Host-based Method and System for Detecting Anomalies in Network Traffic for a Robotic System

Elena Basan
Dept. of Information Security
Southern Federal University
Taganrog, Russia
ebasan@sfedu.ru

Maria Lapina
Dept. of Information Security of Automated Systems
North Caucasus Federal University
Stavropol, Russia
mlapina@ncfu.ru

Dmitry Orel
Dept.Information Security Organizations and Technologies
North Caucasus Federal University
Stavropol, Russia
kde.def@gmail.com

Abstract

This study is devoted to the problem of detecting anomalous behavior of nodes of a robotic system based on network traffic analysis. This article addresses the issue of analyzing changes in the level of network traffic passing through a network node in order to detect denial of service attacks and a black hole attack. To solve this problem, the authors propose to use probabilistic and statistical methods, as well as methods of information theory. The robot wireless network model was developed to collect statistics.

1 Introduction

In this study, stationary robotic systems will be considered. By this definition, we mean intelligent systems that collect, process, analyze data about the environment and are able to perform environmental change. Such systems include teleoperated robots, smart home systems, the Internet of things, Internet robots, as well as sensor networks. In essence, all these systems have similarities: perform the task associated with data collection, complete the previously defined sets of actions, periodically exchange information with each other, have a base station responsible for processing information, exchange data over wireless channels, may have limited computing and energetic resources. Today, the systems "Smart Home", "Smart City" are becoming increasingly popular. They are being introduced both into the daily life of people and into production, as well as into the infrastructure of urban transport, etc. There are a large number of examples of such systems. At the same time, the Strategic Marketing Center Smolin & Partners reports in its report that: The Smart Home category market in Russia will resume growth after the economy transitions to growth. Starting from 2017-18. the annual projected increase will be about 10% from last year. It should be noted that such systems, as a rule, are foreign developments of Chinese, Korean, American, etc. companies [Hag16]. Data received and transmitted between network devices is

In: S. Hölldobler, A. Malikov (eds.): Proceedings of the YSIP-3 Workshop, Stavropol and Arkhyz, Russian Federation, 17-09-2019–20-09-2019, published at http://ceur-ws.org

stored on foreign servers, event and voice recognition systems, as well as other services are located there. At the same time, the company Positive Tecnologies announce the following: This year, analysts call the development of IoT, the Internet of things, one of the main problems. Experts of Positive Technologies made the top 5 of the most dangerous for the user devices with access to the Internet. First of all, this is the heart of the entire home network - a Wi-Fi or 3G-4G router. Experts find up to 10 vulnerabilities every month in these devices. To date, a large number of publications of scientists devoted to the analysis of threats and vulnerabilities of systems "Smart Home". At the same time, in the majority of works by Russian and foreign authors, the main security problems associated with the Smart Home systems are highlighted: Hacker attacks on the central server; Interception of information transmitted via wired and wireless communication channels; Access by an attacker with administrative rights to a central server by stealing passwords and other access control details; Access to the network of unauthorized users, etc.". Every year there are more and more cases of hacking of the systems of Smart Home, often the attacker does not need to have special means to penetrate the system or to get data from the servers [Bas17]. Such a formulation of the problem poses the task of developing fundamentally new methods and means of protecting information systems. Thus, the second problem can be formulated as: the development of means to ensure the security of the Smart Home systems, based on independently obtained fundamental methods, taking into account domestic software and hardware developments [She14].

This study provides a new method for detecting anomalies in the behavior of nodes without using a signature database, a rule base and a template for normal node behavior. Detection of anomalies will be determined by analyzing changes in the fixed parameters of the node with respect to the current behavior of other nodes in the network. This method will reduce the cost of creating and storing databases of signatures, or rules, as well as allow you to get away from the need to build databases, if necessary, to fix new attacks. In addition, this method is not tied to a specific architecture and network structure of the system, for its implementation it is not important to use specific protocols and data structures, it can be adapted for any system.

2 Methods for construction of robotic systems

All methods for constructing stationary robotic systems can be combined into two groups:

1. With one main node (single-hop) - the transmitter transmitter power is sufficient for transmitting the signal to the base station.

2. With several main nodes (multi-hop) - some nodes not only collect information about the observed process, but also collect information from other nodes [Par09].

Each stationary robotic system has a specific set of parameters, such as:

number of network nodes;

data transmission rate over communication channels - the parameter is determined by the channel bit rate. As a rule, the wireless sensor network consists of the same devices, so for all communication channels the bit rate will be the same [Sigh15];

network topology - determined by the presence or absence of a radio channel between any two nodes;

node placement density - the average number of neighbors located close to the node;

network diameter - the minimum number of retransmissions for data transmission between the two most distant network nodes;

delay in the network - the time from the moment of the occurrence of the event to the moment the information about it appears at the base station;

network operation time - network operation without recharging and without changing batteries;

bandwidth - characterized by the amount of information per unit of time.

