1.6%	1.5%	4.7%
90	5.0%	4.1%	3.9%	1.8%	1.6%	1.5%	2.0%
100	5.0%	4.1%	3.9%	1.8%	1.6%	1.5%	0.8%

According to Table 3, the most rational thing is to use the first 14 parameters shown in Table 2, because it is with 14 parameters that the reduction limit of the False Negative errors disappears. It might appear again but at a much low error level, thus it would be permissible. For clarity, the data in Table 2 is presented graphically in Fig. 3. Here, the Detection Efficiency (DE) is calculated as:

$$DE = TP / (TP + FN) \times 100\%. \quad (4)$$

As it can be seen from Fig. 3, the system demonstrates high efficiency of detecting unknown attacks. According to the results of 100 iterations of the analysis, the detection efficiency exceeded 99%. The level of the False Negative and the False Positive errors does not exceed 1%. Moreover, the distributed nature of the system significantly increases its reliability, since each node in it not only works independently,

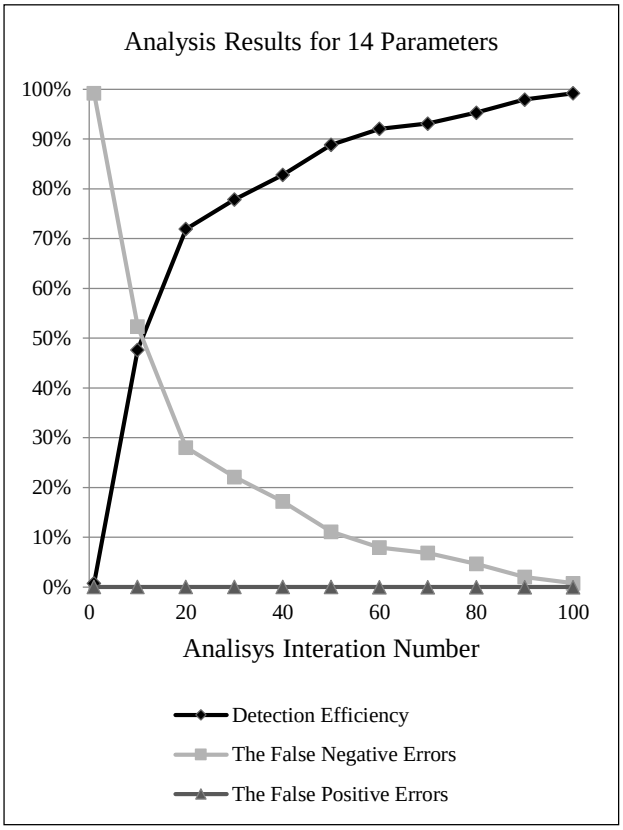

Fig. 3. Analysis Results for 14 Parameters

but at the same time also teaches other nodes to better recognize attacks similar to detected attacks.

The system showed high efficiency in detecting unknown attacks on wireless sensor networks with a low level of False Negative and False Positive errors. The developed system can be applied into any IIoT networks using LEACH protocol.

V. CONCLUSION

Thus, providing security in IIoT systems is an important issue, where the use of intelligent systems is the most perspective solution. Wireless sensor networks are widely used as part of IIoT. The article proposes an approach to detecting threats in WSN, based on the use of a distributed artificial immune system. The architecture of this system is presented.

The dataset WSN-DS, which was considered in detail in [18], was chosen as the analyzed database. This dataset contains LEACH network communication data, as well as normal activity data and attack data. WSN-DS dataset presents 4 types of DoS attacks: Blackhole, Grayhole, Flooding and Scheduling.

In order to improve performance, the space dimension of informative parameters presented in WSN-DS was reduced. For parameters 1 and 4, only ID values were saved. Flag parameters were left unchanged. Integer and fractional parameters were modified as follows. A threshold value was selected for each relevant parameter. If the original value exceeds the threshold, then the result value is 64. If the

original value is strictly equal to 0, then the result value is 0. In all other cases, the result is calculated by (1) and is in the range from 1 to 63.

Then, for each parameter, the percentage of matches between the sets of normal and malicious activity was calculated. The parameters were ranked by the lowest percentage of matches and are summarized in Table 2.

To conduct computational experiments, the distributed AIS was chosen, its architecture, the main functional elements, the interaction scheme of different hosts as a single system were presented. The presented results of computational experiments indicate the rationality of the choice of the first 14 ranked parameters. The system showed a high efficiency in detecting unknown network attacks on WSN. The total percentage of errors False Negative and False Positive does not exceed 1%. The distributed nature of the system allows hosts to train each other, which greatly improves the reliability of the system.

ACKNOWLEDGMENT

The work was supported by the Russian Basic Research Foundation (RBRF) grant №20-08-00668.

REFERENCES

[1] A Full Set of Equipment for Industrial Internet of Things, IPC2U, available at: https://ipc2u.ru/articles/obzory-produktov/industrial-iot/ (accessed 05.02.2020). (in Russian).

[2] G. Falco, C. Caldera and H. Shrobe, IIoT Cybersecurity Risk Modeling for SCADA Systems, IEEE Internet of Things Journal, vol. 5, no. 6, Dec. 2018, pp. 4486-4495.

