https://www.sworldjournal.com/index.php/swj/article/view/swj21-01-037
DOI: 10.30888/2663-5712.2023-21-01-037

UDC 004.056.53:[004.7:004.032.26]

INVESTIGATION OF THE POSSIBILITY OF USING NEUROFUZZY NETWORK TO DETERMINE THE EXTENT OF DoS ATTACK

ДОСЛІДЖЕННЯ МОЖЛИВОСТІ ВИКОРИСТАННЯ НЕЙРОНЕЧІТКОЇ МЕРЕЖІ ДЛЯ ВИЗНАЧЕННЯ СТУПЕНЯ ЗДІЙСНЕННЯ DoS АТАКИ

Pakhomova Victoria / Пахомова Вікторія
c.t.s., as.prof. / к.т.н., доц.
ORCID: 0000-0002-0022-099X
Kovalov Rodion / Ковальов Родіон
second degree holder / здобувач другого ступеня
ORCID: 0009-0005-3780-8550
Ukrainian State University of Science and Technology,
Dnipro, Lazaryan, 2, 49010
Український державний університет науки і технологій,
Дніпро, Лазаряна, 2, 49010

Abstract. *As a research method, ANFIS configurations 4-5-8-16-16-1 were used, where 4 is the number of input neurons; 5 – total number of layers; 8 – the number of neurons of the first hidden layer; 16 – the number of neurons of the second hidden layer; 16 – the number of neurons of the third hidden layer; 1 – the number of resultant neurons created using the Fuzzy Logic Toolbox of the MatLAB system, the resulting characteristic is the degree of confidence that the DoS attack occurred at the following terms: low; medium; high. Using the open database of NSL-KDD network traffic parameters on the created ANFIS, a study of its error at different affiliation functions on samples of different lengths was carried out using different methods of training optimization. It is determined that the smallest value of the ANFIS error was based on the use of the multiparameter Bell function by the Hybrid learning optimization method, and it is enough to have a training sample of 70 examples.*

Keywords: *DoS attack, traffic, NSL-KDD, ANFIS, Bell function, error.*

Introduction

Formulation of the problem. The creation of an effective network attack detection system requires the use of qualitatively new approaches to information processing, which should be based on adaptive algorithms capable of self-learning. The most promising direction in the creation of similar network attack detection systems is the use of neural network technology.

Analysis of the latest research. A review of scientific sources [1-3, 6-8] showed that the following neural networks can be used to detect DoS attacks: Multi Layer Perceptron (MLP); Kohonen network or Self-Organizing Map (SOM); Radial Basis Function network (RBF); Adaptive Network Based Fuzzy Inference System (ANFIS). In addition, none of the methods provides a complete guarantee of detection of DoS attacks, while different neural networks detect different network classes of attacks in different ways.

The purpose of the article is development of a methodology for determining the degree of DoS attack using neurofuzzy technology. In accordance with the goal, the following tasks are set: creation of a neurofuzzy network; determination of optimal parameters on the created neurofuzzy network.

1. Statement of the problem and mathematical apparatus

The category of DoS (Denial of Service) attacks includes a wide range of methods

that are aimed at ensuring that the resource is unavailable to legitimate users. These attacks are used to overload the network or servers, resulting in reduced functionality and availability. The following network attack classes fall into the DoS category: Back; Land; Neptune; Pod; Smurf; Teardrop. As a mathematical apparatus, the ANFIS system, combining the methods of a neural network and the Takagi-Sugeno fuzzy inference logic system, the structure of which is shown in Figure 1.

