ISSN 2347 - 3983
Volume 8. No. 5, May 2020
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter90852020.pdf
https://doi.org/10.30534/ijeter/2020/90852020

Detection of Slow DDoS Attacks based on User's Behavior Forecasting

Vitalii Savchenko[1], Oleh Ilin[2], Nikolay Hnidenko[3], Olga Tkachenko[4], Oleksander Laptiev[5], Svitlana Lehominova[6]

[1]Director of Cybersecurity Institute, State University of Telecommunication, Kyiv, Ukraine, savitan@ukr.net
[2]Professor of Computer Sciences Department, State University of Telecommunication, Kyiv, Ukraine, oleg.ilin@dut.edu.ua
[3]Professor of Computer Sciences Department, State University of Telecommunication, Kyiv, Ukraine, askorpam@ukr.net
[4]Head of Computer Engineering Department, State University of Telecommunication, Kyiv, Ukraine, okar@ukr.net
[5]Professor, Cybersecurity Technical Department, State University of Telecommunication, Kyiv, Ukraine, alaptev64@ukr.net
[6]Head of Cybersecurity Management Department, Cybersecurity Institute, State University of Telecommunication, Kyiv, Ukraine, chiarasvitlana77@gmail.com

ABSTRACT

The article deals with a problem of detecting low and slow distributed denial of service (DDoS) attacks. It is widely known that the detection of slow DDoS attacks differs significantly from volume based attacks, because slow attacks do not increase the intensity of traffic in the network. An assumption about dependency of slow attack from user's behavior is made. A method for detecting such attacks based on research and forecasting of the individual behavioral trajectory of a particular user is proposed. Possibilities of application of such method are proved on the basis of modeling RUDY attacks to HTTP services. The characteristics of forecasting accuracy depending on the accumulated traffic and attack statistics are shown. It is concluded that such method can be used to detect different types of slow DDoS attacks.

Key words : individual prediction, random process, slow DDoS attack, user behavior.

1. INTRODUCTION

Ease to find network vulnerabilities have made DDoS attacks widespread. For effective counteraction to such attacks, two basic measures are required: 1) to diagnose an attack at the earliest stages; 2) to separate a malicious traffic from a normal one after detection. Understanding which user requests are resulted by DDoS attack you can make appropriate settings for firewalls, routers or implement other security measures within network. Parameters of slow DDoS attacks mostly depend on particular user's behavior and for an effective mitigation such user's behavior should be predicted in details. So, forecasting of user's behavior based on traffic parameters becomes the main step in mitigating of slow DDoS intrusions.

1.1 Problem Statement

The feature of slow DDoS attack is the use of vulnerability in TCP protocol where interruptions can be caused either intentionally or unintentionally as a result of delays in the communication channel. As slow DDoS attacks do not cause a sharp increase of traffic, which lead to an instant server's denial of service, so detecting the moment the attack start is almost impossible and the separation of malicious traffic from normal traffic is a significant problem. Therefore, to identify and recognize slow DDoS attacks, it is necessary to develop other approaches and methods. The main problem in detecting slow DDoS attacks is the inability to prevent them, since the determination process is based on the study of existing traffic without the possibility of predicting it depending on users' activity. No doubt, predicting user behavior would make it possible to detect abnormal behavior and prevent the appearance of slow DDoS attacks.

1.2 Related Works Overview

There are a huge number of publications on the detection of slow DDoS attacks. A. Dhanapal and P. Nithyanandam [1] propose a new method for detecting slow HTTP attacks in the cloud. The solution allows to detect Slow HTTP Header Attacks (Slowloris), Slow HTTP Body Attacks (RUDY) or

Slow HTTP Read Attacks. Another their work [2] proposes a new classification model for mitigating attacks in the cloud. At the same time, such approaches do not guarantee the effective detection of attacks in the early stages of their development.

