Influence Of Fractal Dimension Statistical Charachteristics On Quality Of Network Attacks Binary Classification

Oleg Sheluhin
Moscow Technical University of Communication and Informatics
Moscow, Russia
sheluhin@mail.ru

Mikhail Kazhemskiy
Moscow Technical University of Communication and Informatics
Moscow, Russia
m.kazhemskiy@mail.ru

***Abstract*—It is proposed to improve network attacks binary classification efficiency by introducing additional statistics of attacks fractal dimension (FD) among other descriptors and compare the performance of several classifiers. The idea of taking into account additional statistical characteristics of network attacks FD is new (in the past, only the average value of the Hurst parameter was considered) and is the main contribution of the article. The effectiveness of the proposed method is shown by evaluating network attacks and normal traffic binary classification quality with machine learning algorithms in case of using the UNSW-NB15 database. Usage of FD distribution average value, variance, skewness and kurtosis coefficients as additional information features that characterize its form and parameters can increase the efficiency of binary classification by an average of 10%.**

I. Introduction

Taking into account the self-similarity property is observed on a wide time scales, the presence of continuous attacks and abnormal activity in signal changes the self-similar nature of traffic [1]. Statistical analysis of network traffic measurements in computer network shows a clear presence of fractal or self-similar properties as in [2], [3] and [4]. In [5], [6], [7] to solve information security problems fractal analysis is used. In [8] self-similarity of network traffic data in DARPA99 is analyzed and discovered anomalous packet traffic over the certain time period can significantly change the function of self-similarity in traffic process. In [9], the results of research on detecting multifractal dimensions jumps caused by anomalous changes in the properties of telecommunication traffic on the real (current) time scale are presented.

To estimate degree of self-similarity fractal dimension of set (according to Hausdorff) D concepts and the Hurst exponent H characterising the degree of process self-similarity are used, related to each other by ratio: $D = 2\text{-}H$. In the vast majority of works in telecommunications area, the Hurst exponent is used to detect network traffic anomalies, for example in [3], [4], [5].

To build an effective network protection system, a promising direction is the joint use of fractal and intelligent data analysis. As a result, the problem of detecting network anomalies is reduced to binary classification in order to classify the considered time series to one of the two classes {"no attack", "with attack"}.

In [10], on example of the KDD Cup1999 database - [11], [12], evaluating network traffic self-similar properties characterized by Hurst exponent average value used as an additional feature shows positive effect on the quality of binary classification.

In contrast to the well-known works, it is proposed to increase the efficiency of network attacks binary classification by using not only the average value, but also the variance, skewness and kurtosis coefficients that characterize the form and parameters of the FD distribution as additional information features.

The effectiveness of the proposed approach is shown by evaluating network attacks and normal traffic binary classification quality with a wide class of machine learning (ML) algorithms. All evaluations are done on the UNSW-NB15 database [13], [14].

II. The Dataset

In Table I, the dataset statistics are provided which represents the simulation period, the flows numbers, the total of source bytes, the destination bytes, the number of source packets, the number of destination packets, protocol types, the number of normal and abnormal records and the number of unique source/destination IP addresses [13], [14].

TABLE I. Dataset statistics

Stat name		1st day(16 hours)	2nd day (15 hours)
No_of_flows		987627	976882
Src_bytes		4860168866	5940523728
Des_bytes		44743560943	44303195509
Src_Pkts		41168425	41129810
Des_Pkts		53402915	52585462
Protocol types	TCP	771488	720665
	UDP	301528	688616
	ICMP	150	374
	Others	150	374
Normal label		1064987	1153774
Attack label		22215	299068
Unique src_ip		40	41
Unique dst_ip		44	45

The compiled features from the raw data set are shown in Table II. Features 1 to 35 represent the integrated collected information from these packages. Most attributes are generated from packet headers, and additional attributes 36-47 are created based on the flow.