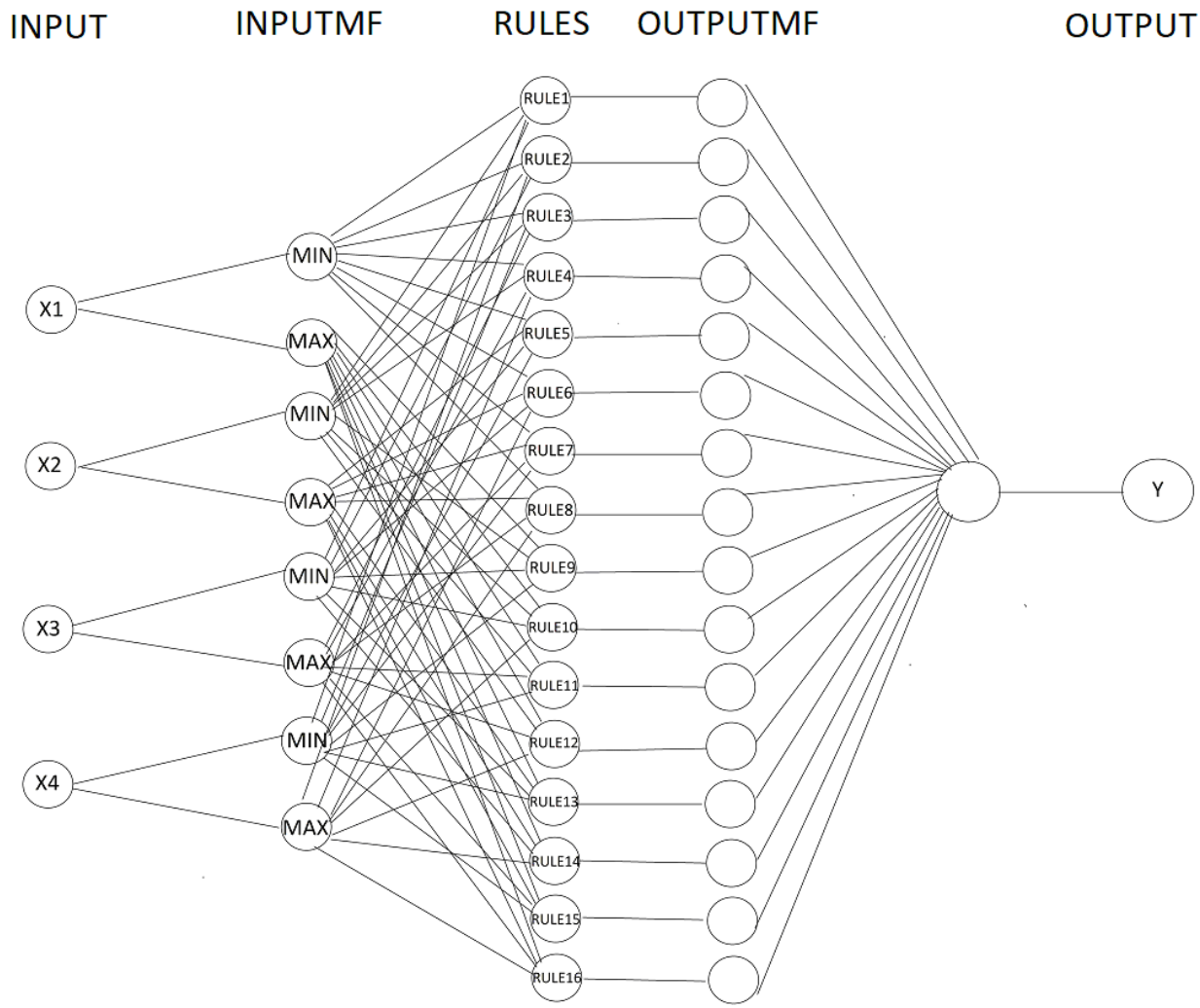

Figure 1 – ANFIS configuration 4-5-8-16-16-1

Authoring

The input neurons of the first layer are the following parameters [5]: X1 (dst_host_count) – the number of connections installed on the host computer to which the attack was directed; X2 (dst_host_srv_count) – the number of unique services available on the host computer targeted by the attack; X3 (count) – the number of requests sent by the attacker within a short time; X4 (srv_count) – the number of requests for a specific service sent by the attacker within a short time.