As a rule, robotic systems are based on the principles of a special network in which each of the nodes (part or nodes) can be a source or receiver of messages and a transit node. In this case, the network is determined by the operation of node selection protocols and routing protocols. Routing in such a network requires a significant investment of resources for transmission of service traffic, and the sum of these costs depends on the nature of the traffic, and the stability of the network structure depends on the characteristics of the traffic. The quality of wireless connections between nodes is strongly influenced by the number of transit routes (hops) in the routes [Mill13]. In mobile networks, nodes can move, and the routing task is more complex. One of the ways to implement mobile networks is to use small mobile robots to transfer sensitive elements to the detection zone. If the algorithm that controls the movement of robots is based on the swarm approach, it consists of a swarm robot. In a typical swarm of robots, only communication between the robots is implemented to ensure the behavior of the swarm. Group communication between mobile robots thus requires protocols that can operate without

central control and handle dynamic topology changes due to the mobility of mobile robots. Multicast is the most important group communication primitive and is critical in applications that require close collaboration between groups (for example, rescue teams, search groups). This is very useful when audio, video, images and other similar data should be passed on to team members. Multicast provides an effective means of sending the same data to multiple recipients. Compared to multiple unicast streams, multicast minimizes channel bandwidth consumption, sender and router processing, and delivery delay. The use of multicast as a group communication primitive in networks of mobile robots can be provided in the following three scenarios: first, multicast messages on networks of mobile robots can be transmitted from one mobile robot to a set of other mobile robots. These nodes can be organized into a multicast group to periodically notify each other about their positions and act as landmarks and navigation signs for other robots that do not know their positions [Vilch18].

To implement messaging between robots, the Multicast Routing Monitor (MRM) Protocol was chosen [Wei99].As a rule, the work of a group of robots is divided into several main stages: the collection of environmental data, the sending of the collected data to neighboring devices or the base station, and the receipt of commands from the base station. To simulate the sending of messages as part of these tasks, the TCP transport layer protocol was chosen.. Allowing to transfer a data stream with a pre-established connection, re-requests data in case of data loss and eliminates duplication when receiving two copies of one packet, thereby ensuring the integrity of the transmitted data and notifying the sender about the results of the transfer. A limited number of nodes were chosen for more accurate modeling and tracking of traffic [Pshikh11]. A base station has been created and nodes have the ability to group together, exchanging packets, and once they have completed their tasks, they will again be grouped.

3 Intrusion Detection System Architecture for a Robotic System

The developed anomaly detection system has a modular structure. Each module performs its functions and communicates with other modules of the system. The basic idea underlying the operation of this system is as follows:

1. The robotic system operates according to the established algorithms and robots perform similar sets of actions. The work of the robot looks static. If robots exchange traffic with the control station or with each other, then this usually occurs within the framework of the request-response form. Thus, the traffic transmitted between network nodes also looks static and lends itself to the normal distribution law. That is, the total amount of traffic is gradually increasing over time [Bas18].

2. If we are talking about the detection of anomalies caused by an active attack, then the consequences of the impact should affect the operation of the network. In particular, the pattern or algorithm of traffic transmission should change significantly and will differ from the situation when the node worked without an attack. First of all, an active network attack affects integrity and availability. If we are talking about the violation of accessibility, then it may be sufficient to conduct a statistical analysis of the transmitted traffic. Violation of integrity may affect the accuracy of the functions performed and the measurements taken. In addition to network traffic, other parameters that are affected by the attack can be analyzed, but this study is limited to analyzing network traffic.

Figure 1 shows the modular architecture of the anomaly detection system. The figure shows that four values fall into the input of the data acquisition module for further analysis. These values are collected on the node itself in real time. The idea of detecting anomalies at the host level is that the robot analyzes itself over the past periods of time.

Next, we consider the features of each module.

1. Data acquisition module. The developed method of analyzing the effectiveness of an active network attack on a system of mobile robots, as one of the results is a set of system parameters that are affected by the attack. So, the information collection module will collect statistical information at the current time and further, using the methods of mathematical statistics, to perform primary information processing. That is, a set of information on the change of fixed parameters will be presented in the form of metrics that will allow us to further assess the presence of signs of abnormal activity. The information collected is divided into four large categories: parameters related to network traffic analysis; parameters associated with a change in the power plan of the device; parameters associated with changes in system characteristics, reflecting the load on the processor and the devices RAM. Each of the parameters is influenced by certain sets of attributes, and then the attributes themselves are analyzed and formalized.