TABLE II. FEATURES OF DATASET UNSW-NB15

#	Feautre name	Description
Flow features		
1	Scrip	Source IP address
2	Sport	Source port number
3	Dstip	Destination IP address
4	Dsport	Destination port number
5	Proto	Transaction protocol
Basic Features		
6	State	The state and its dependent proto, e.g. ACC, CLO, else (-)
7	Dur	Record total duration
8	Sbyte	Source to destination bytes
9	Dbyte	Destination to source bytes
10	Sttl	Source to destination time to live
11	Dttl	Destination to source time to live
12	Sloss	Source packets retransmitted or dropped
13	Dloss	Destination packets retransmitted or dropped
14	Service	http, ftp, ssh, dns...,else
15	Sload	Source bits per second
16	Dload	Destination bits per second
17	Spkts	Source to destination packet count
18	dpkts	Destination to source packet count
Content features		
19	Swin	Source TCP window advertisement
20	Dwin	Destination TCP window advertisement
21	Stcpb	Source TCP sequence number
22	Dtcpb	Destination TCP sequence number
23	Smeansz	Mean of the flow packet size transmitted by the src
24	Dmeansz	Mean of the flow packet size transmitted by the dst
25	Trans_ depth	The depth into the connection of http request/response transaction
26	Res_bdy_ len	The content size of the data transferred from the server's http service.
Time features		
27	Sjit	Source jitter (mSec)
28	Djit	Destination jitter (mSec)
29	Stime	Record start time
30	Ltime	Record last time
31	Sintpkt	Source inter-packet arrival time (mSec)
32	Dinpkt	Destination inter-packet arrival time (mSec)
33	Tcprtt	The sum of 'synack' and 'ackdat' of the TCP.
34	Synack	The time between the SYN and the SYN_ACK packets of the TCP
35	Ackdat	The time between the SYN_ACK and the ACK packets of the TCP.
Additional generated features		
General purpose features		
36	Is_sm_ ips_port	If source (1) equals to destination (3)IP addresses and port numbers (2)(4) are equal, this variable takes value 1 else 0
37	Ct_state_ ttl	No. for each state (6) according to specific range of values for source/destination time to live (10) (11)
38	Ct_flw_ http_mthd	No. of flows that has methods such as Get and Post in http service
39	Is_ftp_l ogin	If the ftp session is accessed by user and password then 1 else 0.
40	Ct_ftp_ cmd	No of flows that has a command in ftp session.
Connection features		
41	Ct_srv_ src	No. of connections that contain the same service (14) and source address (1) in 100 connections according to the last time (26).
42	Ct_srv_ dst	No. of connections that contain the same service (14) and destination address (3) in 100 connections according to the last time (26).
43	Ct_dst_ ltm	No. of connections of the same destination address (3) in 100 connections according to the last time (26).
44	Ct_src_ ltm	No. of connections of the same source address (1) in 100 connections according to the last time (26).
45	Ct_src_ dport_ltm	No of connections of the same source address (1) and the destination port (4) in 100 connections according to the last time (26).
46	Ct_dst_ sport_ltm	No of connections of the same destination address (3) and the source port (2) in 100 connections according to the last time (26).
47	Ct_dst_ src_ltm	No of connections of the same source (1) and the destination (3) address in in 100 connections according to the last time (26).
Labelled features		
48	Attack_ cat	The name of each attack category. In this data set, nine categories (e.g., Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode and Worms)
49	Label	0 for normal and 1 for attack records

Table III represents the distribution of all records of the UNSW-NB15 dataset. The major categories of the records are normal and attack. The attack records are further classified into nine families according to the nature of the attacks.