The second layer (inputmf) has 2*4=8 neurons; two terms (minimum and maximum value) to each of the neurons.

The third layer (rule) has $2^4 = 16$ rules composed similarly to [6], as an example: if X1=MIN1 I X2=MIN2 I X3=MIN3 I X4=MIN4, then there is a low degree of confidence in the implementation of a DoS attack;

if X1=MAX1 I X2=MIN2 I X3=MIN3 I X4=MIN4, then there is a low degree of confidence in the implementation of a DoS attack;
if X1=MIN1 I X2=MAX2 I X3=MIN3 I X4=MIN4, then there is a low degree of confidence in the implementation of a DoS attack;
if X1=MIN1 I X2=MIN2 I X3=MAX3 I X4=MIN4, then there is a low degree of confidence in the implementation of a DoS attack;
if X1=MIN1 I X2=MIN2 I X3=MIN3 I X4=MAX4, then there is a low degree of confidence in the implementation of a DoS attack;
if X1=MIN1 I X2=MIN2 I X3=MAX3 I X4=MAX4, then the medium degree of confidence of carrying out a DoS attack;
if X1=MAX1 I X2=MIN2 I X3=MIN3 I X4=MAX4, then the medium degree of confidence of carrying out a DoS attack;
if X1=MAX1 I X2=MAX2 I X3=MIN3 I X4=MIN4, then the medium degree of confidence of carrying out a DoS attack;
if X1=MIN1 I X2=MAX2 I X3=MIN3 I X4=MAX4, then the medium degree of confidence of carrying out a DoS attack;

…

if X1=MIN1 I X2=MAX2 I X3=MAX3 I X4=MAX4, then a high degree of confidence in carrying out a DoS attack.

The fourth layer (outputmf) is function of belonging to each rule, i.e. their 2^4.

The fifth layer (output) is represented by the resulting neuron Y – the degree of confidence that the attack has taken place. This neuron has three terms: low; medium; high, used to express a level of confidence.

2. Sample preparation

To create samples, an open database NSL-KDD (Network Security Lab-KDD Cup) [5] was used, which contains data on network traffic during its normal activities, as well as during an attack. The training sample consisted of 105 examples: fifteen examples for each network attack class, as well as for the normal state (no network attack, Normal). A fragment of the training sample is shown in Table 1.

Table 1 – Fragment of the training sample (21 of 105 examples)

X1	X2	X3	X4	Y	CLASS
1	1	251	251	1	Neptune
2	2	79	79	1	Neptune
4	4	68	68	1	Neptune
1	2	15	1	1	Land
3	2	16	2	1	Land
1	2	1	6	1	Land
123	6	255	26	1	Back
121	19	255	19	1	Back
166	9	255	9	1	Back
57	16	255	59	1	Pod
2	4	2	38	1	Pod
1	1	1	1	1	Pod

continuation table 1

263	263	255	255	1	Smurf
511	511	255	212	1	Smurf
46	46	255	46	1	Smurf
4	4	255	71	1	Teardrop
21	21	255	21	1	Teardrop
68	68	255	68	1	Teardrop
2	2	150	25	0	Normal
13	1	255	1	0	Normal
5	5	30	255	0	Normal

A source: [5]

3. Creation, training and testing the ANFIS

With the help of the Fuzzy Logic Toolbox package, MatLAB created ANFIS configuration 4-5-8-16-16-1 [4], which is shown in Figure 2.