2. Abnormality detection module. This module receives data from the module for collecting and processing information and detects the presence of abnormal activity. The module is based on a developed method for

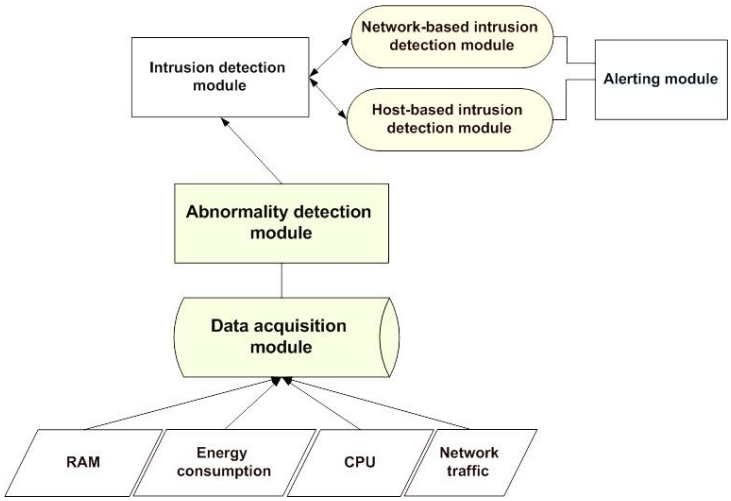

Figure 1: Architecture of modular anomaly detection system for a group of robots, including the set of parameters that are analized

detecting anomalies, which can be detected by an attacker without building a pattern of normal behavior and creating a database of signatures [Fag14]. The use of the method is possible due to the presence of a group of nodes that perform approximately the same functions: they fix parameters and transfer information to the central node. Network nodes, as a rule, act according to a predefined scenario and their actions should not go beyond the scope of the intended behavior. Depending on how strongly and which parameter deviated, it is possible to determine the probability of how malicious the node is and what kind of attack it conducts. This is possible due to the use of probabilistic methods, in particular, the construction of confidence intervals for the distribution function and the estimation of the probability of a parameter falling into a confidence interval. These calculations are made solely on the basis of those indicators that are obtained from the nodes - sensors and intelligent nodes for the past and current time interval, and there is no need to create a reference behavior pattern. Nodes themselves determine which of them behaves differently than others.

3. Intrusion detection module in the robotic system. After the information obtained has been normalized and processed in the previous two stages, it can be concluded which attack is carried out, to reveal its intensity and to determine the object of impact. The identification of system-specific attack criteria is an important task, since they will differ significantly from the standard sets of criteria of classical information systems. This is primarily connected with non-standard traffic patterns, protocols and standards. In addition, many attacks by an attacker can be associated with the exhaustion of resources of network nodes, thus cyber-parameters should be taken into account when an attack is detected. The intrusion detection and intrusion detection module, in turn, will be divided into two subsystems: the subsystem for analyzing changes in the site's own parameters; subsystem analysis of attacks and intrusions in the network. In other words, the creation of a system for detecting attacks and intrusions at the node level and not at the network level is implied. To implement the attack detection subsystem at the node level, information theory methods will be used, in particular, the Kullback-Leibler divergence measurements. Measuring this distance makes it clear how much the current distribution differs from the previously obtained. Thus, a node can build a normal distribution to change its parameters and compare the deviation from the previous one over time (or from that obtained in laboratory conditions) and draw a conclusion about whether there is an influence on one or several analyzed parameters, and depending on the degree of deviation to make a conclusion about the intensity of exposure.

4. Alerting module. This module allows you to notify both the node itself and other nodes about the presence of an attack. In addition, if we are talking about a system of detection of intrusions at the host level, it is necessary to provide for the adoption of measures to block or limit the operation of the node.

Thus, a node can analyze both its own behavior and the behavior of other network nodes. At the same time, the same metrics will be analyzed; they will be processed in a similar way. The only difference is that by analyzing their behavior, the node is described by the data that it collected itself and by analyzing the behavior of its neighbors by the data that it received via the wireless channel.

4 Host-Based Anomaly Detection Module

This study takes into account only metric network traffic. That is, the search for anomalies will be carried out in network traffic. In early studies, the network node load factor was analyzed as an integer value. In this study, we propose to divide network traffic into: sent packets, forwarded, received, and dropped. When the network load of the entire node is estimated, it is not fully understood what is connected with the change in this indicator. For example, if an attacker conducts a denial of service attack, then the number of packets that it sends out far exceeds the normal state. At the same time, the victim of the attack receives more requests and the traffic of the victim also increases. If you analyze the entire traffic or the network load of the node, you can conclude that the anomaly is present, but the victim and the attacker may be disqualified from the job, and the source of the attack may not be detected.