Th. Lukaseder et al. [3] introduce a system that can detect and mitigate attacks within the network infrastructure. This work is continued by article of H. Abusaimeh et al. [4] which explores a Side-Channel protection model. The main identification parameters in both models are the packet transfer rate and the uniform distance between the packets which does not allow to preempt the actions of attackers. C.L. Calvert and T.M. Khoshgoftaar [5] are considering data sampling to create different class distributions to counteract the effects of the highly unbalanced slow HTTP DoS datasets. At the same time, a significant number of samples (authors use 1.89 million copies of attacks) is quite difficult to achieve in reality. A study of B. Cusack and Z. Tian [6] develops a metric based system to detect traditional slow attacks that can be effective with limited resources based on similarity research and the introduction of the Euclidean metric. This approach is quite effective only if there are a large number of such samples of slow attacks, and with a large variety of such an approach is unlikely to be effective.

Publication of Ie. V. Duravkin et al. [7] defines the quality parameters of TCP connections typical for slow HTTP attacks. The derived formulas estimate the probability and time of the web server transitioning to the overload. Despite a detailed study, such attack detection is based on observation statistics and does not address forecasting. The article of I. V. Ruban et al. [8] proposes an algorithm for detecting slow DDoS attacks based on traffic patterns depending on the server load status. At the same time, the process of making decisions is not considered. The work of Ya. V. Tarasov [9] reviews various scenarios and proposes a hybrid neural network to detect DDoS attacks. But, the method and general technique for detecting low-intensity DDoS attacks are not considered.

Y. M. Krakovsky and A. N. Luzgin [10] consider interval prediction based on a probabilistic neural network with dynamic updating of the smoothing parameter. But the problem of model dynamics remains unsolved. The article of S. Lysenko and V. Tkachuk [11] presents a new method for detecting RUDY DDoS attack based on the self-similarity of network traffic. The work does not take into account the diversity of training samples and the training set obtaining process. M. Idhammad et al. [12] presents a system for detecting HTP DTP attacks in the cloud, based on information entropy indicators and random trees. This approach is quite effective, although it does not address the issues of predicting the development of an attack.

Thus, the majority of works devoted to counteracting slow DDoS attacks do not address the issues of predicting user behavior and therefore are not effective enough to detect attacks in the early stages. The **aim** of this work is to form a system for detecting slow DDoS attacks based on predicting the behavior of network users. To successfully solve the identified problem, it is necessary to build a model and technology for predicting user behavior, taking into account the history of their interaction with the server, and also offer a topology for recognizing slow DDoS attacks.

2. DEVELOPMENT OF INTRUSION DETECTION SYSTEM FOR SLOW DDOS ATTACKS

2.1 Traffic Parameter for Slow DDoS Attack Detection

The main components of the slow DDoS attack detection system were proposed in Ie. V. Duravkin et al. [7]. The architecture of such a system should consist of four modules: 1) traffic collection; 2) calculation of traffic parameters; 3) calculation of network statistics; 4) attack classifier.

The workflow of such system should consist of the following steps:

1. The traffic collection module for a certain period of time registers the traffic parameters necessary for further calculations: IP addresses of the sender and receiver; TCP window size; package arrival time.

2. In the module for traffic parameters calculating for each IP address the following traffic characteristics are calculated, e.g.
the average delay between transmitted packets

$$\overline{T} = \frac{\sum_{i=1}^{k} (t_{i+1} - t_i)}{k-1}, \quad (1)$$

where:

t_i – the i-th package arrival time;

t_{i+1} – the i+1-th package arrival time;

k – the number of packets received during the analyzed period.

The built-in timer allows to record the beginning and end of the session, which makes it possible to track the duration of open connections. Other parameters can also be used for forecasting, depending on the system settings.

3. In the attack classification module, a decision about the existence of a possible slow HTTP attack can be made after comparison of the obtained indicators with threshold values.

With this approach, a decision on the presence or absence of a slow DDoS attack can be made only after collecting enough statistics. At the same time, in such a situation, protection systems often do not have time for proactive actions. To solve this problem, the decision on the presence of a slow DDoS attack should be based on a forecast of user behavior, which can be generated based on a study of statistics of similar actions by other users. Thus, it is advisable to add a situation forecast block to the considered action algorithm.