TABLE III. DATASET RECORDS DISTRIBUTION

Type	Samples count	Description
Normal	2,218,761	Natural transaction data.
Fuzzers	24,246	Attempting to cause a program or network suspended by feeding it the randomly generated data.
Analysis	2,677	It contains different attacks of port scan, spam and html files penetrations.
Backdoors	2,329	A technique in which a system security mechanism is bypassed stealthily to access a computer or its data.
DoS	16,353	A malicious attempt to make a server or a network resource unavailable to users, usually by temporarily interrupting or suspending the services of a host connected to the Internet.
Exploits	44,525	The attacker knows of a security problem within an operating system or a piece of software and leverages that knowledge by exploiting the vulnerability.
Generic	215,481	A technique works against all blockciphers (with a given block and key size), without consideration about the structure of the block-cipher.
Reconnais.	3,987	Contains all Strikes that can simulate attacks that gather information.
Shellcode	1,511	A small piece of code used as the payload in the exploitation of software vulnerability
Worms	174	Attacker replicates itself in order to spread to other computers. Often, it uses a computer network to spread itself, relying on security failures on the target computer to access it.

Using features 1-5 and 29, 30, flow data was selected from source data, separated by each category. Fig. 1-9 show graphs

of incoming / outgoing packets per second dependencies for each attack category over two days.