4.1 Definition of the format of the analyzed data

To collect statistical information, software was developed that includes the following components:

as a network sniffer, we will use the tcpdump program, since it has a powerful tool for intercepting traffic, a lot of conveniently configured filters, and a console interface;

programming language - Python, as it is supported by default by operating systems of the unix family, is interpretable, which eliminates hardware dependency, unlike compiled ones.

The program fulfills the following functional requirements:

call the network tcpdump sniffer from the main thread to intercept traffic on the node and then save the result to a file with the pcap resolution;

analyze the resulting file, highlighting the information that will generate statistical data for further calculation of the main indicators of the network;

to calculate the network performance indicators, compare them with the permissible values obtained in the study of a normally functioning network;

in case of deviation of the obtained values of the indicators from the permissible values, notify the user about the abnormal behavior of the node, about the type of possible attack;

save to the text file all the collected statistical information, calculated values of the indicators for the subsequent analysis of the incident by the user.

The collected values must be processed. We assume that the number of packets passing through the node is a random continuous value. This parameter is measured at the current time interval, and is summarized at all intervals, so we have a value that grows evenly throughout the entire time interval. The analysis of statistical data was carried out on the subject of what type of distribution they correspond to. To do this, quantile-quantile diagrams were constructed. This diagram estimates the change in the number of received packets by the node. Such diagrams were built for each type of package. It can be seen from the diagram that the statistics obtained are distributed uniformly near a straight line [Grjib08]. The idea is that if the values are distributed around a straight line, then the collected data correspond to the normal distribution.

Next, you need to calculate the values of the normal distribution. The difficulty lies in the fact that the distribution needs to be calculated continuously and to evaluate possible changes. To estimate the degree of difference between distributions, the use of Kullback-Leibler divergence is proposed. Divergence is a measure of the difference between two distributions and is used in information theory.

In order to build a normal distribution, it is necessary to build up an initial data set. The normal distribution will be more accurate than for more random variables it is constructed. For a small number of random variables, the normal distribution takes the form of a straight line. Experimentally, it was revealed that the minimum number of random variables needed to build a normal distribution is 6. Further, after the first normal distribution is constructed, it is necessary to compare it with subsequent distributions. It is proposed to use the concept of a sliding window. When we receive new measurements and after a certain period we build a new one. This takes into account the four previous values and two new ones. Thus, the sliding window is equal to two time intervals. It is assumed that the value of received / sent / received/discarded packets is fixed for a certain period of time; you can record values every second, less often depending on the network requirements.

Next will be presented a technique for detecting anomalies in a robotic system based on statistical analysis of network traffic.

4.2 Method for detecting anomalous activity based on statistical analysis of network traffic

1. At the first stage, it is necessary to construct a normal distribution of a random variable. To build a normal distribution, we define the conditions associated with the minimum necessary number of random variables to build up the distribution, as well as the size of the sliding window, in order to preserve the condition for dynamic construction of normal distribution.

$$\begin{cases} (s, r, f, d) \geqslant m \\ \Delta t = i \end{cases} \tag{1}$$

where are s - the sent packets, r - these are the received packets, f - these are the packets that were forwarded through the node, d - these are the packets that were dropped by the node, m - this is the threshold value that determines the minimum required number of random values of these parameters to build a normal distribution; (Δt) - this is the value of the time window parameter;i - the number of time intervals that will be taken into account when building each new distribution.

To date, the principle of determining the threshold values of m and has not been developed. In this study, it was established that the minimum number of random variables is advisable to take equal to six, that is, each normal distribution will be based on six values. The value of the sliding window is chosen to be two, that is, when each new normal distribution is constructed, four values of a random variable obtained at previous intervals and two values obtained at the last intervals will be used. This way of posting the normal distribution is due to the following. The normal distribution smoothes the change in the random variable and the time of the beginning of the attack may not be fixed if the distribution is rebuilt every time interval. If the scatter of values is significant, then the attack will be fixed due to the fact that the standard deviation will increase dramatically and it may even exceed the value of the expectation. Therefore, the minimum size of a sliding window for six values is proposed to choose equal to two.