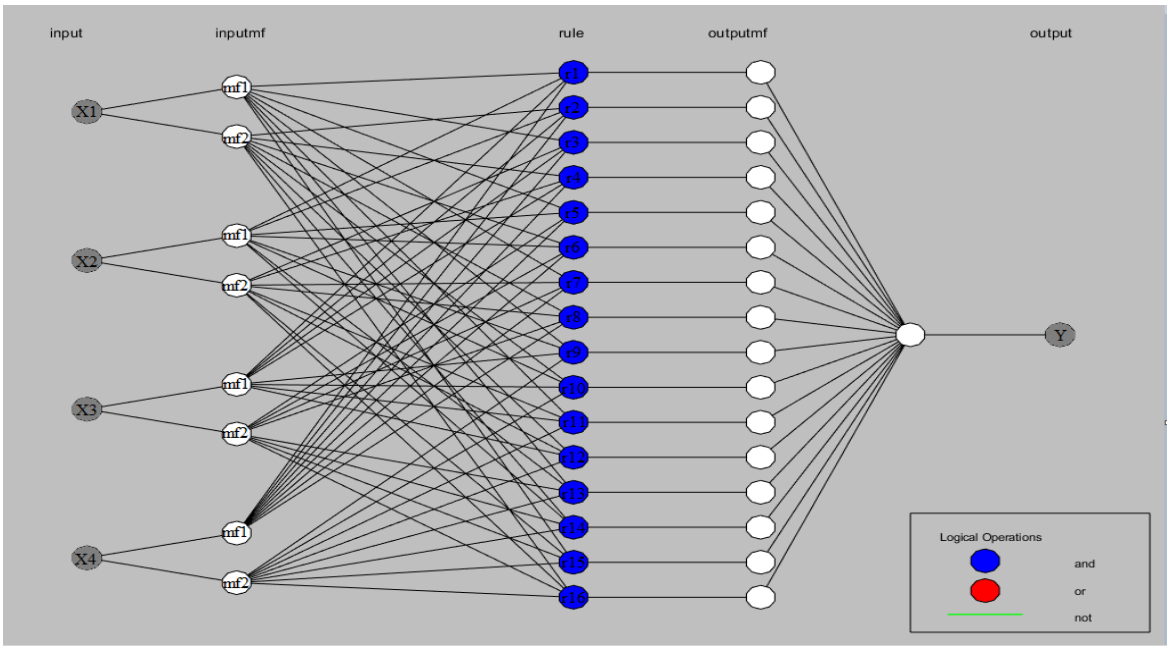

Figure 2 – Created by ANFIS in the MatLAB system

Authoring

The results of ANFIS training and testing are presented in Figure 3. As can be seen from the figure, the error of ANFIS was 0.51 during training and 0.56 during testing.