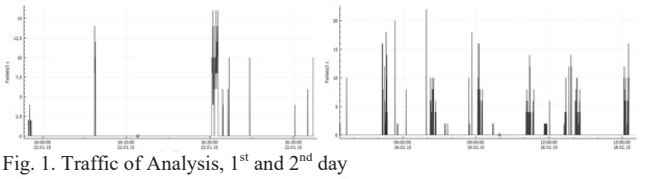

Fig. 1. Traffic of Analysis, 1st and 2nd day

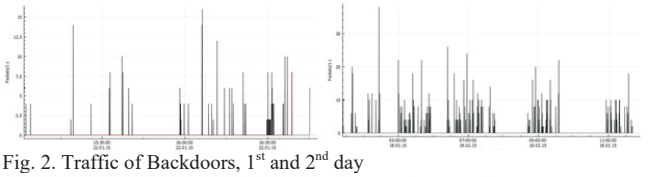

Fig. 2. Traffic of Backdoors, 1st and 2nd day

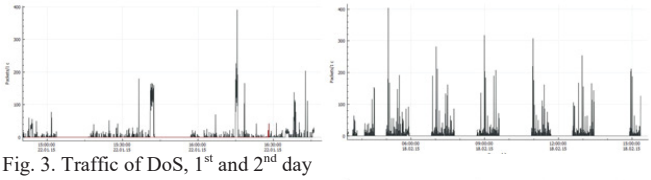

Fig. 3. Traffic of DoS, 1st and 2nd day

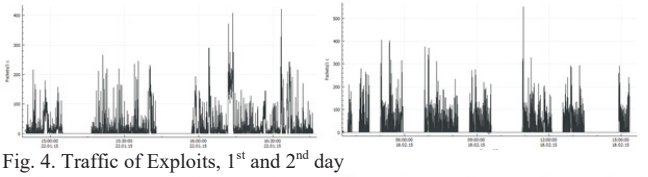

Fig. 4. Traffic of Exploits, 1st and 2nd day

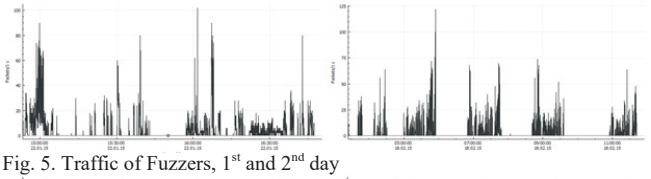

Fig. 5. Traffic of Fuzzers, 1st and 2nd day

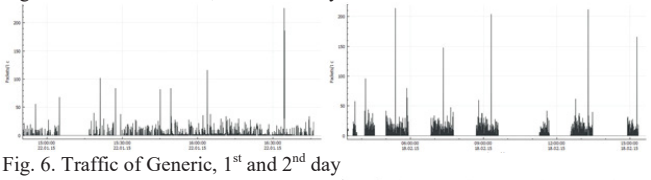

Fig. 6. Traffic of Generic, 1st and 2nd day

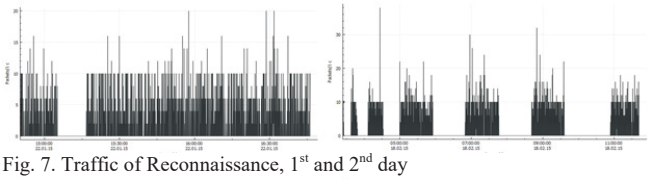

Fig. 7. Traffic of Reconnaissance, 1st and 2nd day

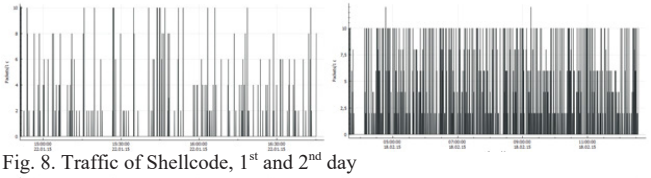

Fig. 8. Traffic of Shellcode, 1st and 2nd day

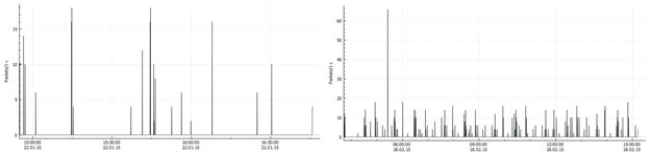

Fig. 9. Traffic of Worms, 1st and 2nd day

These dependencies allow us to evaluate the FD of attacks and normal traffic at the training stage and then use them at the classification stage.

III. Structure of Training and Testing Samples in UNSW-NB15 Dataset

In the analyzed dataset, testing and training samples do not contain features 29-30, as well as features 1-4. In total, the samples contain ***175341*** and ***82332*** training and test records, respectively. Entries distribution by category is shown on Fig. 10 and Fig. 11.

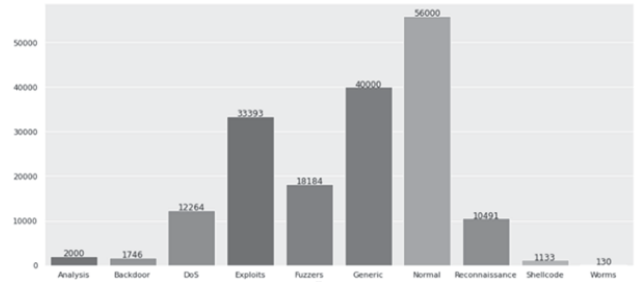

Fig. 10. Training set records distribution for all categories

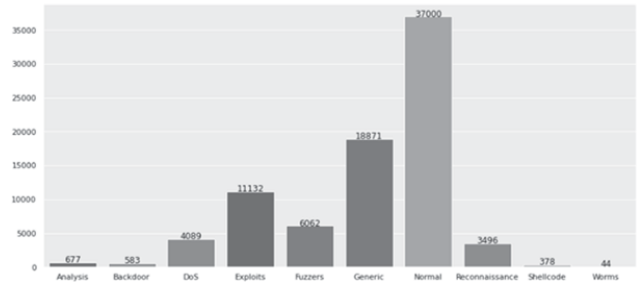

Fig. 11. Testing set records distribution for all categories

As can be seen from Fig. 10 and Fig. 11, classes in UNSW-NB15 dataset are not balanced, i.e. the number of records with one classes (e.g. *Normal*, *Generic*) is several times higher than the number of records for other classes (e.g. *Worms*, *Shellcode*). With high probability it could lead to "*under-fitting*" of machine learning algorithms and, consequently, to errors in of rare categories classification.

Some ML algorithms work with all features as with a single vector, and since the values of the features are diverse, this can lead to incorrect calculations [10]. To eliminate this drawback, a *normalized* data set was used where features values could be from 0 to 1. Normalization was performed using the *Min-Max* principle in accordance with the formula (1).

$$x' = \frac{x - min(X)}{max(X) - min(X)} \quad (1)$$

where *min(X)* and *max(X)* are the minimum and maximum values of the feature in the entire data set.