2. After the initial conditions are determined, a normal distribution is constructed.

2.1 To build the distribution after data collection, it is necessary to calculate the mathematical expectation and standard deviation, as shown in formulas 2,3.

$$M(s) = \int_{\Delta t+i}^{\Delta t-m} sf(s)ds, \quad M(s) = \int_{\Delta t+i}^{\Delta t-m} rf(s)dr, \tag{2}$$

$$D(s) = \int_{\Delta t+1}^{\Delta t-m} \left[s - M(s)\right]^2 f(s)ds, \quad D(s) = \int_{\Delta t+1}^{\Delta t-m} \left[r - M(r)\right]^2 f(r)dr, \tag{3}$$

$$\sigma(s) = \sqrt{D(s)}, \quad \sigma(r) = \sqrt{D(r)}. \tag{4}$$

where M(s), M (r) is the mathematical expectation for send and received packets, it is also calculated in similar way for forwarded and received packets; D (s), D(r) - it is the variance of the random variable for the indicators adopted and sent pacts, also calculated for all types of packages; Third level headings must be flush left, initial caps and bold. $\sigma(s), \sigma(r)$ - this is the standard deviation for the same parameters. 2.2 Next is the normal distribution. First, the distribution for the first six intervals is constructed, then the four past intervals and the last two are taken, the formula for the calculation is presented below:

$$f(s) = \frac{1}{\sigma_s\sqrt{2\pi}} e^{\frac{(s-M_s)}{2\sigma_s^2}}, \quad f(r) = \frac{1}{\sigma_r\sqrt{2\pi}} e^{\frac{(r-M_r)}{2\sigma_r^2}}, \tag{5}$$

where *f(s), f(r)* - the normal distribution function of a random variable (in this case, the number of sent and received packets) at specified time intervals.

The result of the construction of the normal distribution is shown in Figure 2. This figure specifically shows the normal distributions for the parameter sent packets for the normal network node, which worked under normal conditions and was not affected either by the attacker or by any other external influence. These results of constructing a normal distribution were obtained during the simulation of the network operation of 120 seconds, while the time range was divided into intervals of 10 seconds. In particular, the distribution highlighted in blue was obtained in the range from 0-50 seconds; distribution highlighted in red from 20-80 seconds; the distribution highlighted in green is obtained in the interval from 40-100 seconds and respectively the last violet distribution from 60-120 seconds.

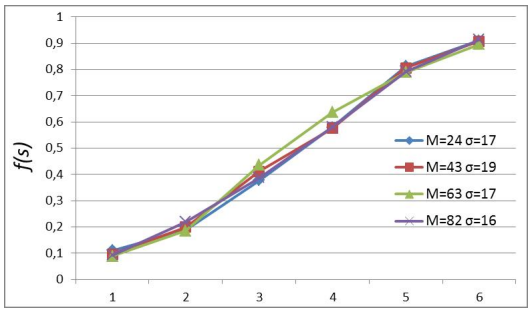

Figure 2: The result of the calculation of the Normal Distribution function is presented in this paper for the node that sent the packets without affecting the attack

From figure 3 it can be seen that the distributions practically overlap each other and there is no difference between them. At all-time intervals, the probability values coincide. So the node behaved the same on all time intervals. However, this information does not provide a complete picture that would allow an assessment of the presence of abnormal behavior. As a rule, a confidence interval is used to detect deviations from the normal distribution. But in the case of anomalies, this method gives false positives. Therefore, in this paper we will use another mathematical method.

3. Calculation of Kullback-Leibler divergence for analyzing the degree of differences between probability distributions. The idea of using this distribution is to compare how much the distribution just received has changed from the previous one. Kullback-Leibler divergence is used in information theory as a measure of the distance from each other of two probability distributions [14] defined on a common space of elementary events. With this measure, you can find out how much the behavior of the node has changed over the current period of time. Certainly peak mismatches have a place to be. We assume that the attack is signaled when peak mismatches occur on more than three time intervals in a row. That is, it will calculate Kullback - Leibler divergence.

$$D_{KL}(f_n(s)|f_{n-1}(s)) = \int f_n(s) * ln\frac{f_n(s)}{f_{n-1}(s)}ds, \tag{6}$$

$$D_{KL}(f_n(r)|f_{n-1}(r)) = \int f_n(r) * ln\frac{f_n(r)}{f_{n-1}(r)}dr, \tag{7}$$

where D_{KL} - this value Kullback Leibler divergence obtained by estimating the difference between the normal distribution obtained in the last 6 time intervals $f_n(s)$ and with the distribution that was obtained during the previous intervals $f_{n-1}(s)$, the same calculations are made for each type of packet, including for received packets, as shown in formula 7.

That is, starting from Figure 3, the difference will be the computation between the blue and red distributions, between red and green, and between green and purple.

The resulting histogram confirms that the discrepancy between the distributions tends to zero, as it can be seen from Figure 3 . Figure 3 shows results of Kullback Leibler divergence calculation. Therefore, we assume that if the resulting divergence value can be rounded to zero, then the attack is not carried out. Further, in order to determine the boundaries of the divergence, denial-of-service attacks were conducted with varying degrees of intensity. And also carried out the Black-Hole attack.