2.2 User's Behavior Forecasting

The user's behavior in the network forms an individual trajectory of changes in the traffic parameters of this user. Such trajectories will be typical for both normal traffic and in the case of a slow DDoS attack. To determine the appropriate time to start neutralizing a slow DDoS-attack, it is necessary to solve the problem of individual prediction of its time trajectory. Prediction of traffic parameters by individual trajectory has already been studied in research of V. Savchenko et al. [13], in which traffic parameters were determined at long intervals (week, month). This approach was also partially used for protection of information in social networks (V. Savchenko et al. [14]) and for forecasting in a multi-agent environment (P. Shchipansky et al. [15]). At the same time, the accuracy of the system in recognizing a slow DDOS-attack should be much higher and therefore this approach needs to be improved.

In this case, the input data can be used to observe the traffic parameters, e.g. average time interval between transmitted packets, the delay between packets in the session etc., which form a vector of parameters $X = (X_1, X_2, ..., X_H)$. Fulfillment of the condition $X \in S_0$, where S_0 is the tolerance area of the vector *X*. The random process $X(t)$ is formed with time intervals between packets or delays between packets in time and describes the evolution of network parameters over time. Assume also that the process $X(t)$ is statistically determined at $t \geq t_1$, where t_1 is the moment of the beginning of observations. Figure 1 shows a specific trajectory ω and the control moment of observation $t_k \geq t_1$.

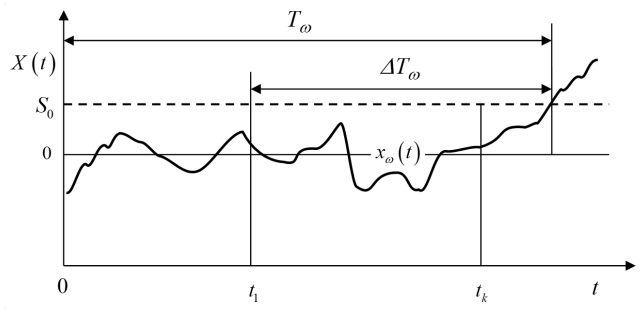


Figure 1: Traffic forecasting

Information about the beginning of the trajectory of the traffic parameter ω can be specified as $x_\omega(t) \in S_0, t_1 \leq t \leq t_k$ and obtained from observations of the traffic parameter $X(t)$. In this case, the problem of individual prediction of the traffic parameter behavior is formulated as the problem of determining the a posteriori distribution of the process output time T_ω. $X(t)$ outside the tolerance area S_0 regarding implementation $x_\omega(t)$ can be presented as a task

$$P^{ps}(s) = P\{X(s) \in S_0 / x_\omega(t)\}, t_1 \leq t \leq t_k, s \geq t_k. \qquad (2)$$

Formula (2) gives a probability that a particular trajectory of the parameter ω guaranteed to fall within the allowable range $s > t_k$, if by the time t_k including its condition was described as $x_\omega(t), t_1 \leq t \leq t_k$. That is, the problem of individual trajectory forecasting of the network traffic parameter is solved. To solve the problem of forecasting the researched process should be represented by the formula

$$X(t) = m(t) + \sum_v V_v \phi_v(t), \qquad (3)$$

where $m(t)$ – mean function of the process;

$\phi_v(t)$ – non-random (coordinate) time functions;

V_v – random, uncorrelated coefficients ($M[V_v] = 0$, $M[V_v, V_\mu] = 0$, $v \neq \mu$).

This view, proposed in [13] allows you to apply it to any traffic parameter that can be represented as a time series. The process $X(t)$ can be written as a random sequence $X(t_i) = X(i), i = \overline{1,I}$ in a discrete series of observations t_i:

$$X(i) = m(i) + \sum_{v=1}^{i} V_v \phi_v(i), i = \overline{1,I}, \qquad (4)$$

where V_v – random coefficient with parameters $M[V_v] = 0$, $M[V_v, V_\mu] = 0$, $v \neq \mu$; $M[V_v^2] = D_v$;

$\phi_v(i)$ – non-random coordinate function, $\phi_v(v) = 1$, $\phi_v(i) = 0$ while $v > i$.