IV. Classification Algorithms and Metrics

The following classification algorithms were used to classify the dataset [15]:

- ***K-Nearest Neighbors method*** (neighbors, ***k-NN***). A *normalized* dataset was used.
- ***Logistic regression (LR)***. SAGA algorithm was used to solve the logistic regression equation. A *normalized* dataset was used.

- ***Naive Bayes*** (***NB***). A *normalized* dataset was used.
- ***Decision Tree Classifier*** (***DTC***). Gini coefficient was used as an evaluation function. Data *normalization is not required*. During empirical analysis, it was found that the best result of the algorithm is achieved with the *number of features 31* and the *depth of the tree 29*.
- ***Random Forest*** (***RF***). Due the fact that the basis of the algorithm is a decision tree, *normalization is not required*. The best result for the analized dataset was obtained by dividing the data into *100 subsets*.
- ***Ada Boost*** (***AB***). Due the fact that the basis of the algorithm is a decision tree, *normalization is not required*. The best result for the dataset was obtained by dividing the data into *1000 subsamples*.

The following metrics are most often used in machine learning tasks to evaluate the efficiency of built models: *precision*, *recall*, *F-score*, *ROC-curves* (*Receiver Operating Characteristic curve – error curve*), *AUC-ROC* and *AUC-PR* (*Area Under Curve* - area under *error curve* and area under the *precision-recall curve*)

After the classification it is possible to obtain four types of results: *TP* (True Positive), *TN* (True Negative), *FP* (False Positive), *FN* (False Negative). These results can be represented as the *confusion matrix*.

V. Additional fractal feautures of attacks

To improve the efficiency of binary classification of the analyzed data set it is proposed to introduce additional features (attributes) for each of detected attacks, in contrast to [10]. As such attributes it is proposed to use experimentally obtained FD (Hurst index) statistical characteristics of the average value of the FD $\boldsymbol{M_H}$, the variance of the Hurst index $\boldsymbol{D_H}$, skewness $\boldsymbol{K_{sk}}$ and the $\boldsymbol{K_k}$ kurtosis characterizing FD values probability density function *w(H).*

Fractal dimension estimation ***D*** or *Hurst exponent* ***H*** depends on many factors, and is a complex task in itself, since there are always constraints associated with a finite set of data when working in real conditions. Most often for the Hurst exponent estimation the analysis of the rescaled range series (R/S-method), the analysis variance changes and wavelet analysis [1], [2] are used.

A normalized dimensionless measure capable to describe time series variability is called *rescaled range series* (R/S). For a given set of observations X with mean $\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$, where n is the number of observations, introduced the concept of range (the difference between the maximum and minimum deviation $R(n) = max\Delta_j - min\Delta_j$, where $1 \le j \le n$,

$$\Delta_k = \sum_{i=1}^{k}(X_i - k\bar{X}) \,, \forall k = \overline{1,n}\; S(n) = \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X})^2.$$

It is known for many natural phenomena the mathematical expectation of the normalized range is approximately equal to cn^H at $n \to \infty$,where $\boldsymbol{c}$ is a positive constant independent of n. Then the exponent H can be obtained by drawing a graph of the dependence $\log(M\left(\frac{R(n)}{S(n)}\right)$ of $\log(n)$, and using the obtained points, to choose by the least squares method a straight line with a slope H [7]. The empirical law $H = \ln\left(\frac{R}{S}\right) / \ln\left(\frac{n}{2}\right)$ is used to determine the quantitative value of H. When performing numerical calculations, only implementations of attacks for which the number of observations is n >100 were taken into account. In this case, the error in estimating the Hurst index did not exceed 5%.

The results of fractal analysis of attacks presented in Table I and normal traffic in the absence of attacks by R/S-method are given in Table IV. It shows results of evaluating Hurst exponent H statistical parameters for all attacks with the number of attack implementations (intervals) equal to N. The calculations did not take into account the attacks of Shellcode and Worms, because they did not have the long intervals necessary for evaluating the parameters of the FD.