The resulting histogram confirms that the discrepancy between the distributions tends to zero. Therefore, we assume that if the resulting divergence value can be rounded to zero, then the attack is not carried out. Further, in order to determine the boundaries of the divergence, denial-of-service attacks were conducted with varying degrees of intensity. And also carried out the Black-Hole attack.

4. The next step is to define the boundaries of the values that will signal an attack. The next step is to define the boundaries of the values that will signal an attack. For each site on which the normal distribution is built, certain limits of confidence intervals are characteristic, respectively, the greater the deviation of the obtained value from the confidence interval, the greater the likelihood of an attack. In a normal situation, the divergence should go to zero, but deviations in the network are still possible, therefore, the limit value that corresponds to the nominal value is 0.2. The maximum deviation in behavior that can be fixed, based on the fact that the value of the normal distribution varies from 0 to 1, equal to 4.6. The maximum deviation in behavior that can

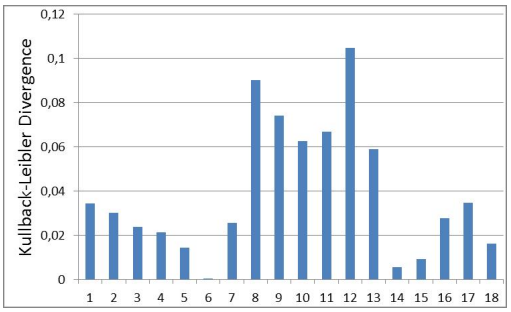

Figure 3: Histogram showing the result of calculating the Kullback-Leibler divergence for the nodes during normal network operation

be fixed, based on the fact that the value of the normal distribution varies from 0 to 1, equal to 4.6. A deviation greater than 0.15 already indicates a significant difference in the received distributions and may become the first sign of an attack. When the value of divergence tends to a value from 0.5-1, this already indicates a significant change in the type of normal distribution. Of course, single deviations can be observed, an important condition is the presence of at least three consecutive intervals with similar deviations. Thus we define the conditions for fixing the attack.

$$\begin{cases} f(s)_{(\Delta t-1)}, & f(s)_{\Delta t}, & f(s)_{(\Delta t+1)} > 0,2; \\ f(r)_{(\Delta t-1)}, & f(r)_{\Delta t}, & f(r)_{(\Delta t+1)} > 0,01; \\ f(f)_{(\Delta t-1)}, & f(f)_{\Delta t}, & f(f)_{(\Delta t+1)} > 0,01. \end{cases} \tag{8}$$

At the same time indicators for the remaining packages may not change. The main thing is to increase the number of sent packets. The situation is more complicated with the victim node. When modeling denial-of-service attacks with varying degrees of intensity, the victim node experienced significant changes in metrics, received packets, as well as redirected packets. The growth of redirected packets can be observed when a denial of service attack is directed not at one node but against a group of nodes, and a side effect is the need to redirect nodes to a large number of packets. We define the conditions that allow us to determine the node-victim.

$$\begin{cases} f(s)_{(\Delta t-1)}, & f(s)_{\Delta t}, & f(s)_{(\Delta t+1)} \leqslant 0,2; \\ f(r)_{(\Delta t-1)}, & f(r)_{\Delta t}, & f(r)_{(\Delta t+1)} > 0,2; \\ f(f)_{(\Delta t-1)}, & f(f)_{\Delta t}, & f(f)_{(\Delta t+1)} > 0,2. \end{cases} \tag{9}$$

4.2.1 Evaluation of attack detection

Next, we simulated attacks with varying degrees of intensity on a group of robotic devices. An attack with a low degree of intensity is more difficult to detect, since it does not greatly affect the traffic pattern. Nevertheless, such an attack was discovered by the developed methods within 30 seconds from the beginning of its implementation. Figure 4 shows the calculation of Kullback-Leibler divergence for normal traffic and for an attacker who is conducting a low-intensity attack.

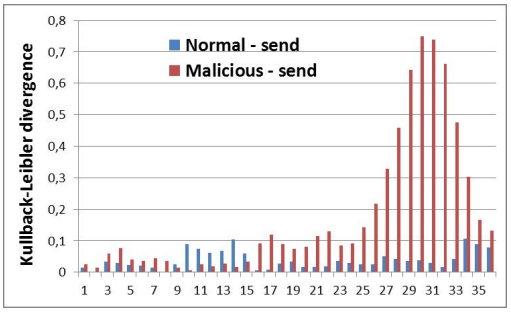

Figure 4: Histogram showing the result of calculating the Kullback-Leibler divergence for the nodes during light denial of service attack for normal node and for malicious

The figure shows that there is an increase in the deviation between the normal distributions at the attacker's node and the maximum value reaches 0.7. Moreover, since the attack is not very intense, that is, the node sends not many more packets than during normal operation, a value close to 0.2 is observed for a long time.