Formulas for variance and correlation function can be written as

$$D(i) = \sum_{v=1}^{i} D_v \phi_v^2(i), i = \overline{1,I}; \qquad (5)$$

$$D(i,j) = \sum_{v=1}^{inf(i,j)} D_v \phi_v(i) \phi_v(j), \; i,j = \overline{1,I}. \qquad (6)$$

Therefore, the representation of random processes of traffic parameters (2) allows to solve the problem of detecting a slow DDoS-attack based on predicting user behavior.

2.3 Algorithm for Detecting a Slow DDoS Attack based on User's Behavior Forecasting

1. Determine the characteristics of an a priori random process $X(t)$ in the form of an equation (4) on a discrete series of points $t_i, i=\overline{1,I}$. For this purpose it is necessary to generate results of parameter control $x(\mu), \mu=\overline{1,k}$ in the form of a time series where the results of the observations will be correlated with the moments of time $t_\mu, \mu=\overline{1,k}, k<I$. Based on the results of such control a forecast of the a posteriori trajectory $X^{ps}(t)$ arises from the a priori $X(t)$ taking into account the observations of the parameter.

2. Establish the value of the implementation process $x(1)$, obtained as a result of control at the time $\mu=1$, and presenting as

$$x(1)=m(1)+v_1. \quad (7)$$

3. Concretize the meaning v_1 of random coefficient V_1 by formula (7), which corresponds to the result of the first observation. Concretization of meaning V_1 leads to a change in the density of the distribution of the remaining coefficients $V_i, i=\overline{2,I}$.

4. Establish the type of a posteriori random process that is at the moment $i=1$ pass through a point $x(1)$ by substituting the value V_1 from (7) to the formula (4):

$$X^{(1)}(i)=m(i)+\left(x(1)-m(1)\right)\phi_1(i)+\sum_{v=2}^{i}V_v\phi_v(i), i=\overline{1,I}. \quad (8)$$

5. Determine the mean function for the process (8) that is at the moment $i=1$ pass through a point $x(1)$:

$$m^{(1)}(i)=m(i)+\left(x(1)-m(1)\right)\phi_1(i), i=\overline{1,I}, \quad (9)$$

6. Establish a general dependency for the process that is currently passing through the point $x(1)$:

$$X^{(1)}(i)=m^{(1)}(i)+\sum_{v=2}^{i}V_v\phi_v(i), i=\overline{1,I}. \quad (10)$$

7. If $\mu=k$ then go to step 9; otherwise – to the next step.

8. Establish the value of the implementation process $x(2)$, obtained as a result of control at the time $\mu=2$, and with (10) present it in the form:

$$x(2)=m^{(1)}(2)+v_2.$$

Return to step 3. Repeat operations, as for the case $\mu=1$, and get

$$m^{(2)}(i)=m^{(1)}(i)+\left(x(2)-m^{(1)}(2)\right)\phi_2(i), i=\overline{1,I}; \quad (11)$$

$$X^{(2)}(i)=m^{(2)}(i)+\sum_{v=3}^{i}V_v\phi_v(i), i=\overline{1,I}. \quad (12)$$

9. Determine the extrapolation operator for the mean function of the a posteriori random process and an arbitrary number $k<I$ moments of control:

$$m^{(0)}(i)=m(i), i=\overline{1,I},$$

$$m^{(k)}(i)=m^{(k-1)}(i)+\left(x(k)-m^{(k-1)}(i)\right)\phi_k(i), i=\overline{1,I}; \quad (13)$$

$$X^{(k)}(i)=m^{(k)}(i)+\sum_{v=k+1}^{i}V_v\phi_v(i), i=\overline{1,I}. \quad (14)$$

Thus, formulas (13) - (14) fully describe a linear a posteriori process in which (13) is the mean function of this process at points t_i.

10. Construct the traffic parameter forecast and determine the moment when the parameter goes over the critical values. Classify the traffic as a slow DDoS attack and implement security measures.