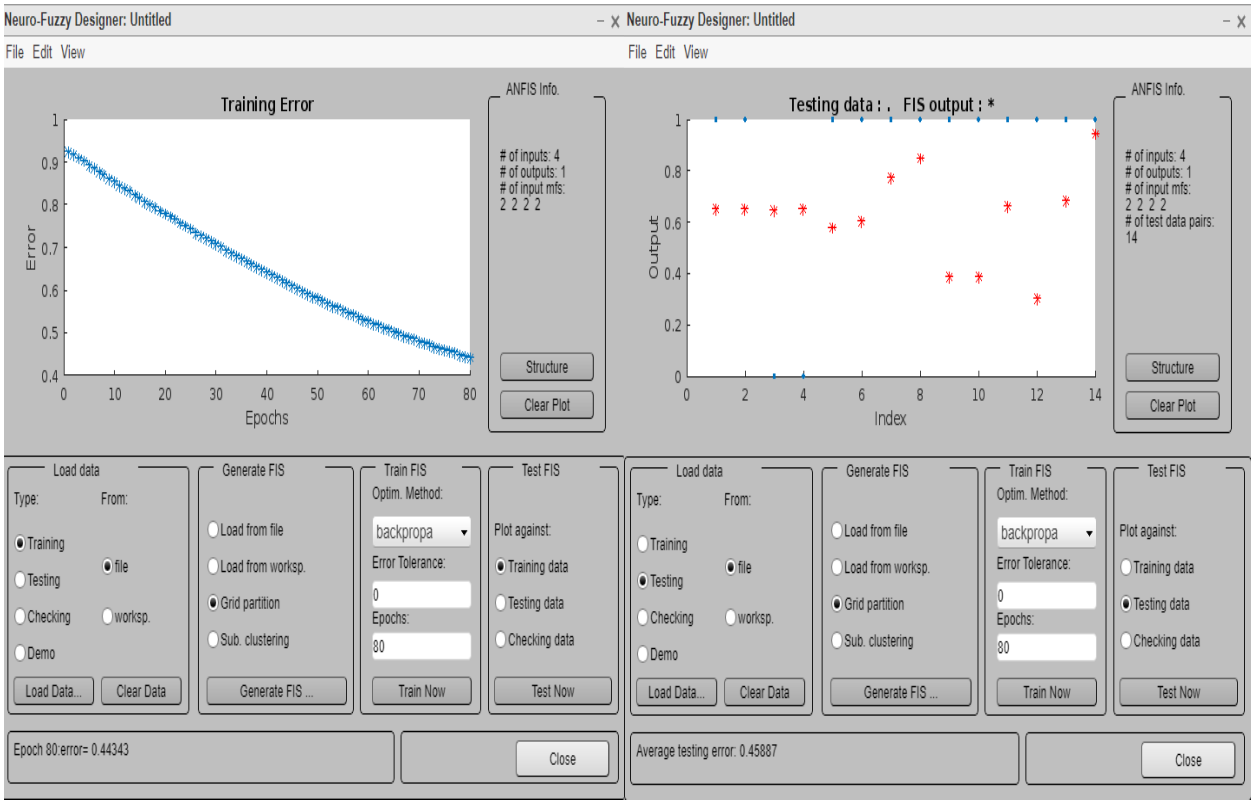

Figure 3 – ANFIS Training & Testing

Authoring

4. Exploration of ANFIS parameters

On the basis of the created ANFIS, an error study was carried out in various functions of neuronal affiliation; as an example in Figure 4 the learning outcomes of ANFIS using various neuronal affiliation functions are summarized in Table 2.

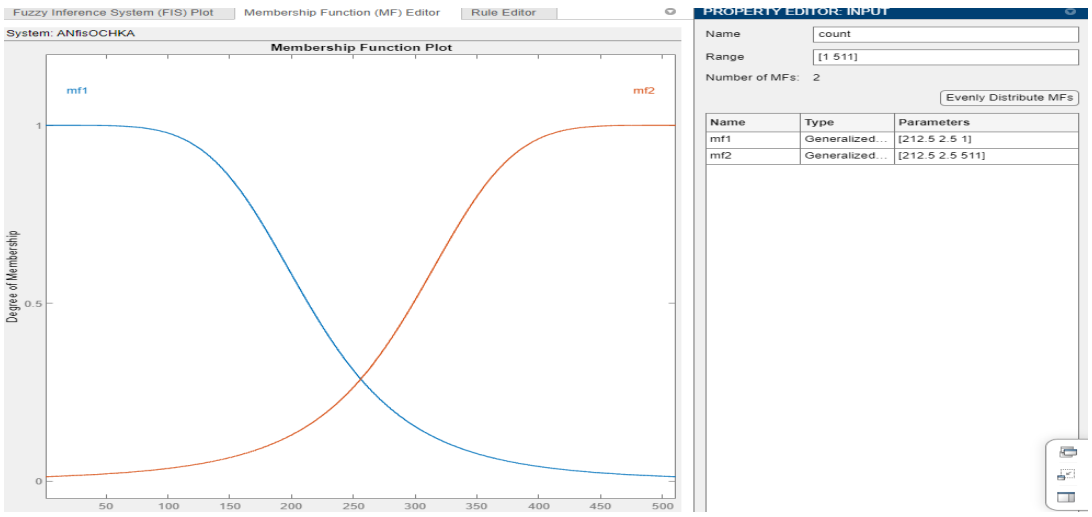

Figure 4 – Affiliation Function Editor

In addition, during the training and testing of ANFIS, studies of its error were carried out on samples of different lengths (28, 70 and 105 examples) using various methods of learning optimization: Backpropa (a method of backpropagation of an error