TABLE IV. Statistical characteristics of *w*(*H*) for attacks

Type	(N)	M_H	D_H	K_{sk}	K_k
Normal	20	0.6949	0.0009	0.3137	0.4431
Analysis	9	0.6685	0.0084	0.2493	1.1772
Backdoors	8	0.6121	0.0030	0.8678	0.4829
DoS	18	0.5900	0.0051	1.0607	2.5674
Exploit	21	0.7251	0.0060	0.4528	0.3805
Fuzzers	23	0.6891	0.0045	0.1573	1.2453
Generic	15	0.6726	0.0083	0.3544	1.3438
Reconnaissance	9	0.6026	0.0013	0.1751	1.0603

According to obtained results, four new features were added from for training and testing samples - Table V.

TABLE V. Additional features from Table IV

№	Feature	Description
Additional features of FD		
50	hurst_avg	Expectation M_H of FD for distribution *w*(*H*)
51	hurst_desp	Variance D_H for distribution *w*(*H*)
52	hurst_skew	Skew K_{sk} for distribution *w*(*H*)
53	hurst_kurtosis	Kurtosis K_k for distribution *w*(*H*)

If there is no entry in Table IV, the additional feature was assumed to be **zero**. This means that there is no additional attribute for this attack category, and it is not taken into account in the classification process.

Table VI shows how many records have the values *hurst_** ≠ 0. Parameter *hurst_** (50-53) is set for all traffic categories defined in Tavle IV. So values hurst_* = 0 will only apply to *Shellcode* and *Worms* categories.

TABLE VI. Ratio of entries for parameter *hurst_**

Feature	Sample	hurst ≠ 0 records count	% of all records
hurst_*	Training	172332	98%
	Testing	81327	99%

Fig. 12 and Fig. 13 show histograms that allow us to evaluate importance of features with entered additional parameters of the FD. The importance of the introduced features was calculated using the Gini coefficient [15], which is the basis for decision-making algorithms "*Decision Tree*" (Fig. 12) and "*Random forest*" (Fig. 13).

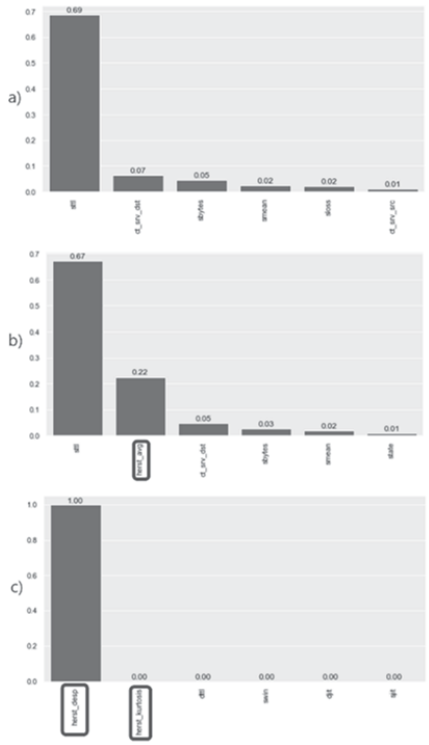

Fig. 12. Importance of the first 6 features for the "***Decision Tree***" algorithm in the classification issue ***a)*** excluding FD; ***b)*** with parameter hurst_avg(50) only; ***c)*** with all statistical parameters of FD from Table V

Comparison of histograms at Fig. 12a and Fig. 12b shows the "*Decision Tree"* algorithm takes into account only one additional attribute – the average value of Hurst index *hurst_avg,* puts it in second place in classification. However, if it is possible to evaluate additional parameters of the FD, the *hurst_desp* and *hurst_kurtosis* attributes will have the greatest impact on classification. According to Fig. 12c, the *hurst_desp* attribute characterizes the spread of the Hurst index relative to the average value. The *hurst_kurtosis* parameter, which characterizes the distribution form of the Hurst index *w*(*H*), has an important but significantly lower value.