11. The end of the algorithm.

Thus, formula (13) allows to optimally solve the problem of extrapolation of the process, and formula (14) - to reproduce the a posteriori random process on the basis of modeling. The specified linear analytical model of a posteriori random process on the basis of such representation allows to solve problems of forecasting of network traffic parameters. The decision to have a slow DDoS attack is has to be made for each sender's IP address based on a comparison of the predicted parameters with the critical values to determine the time of entry of the parameter into the zone of critical values. Such approach takes into account the statistics of the particular user's behavior as well as other users' behavior in similar circumstances in the case of a slow DDoS attack.

3. MODELING OF ALGORITHM FOR DETECTING SLOW DDOS ATTACKS BASED ON USER'S BEHAVIOR FORECASTING

Simulation of slow DDOS attack detection based on user behavior forecasting was performed for RUDY attack. For simplicity, only one case of attack was considered against the background of normal traffic as it is shown in Figure 2. The average delay between transmitted packets is considered as the investigated parameter.

RUDY is an attack on a network server designed to cause a web server to crash by sending long requests. The attack is performed using tool that scans the target website and detects embedded web forms.

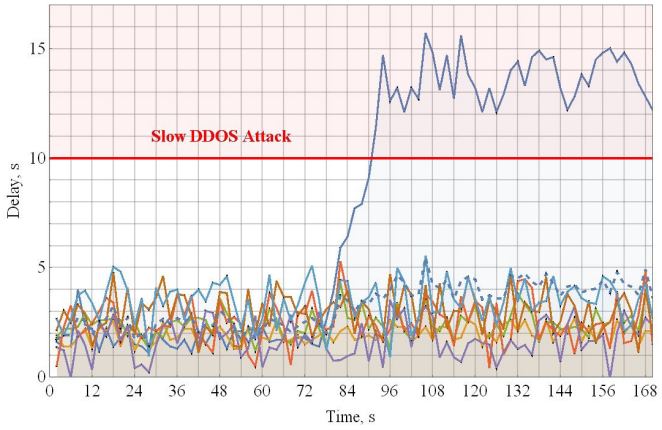


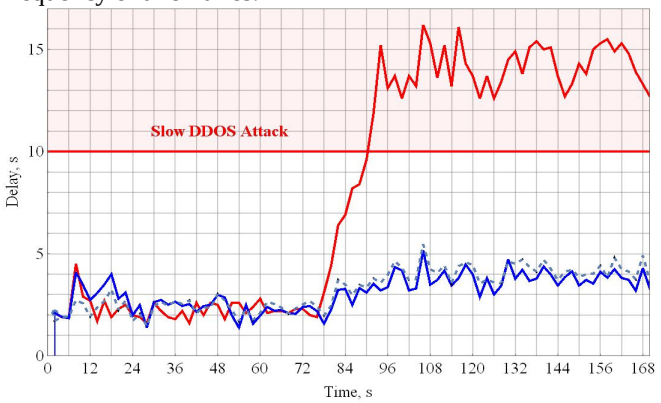
Figure 2: Traffic patterns

Once the forms were discovered, RUDY sends legitimate HTTP POST-requests with an abnormally long content-length header field, after which it begins to enter information by one byte per packet. This type of attack is difficult to detect because of miserable fluctuations in incoming traffic.

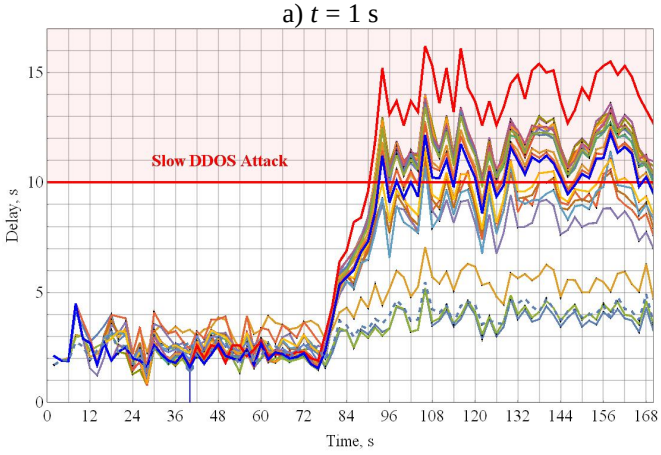
To the process shown in Figure 2, the forecasting technique (4) - (14) was applied, taking as initial values of observations individual points of the time series, which correspond to the partial trajectory № 1 (Figure 2, the blue curve). Taking this curve as a control, as the initial data of observations were taken the first values of the time series, which correspond to t = 1, 40, 90 s observations.