Then an attack of medium intensity was carried out, while it had an impact on the network. The node began to receive and redirect more packets and wasted energy accordingly. Figures 5 show the result of calculating the divergence for sent and received packets for both the victim and the attacker. And also for comparison, the calculation of divergence for a normal network node is given.

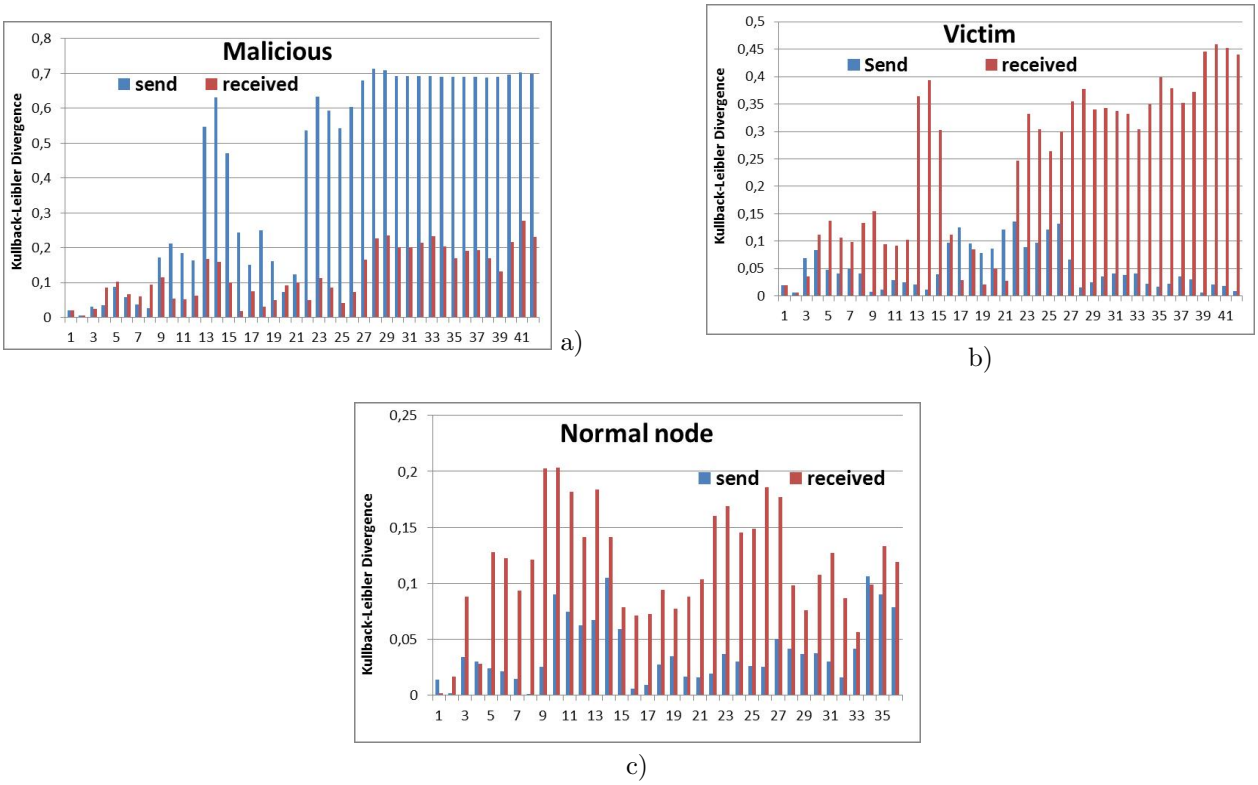

Figure 5: Histogram showing the result of calculating the Kullback-Leibler divergence for the nodes during medium intense denial of service attack for (a) malicious node (b) victim node (c) normal node

The first histogram depicts the deviation between normal distributions for the attacker. Attack is detected much faster than light attack. From the histogram there is an increase in the sent packets and the maximum value is also fixed equal to 0.7. It can be said that this is a clear sign of this attack, the slow growth of the divergence value, which is fixed at 0.2 and then reaching the value of 0.7 and fixing on it. These changes can be used as necessary for signature analysis, neural network training and classifiers. In this case, as can be seen from the figure for the attacker, the growth of the packets received was not observed, as was assumed under the given conditions. Next, consider what happened to the victim. If you carefully compare the figures (a) and (b), then you can notice some symmetry. The histograms are very similar, only for the victim's site are similar deviations recorded for received packets. In this case, the level of sent packets was within the acceptable limits of the established norm. Special attention should also be paid to the divergence of a normal node. First, for sent packets, it does not exceed 0.1 in the general case. This is due to the fact that the node sends an approximately equal number of packets each interval with small deviations, which do not significantly affect the form of the nominal distribution. As for the received packets, then the deviations reach 0.2. This is generally acceptable. It should be noted that this situation arises from the fact that a node receives packets from different nodes in different periods of time, which may not coincide, so there may be a slight difference between the distributions of a given value. However, these gaps are minimal and uniformly overall. The pattern of deviation of the indicator is clearly different from the previous two figures.