based on the ideas of the fastest descent method); Hybrid (a hybrid method that combines the backpropagation method of error with the method of least squares), Figure 5. As you can see from the figure, the smallest ANFIS error values are achieved when using Hybrid, and the sample length should be at least 70 examples.

Table 2 – ANFIS Training on Different Affiliation Functions

Affiliation Function	Function designation	Value of error
Triangular function	trimf	0.59608
Trapezoidal function	trapmf	0.59608
Bell function	gbellmf	0.51037
Gaussian function	gaussmf	0.52625
Gaussian function with two peaks	gauss2mf	0.60273
Parallel Boundary-Interval function	pimf	0.63198
Bilateral Sigmoidal function	dsigmf	0.62694

Authoring

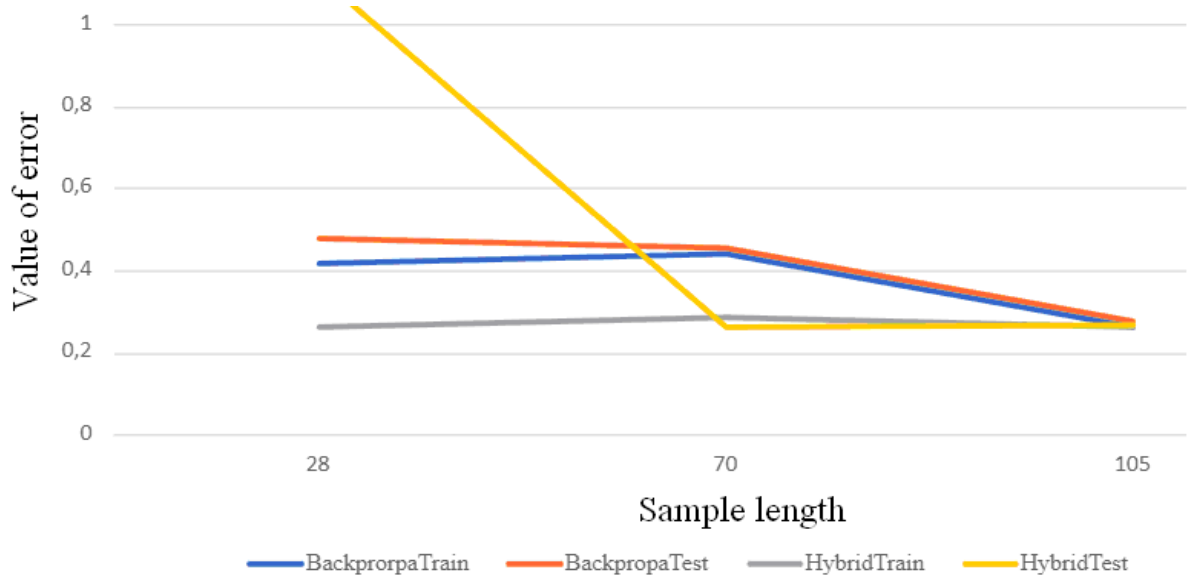

Figure 5 – ANFIS Error on Samples of Different Lengths

Authoring

Conclusions

To determine the degree of confidence of a DoS attack using the NSL-KDD database, it was created using the Fuzzy Logic Toolbox of the MatLAB ANFIS system, the Bell function was taken as a function of neuronal affiliation. On the basis of the created ANFIS, an error study was carried out on samples of different lengths (28; 70; 105 examples) using different optimization methods: Backpropa and Hybrid. The smallest error value of ANFIS was according to the Hybrid method, and it is enough to have a sample of 70 examples.