Figure 3a shows the results of forecasting at t = 1 s. Few number of initial control data can only reproduce the process as a whole (mean process curve), but the specific values of the predicted traffic will be very different from the real ones (control trajectory). So, knowing the average parameters of network traffic and the entry point to the forecast, you cannot accurately predict the future behavior of the system. So the method "selects" the required trajectory depending on the entry point and the average trajectory.

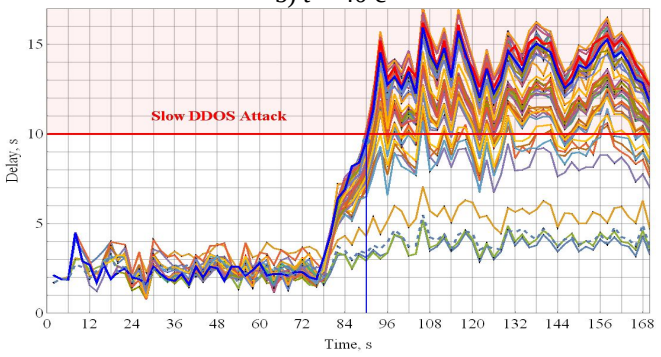
Increasing the number of observations to t = 40, 90 s (Figure 3b) increases the reliability of further prediction and at t = 90 s we can talk about a fairly accurate prediction $P^{ps}(s) = P\{X(s) \in S_0 / x_\omega(t)\} \geq 0{,}99$. In Figure 3b and 3c curves of other colors show how forecasting will be carried out when receiving data from other control points $t_\mu, \mu = \overline{1,k}, k < I$, preceding the moment t_k. That is, the probability of error in choosing the correct trajectory depends on the amount of initial data observed. It is logical to assume that in this case the accuracy of prediction will depend too much on the characteristics of the behavior of the trajectory, which lead to anomalous traffic, as well as on the observed frequency of anomalies.



a) t = 1 s



b) t = 40 c



c) t = 90 c

Figure 3: User's behavior forecasting with observation time t = 1, 40, 90 s:
— forecast value; — compared value; --- mean value

For this model example, the question of the required number of observations and the accuracy of predicting the trajectory of the attack is interesting. Table 1 shows the increase in prediction accuracy with number of user's behavior control points that reaches <1% at 90 s. Thus, the simulation results confirm the adequacy of the predictive model for determination of a slow DDoS attacks based on the individual user's behavior forecasting. In this case, the random process is precisely determined at the control points and provides a minimum of the mean square error in the intervals between these points.

Table 1: The average deviation of the forecast from the control implementation

Observation time, s	1	10	20	30	40	50	60	70	80	90
Deviation, %	86	64	44	28	19	12	6	3	2	< 1

4.CONCLUSION

1. Slow DDoS attacks are becoming more common due to the simplicity of implementation and the complexity of their detection. Detection of an attack by existing methods is ineffective due to the delayed nature of the response to the attack in the case of observation and analysis of traffic parameters. A more promising approach is to predict user behavior based on the statistics of previous attacks.

2. Forecasting individual user behavior provides a solution to the problem of detecting slow DDoS attacks based on the algorithm for finding unknown future values for a time series of traffic parameters. The proposed method combines the advantages of artificial intelligence and statistical analysis and is capable of self-learning in the case of attack statistics supplementing. This approach makes it possible to accurately determine the random process at the control points and provide a minimum of the mean square of the approximation error in the intervals between these points.

3. The direction of further research in the field of counteraction to slow DDoS attacks can be a wide range of issues to improve the method to enable the possibility of forecasting at intervals beyond the available statistics including conditions of high noise data or their partial absence.