[3] K. Gunnarsdóttir, M. Arribas-Ayllon, Discussion Paper: Ambient Intelligence: a Narrative in Search of Users, Lancaster University, available at: https://eprints.lancs.ac.uk/id/eprint/74291 (accessed 21.01.2020).

[4] P.M. Ivanov, O.B. Makarevich, Z.V. Nagoev, Automatic Forming of Context of Situations in Ambient Security Systems on a Basis of Multiagent Cognitive Architectures, News of SFedU. Technical Science, no 12 (149), 2013, pp. 33-39. (in Russian).

[5] Recommendations to Ensure the Security of IoT Devices Have Been Developed, Anti-Malware, 2017, available at: https://www.anti-malware.ru/news/2017-12-26-1447/25175 (accessed 03.03.2020). (in Russian).

[6] Good Practices for Security of Internet of Things in the Context of Smart Manufacturing, European Union Agency for Cybersecurity, 2018, available at: https://www.enisa.europa.eu/publications/good-practices-for-security-of-iot (accessed 28.01.2020).

[7] Draft NISTIR 8200. Interagency Report on Status of International Cybersecurity Standardization for the Internet of Things (IoT) / Interagency International Cybersecurity Standardization Working Group, 2018, available at: https://nvlpubs.nist.gov/nistpubs/ir/2018/NIST.IR.8200.pdf (accessed 07.03.2020).

[8] Methodological Recommendations MR 26.4.001-2019 Information Technology. Cryptographic Information Security. Secure Exchange Protocol for Industrial Systems, Technical Committee for Standardization "Cryptography and Security Mechanisms", 2019, Moscow, Russia, available at: https://tc26.ru/standarts/metodicheskie-rekomendatsii/mr-26-4-001-2019-protokol-zashchishchennogo-obmena-dlya-industrialnykh-sistem-crisp-1-0-.html (accessed 07.03.2020). (in Russian).

[9] J. Sen, Security in Wireless Sensor Networks, in Wireless Sensor Networks: Current Status and Future Trends, S. Khan, A.-S. K. Pathan, and N. A. Alrajeh, Eds., CRC Press, New York, USA, 2012, pp. 407-460.

[10] N. Farooq, I. Zahoor, S. Mandal, and T. Gulzar, Systematic Analysis of DoS Attacks in Wireless Sensor Networks with Wormhole Injection, International Journal of Information and Computation Technology, vol. 4, no. 2, 2014, pp. 173–182.

[11] A. Mitrokotsa, T. Karygiannis, Intrusion Detection Techniques in Sensor Networks, Wireless Sensor Network Security, Cryptology and Information Security Series, pp. 251–272, IOS Press, 2008.

[12] KDD Cup 1999 Data, available at: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html (accessed 03.03.2020).

[13] NSL-KDD dataset, available at: https://www.unb.ca/cic/datasets/nsl.html (accessed 03.03.2020).

[14] I. V. Sharabyrov, Local Wireless Networks Attack Detection System Based on Intelligent Data Analysis, PhD Thesis, Ufa State Aviation Technical University, Ufa, Russia, 2016. (in Russian).

[15] C. Kolias, G. Kambourakis, A. Stavrou and S. Gritzalis, Intrusion Detection in 802.11 Networks: Empirical Evaluation of Threats and a Public Dataset, IEEE Communications Surveys & Tutorials, vol. 18, no. 1, Firstquarter 2016, pp. 184-208.

[16] AWID Dataset Description, available at: http://icsdweb.aegean.gr/awid/features.html (accessed 04.03.2020).

[17] V.I. Vasilyev, A.M. Vulfin, V.M. Kartak, A.D. Kirillova, K.V. Mironov, System of Attacks Detection in Wireless Sensor Networks of Industrial Internet of Things, Proceedings of the Institute for System Analysis of the Russian Academy of Sciences, no 4, 2019, pp.70-78. (in Russian).

[18] I. Almomani, B. Al-Kasasbeh, M. AL-Akhras, WSN-DS: A Dataset for Intrusion Detection Systems in Wireless Sensor Networks, Journal of Sensors, vol. 2016, available at: https://www.hindawi.com/journals/js/2016/4731953/ (accessed 01.03.2020)

[19] V. Vasilyev, R. Shamsutdinov, Distributed Intelliget System of Network Traffic Anomaly Detection Based on Artificial Immune System, Proceedings of the 7th Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS 2019), May 28-29, 2019, Ufa, Russia, Advances in Intelligent System Research, vol 166, pp. 40-45.

[20] A. O. Tarakanov, Y. A. Tarakanov, A Comparison of Immune and Genetic Algorithms for Two Real-Life Tasks of Pattern Recognition, Int. J. of Unconventional Computing, 2004, vol. 1.4, pp. 357-374.

[21] A. O. Tarakanov, Y. A. Tarakanov, A Comparison of Immune and Neural Computing for Two Real-Life Tasks of Pattern Recognition, International Conference on Artificial Immune Systems, Catania, 2004, pp. 236-249.