5 Conclusion

Thus, it should be noted that in this work an intrusion detection system based on a network node and a method that allows detecting anomalies was presented. The main idea of this method is that by analyzing itself a node can detect non-standard behavior. As it was proved by an experimental study, a node can fix deviations between probability distributions built at different time intervals and find out that it is conducting an attack or is a victim. Due to the fact that the normal distribution is used, the collected statistics are smoothed out and small jumps in the collected data are ignored. This reduces the occurrence of errors associated with his refusal. Moreover, when constant changes occur and an attack is clearly carried out, the type of distribution begins to change and fix the anomaly is easy. During the attacks, the simulation lasted 200 seconds, the first 100 seconds of the attack did not take place and the node behaved normally, and then the attack began. This study will be expanded and conducted for other types of attacks. In addition, it is planned to increase the number of analyzed parameters.

5.0.1 Acknowledgements

This work was partially supported by the Russian Foundation for Basic Research No.17-07-00106.

References

[Hag16] M. Hagele. Robots Conquer the World [Turning Point]. *IEEE Robotics & Automation Magazine.* 23(1):120 118. 2016.

[Bas17] A.S. Basan, E.S. Basan, O.B. Makarevich. Analysis of Ways to Secure Group Control for Autonomous Mobile Robots. *Proceedings of 10th International Conference On Security Of Information And Networks (SIN 2017).* 1-6. 2017.

[She14] S. Shetty, T. Adedokun, and L.-H. Keel. Cyberphyseclab: A testbed for modeling, detecting and responding to security attacks on cyber physical systems. *2014 ASE BIGDATA/SOCIALCOM/CYBERSECURITY Conference*Stanford University. 1-9. 2014

[Par09] M. J. Parsons, P. Ebinger. Performance evaluation of the impact of attacks on mobile ad hoc networks. *Proceedings : Field Failure Data Analysis; Embedded Systems and Communications Security, in conjunction with 28th IEEE International Symposium on Reliable Distributed Systems* Niagara Falls, New York, U.S.A. 40-48. 2009.

[Sigh15] M. Singh, Md. A. Khan, V. Singh, A. Patil, S. Wadar. Attendance management system. *2nd International Conference on Electronics and Communication Systems (ICECS).* 418 - 422. 2015.

[Mill13] J. Milliken; V. Selis, K. M. Yap. Impact of Metric Selection on Wireless DeAuthentication DoS Attack Performance. *IEEE Wireless Communications Letters.* 2(5): 571 574. 2013.

[Vilch18] V. M. Vilches, L. A. Kirschgens, A. B. Calvo. A.H. Cordero , R. I. Pison, et al. Introducing the robot security framework (RSF),a standardixed methodology ti perform security assessments in robotics. *Symposium on Blockchain for Robotic Systems.* MIT Media Lab, Cambridge, United States. 1-19. 2018.

[Wei99] L. Wei, and D. Farinacci. Multicast Routing Monitor (MRM). *IETF Internet-Draft, draft-ietf-mboned-mrm-.txt.* 1-22. February 1999.

[Pshikh11] V. Pshikhopov, A. Ali. Hybrid motion control of a mobile robot in dynamic environments. *Proc. of the 2011 IEEE International Conference on Mechatronics.* Istanbul, Turkey. 540545. April 2011.

[Bas18] A.S. Basan, E.S. Basan, O.B. Makarevich. Method of Detecting and Blocking an Attacker in a Group of Mobile Robots. *Advances in Intelligent Systems and Computing.* Vol. 875. 340-349.

[Fag14] A. Fagiolini, G. Dini, A. Bicchi. Distributed Intrusion Detection for the Security of Industrial Cooperative Robotic Systems. *Proceeding of the 19th World Congress The International Federation of Automatic Control Cape Town.* South Africa.7610-7615. 2014.

[Grjib08] A. Grjibovski. Data types, control of distribution and descriptive statistics. *Human ecology. Publisher: Northern State Medical University (Arkhangelsk).* ISSN: 1728-0869. 52-58. 2008.