References

1. Alguliyev R. M., Imamverdiyev Y. N. & Sukhostat L. V. (2018). An improved ensemble approach for DoS attacks detection. *Radioelectronics, informatics, upravlins* [Radio Electronics, Informatics, Control]. No. 2. pp. 73-82. DOI: 10.15588/1607-3274-2018-2-8

2. Amini M., Rezaeenour J. & Hadavandi E. (2016). A Neural Network Ensemble Classifier for Effective Intrusion Detection using Fuzzy Clustering and Radial Basis Function Networks. International Journal on Artificial Intelligence Tools. Vol. 25. Iss. 02. pp. 1–32. DOI: https://doi.org/10.1142/s0218213015500335

3. Karpinski M., Shmatko A., Yevseiev S., Jancarczyk D. & Milevskyi S. (2021). Detection of Intrusion Attacks Using Neural Networks. *Miznar. nauk.-pract. conf. «Information bezpeka tha information technology», Kharkiv-Odessa* [Intl. Scin.-Pract. Conf. Information Security and Information Technologies, Kharkiv-Odesa]. pp. 117-124.

4. Kovalov R. (2023). Detecting DoS network attacks using neural network technology. *Thesis for obtaining a bachelor's degree*: speciality 125 – cybersecurity / supervisor doc. Victoria Pakhomova; Ukrainian State University of Science and Technology. Dnipro. 38 p.

5. NSL-KDD dataset. URL: https://www.unb.ca/cic/datasets/nsl.html

6. Pakhomova V. M. & Maslak A. V. (2022). Network attack detection using KDDCup99 database and neuron fuzzy technology. *Vceni zapiski tavriysky natsionalnogo university imeni V.I. Vernadskogo*. Seria: technical nauki [Scientific Notes of V. I. Vernadsky Taurida National University. Series: Technical sciences]. Vol. 33(72). No. 5. pp. 135-140. DOI: https://doi.org/10.32872/2663-5941/2022.5/19

7. Pakhomova V. M. & Motylenko V. A. (2022). Studying the possibility of using RBF for determining Smurf attacks based on the KDDCup database. *Vceni zapiski tavriysky natsionalnogo university imeni V.I. Vernadskogo*. Seria: technical nauki [Scientific Notes of V. I. Vernadsky Taurida National University. Series: Technical sciences]. Vol. 33(72). No. 6. pp. 115-121. DOI: https://doi.org/10.32872/2663-5941/2022.6/20

8. Saied A., Overill R. E. & Radzik T. (2016). Detection of known and unknown DDoS attacks using Artificial Neural Networks. Neurocomputing. Vol. 172. pp. 385-393. URL: https://doi.org/10.1016/j.neucom.2015.04.101

Анотація. *У якості методу дослідження використана ANFIS конфігурації 4-5-8-16-16-1, де 4 – кількість вхідних нейронів; 5 – загальна кількість шарів; 8 – кількість нейронів першого прихованого шару; 16 – кількість нейронів другого прихованого шару; 16 – кількість нейронів третього прихованого шару; 1 – кількість результуючих нейронів, що створена за допомогою пакета Fuzzy Logic Toolbox системи MatLAB, за результуючу характеристику взято ступень впевненості, що DoS атака відбулася за наступними термами: низький; середній; високий. З використанням відкритої бази даних параметрів мережевого трафіку NSL-KDD на створеній ANFIS проведено дослідження її похибки при різних функціях приналежності на вибірках різної довжини за різними методами оптимізації навчання. Визначено, що найменше значення похибки ANFIS склало на основі використання багатопараметричної функції Белла за методом оптимізації навчання Hybrid, при цьому достатньо мати навчальну вибірку із 70 прикладів.*

Ключові слова: *DoS атака, трафік, NSL-KDD, функція Белла, похибка.*