REFERENCES

1. A. Dhanapal and P. Nithyanandam. **The Slow Http Distributed Denial of Service Attack Detection in Cloud.** *Scalable Computing: Practice and Experience*. Volume 20, Number 2, pp. 285–298, 2019. https://doi.org/10.12694/scpe.v20i2.1501
2. A. Dhanapal_and P. Nithyanandam. **The Slow HTTP DDOS Attacks: Detection, Mitigation and Prevention in the Cloud Environment.** *Scalable Computing: Practice and Experience*. Volume 20, Number 4, pp. 669–685, 2019. https://doi.org/10.12694/scpe.v20i4.1569
3. T. Lukaseder, S. Ghosh, F. Kargl. **Mitigation of Flooding and Slow DDoS Attacks in a Software-Defined Network**. 16 August 2018. https://arxiv.org/pdf/1808.05357.pdf
4. H. Abusaimeh, H. Atta, H. Shihadeh. **Survey on Cache-Based Side-Channel Attacks in Cloud Computing.** *International Journal of Emerging Trends in Engineering Research*. Volume 8, No.4, p.1019-1026, April 2020.
5. C. L. Calvert, T. M. Khoshgoftaar **Impact of class distribution on the detection of slow HTTP DoS attacks using Big Data**. *Journal of Big Data.* 6, 67, 2019. https://doi.org/10.1186/s40537-019-0230-3
6. B. Cusack, and Z. Tian. **Detecting and tracing slow attacks on mobile phone user service**. In Valli, C. (Ed.). *The Proceedings of 14th Australian Digital Forensics Conference*, 5-6 December 2016, Edith Cowan University, Perth, Australia. pp. 4-10, 2016.
7. Ie. V. Duravkin, A. Carlsson, A. S. Loktionova. **Method of Slow-Attack Detection.** *Information processing systems*, issue 8 (124), pp. 102-106, 2014.
8. I.V. Ruban, D.W. Pribylnov, E.C. Loshakov. **A method of detecting a low-speed denial-of-service attack.** *Science and technology of the Air Force of the Armed Forces of Ukraine*, № 4(13). 85-88, 2013.
9. Ya. V. Tarasov. **Investigation of the application of neural networks for the detection of low-intensity DDoS-attacks of the application level.** *Cybersecurity issues* №5(24), 23-29, 2017. https://doi.org/10.21681/2311-3456-2017-5-23-29
10. Y. M. Krakovsky, A. N. Luzgin. **The cyberattack intensity forecasting to information systems of critical infrastructures.** *Problems of smart cities and sustainable development of territories*. SAFETY2018, Ekaterinburg, October 4-5, 34-42, pp. 180-187, 2018.
11. S. Lysenko, V. Tkachuk. **Method and software for detecting r.u.d.y. attack based on the usage of the algorithm of determining traffic self-similarity.** *Herald of Khmelnytskyi national university*, Issue 3, p. 273, 2019.
12. M. Idhammad, K. Afdel, and M. Belouch. **Detection System of HTTP DDoS Attacks in a Cloud Environment Based on Information Theoretic Entropy and Random Forest.** *Security and Communication Networks*. Volume 2018, Article ID 1263123, 13 p. https://doi.org/10.1155/2018/1263123
13. Vitalii Savchenko, O. Matsko, O. Vorobiov, Y. Kizyak, L. Kriuchkova, Y. Tikhonov, A. Kotenko. **Network traffic**

forecasting based on the canonical expansion of a random process. *Eastern European Journal of Enterprise Technologies*. VOL 3, NO 2 (93). p. 33-41, 2018.
https://doi.org/10.15587/1729-4061.2018.131471

14. Vitalii Savchenko, V. Akhramovych, A. Tushych, I. Sribna, I. Vlasov. **Analysis of Social Network Parameters and the Likelihood of its Constraction**. *International Journal of Emerging Trends in Engineering Research*. Volume 8, No. 2, p. 271-276, February 2020. https://doi.org/10.30534/ijeter/2020/05822020
15. Pavlo Shchypanskyi, Vitalii Savchenko, Oleksii Martyniuk, Ihor Kostiuk. **Air Defense Planning from an Impact of a Group of Unmanned Aerial Vehicles based on Multi-Agent Modeling.** *International Journal of Emerging Trends in Engineering Research*. Volume 8. No. 4, p. 1302-1308, April 2020.