As can be seen from Fig. 13a for the "*Random Forest*" algorithm, in the absence of a FD feature, the quality of classification is affected by a larger number of features, compared to the "*Decision Tree*" (Fig. 12a) algorithm. However, even in this case, taking into account only one additional attribute *hurst_avg* puts it in the first place in importance.

If it is possible to evaluate additional parameters of FD that characterize the shape and parameters of the Hurst index distribution *w*(*H*), the *hurst_desp* and *hurst_skew* attributes will have the greatest impact on classification. The parameters *hurst_avg* and *hurst_kurtosis*, which characterize the shape of the distribution *w*(*H*), have a slightly lower value and have the 5th and 6th places in terms of importance.

As can be seen from the presented histograms, additional statistical attributes from Table V have a significant impact on the decision-making algorithm.

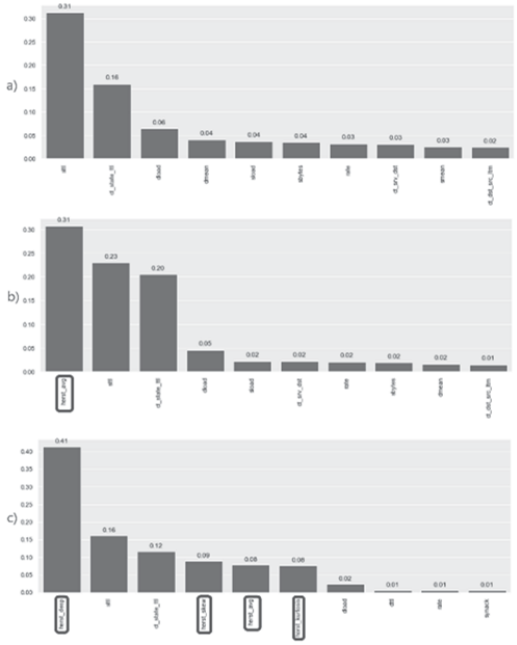

Fig. 13. Importance of the first 6 features for the "***Random Forest***" algorithm in the classification issue ***a)*** excluding FD; ***b)*** with parameter hurst_avg(50) only; ***c)*** with all statistical parameters of FD from Table V

VI. RESULTS OF BINARY CLASSIFICATION

For binary classification, all attack categories were labeled as "Attack" category. As a result, the classification is reduced to the task of identifying two classes: **Attack** and **Normal**. Let's consider the results of a comparative analysis of influence of statistical characteristics of *w(H)* on quality of binary classification. Three operating cases were analyzed.

1) Classification only when using the initial features 1...49 from Table I. Classification results are shown on Fig. 14-15 (a).

2) Classification when adding one additional 50th feature – ***hurst_avg*** (Table V - 50) - the average value of the Hurst indicator to the set of features 1...49. Classification results corresponding to this case are shown in Fig. 14-15 (b).

3) Classification when adding all four statistical features 50...53 from Table V – ***hurst_stat*** (*hurst_avg*, *hurst_desp*, *hurst_skew*, *hurst_kurtosis*) to the set of features 1 ... 49. The classification results corresponding to this case are shown in figures Fig. 14-15 (c).

Features 1-4, 29, 30 from Table II were not used in classification because they were not presented in initial training and test samples.

As can be seen from the presented results, the effectiveness of using additional attributes ***hurst_stat*** (FD statistical parameters of traffic) is most noticeable for "*K-nearest neighbors"* classification and "*Logistic regression"* algorithms. For these algorithms, the gain from using additional attributes ***reaches 21%*** for the *precision* metric in the presence of attacks and ***41%*** in the absence of them.

The gain in the *f1-score* metric is more modest and is about ***7%***. For the *AUC-P*R metric, the gain is ***7-8%***. The greatest effect is achieved by using the average value of the fractal dimension – M_H as an additional attribute. When using classification algorithms "*Decision Tree*" and "*Random Forest*", the gain in classification from using additional attributes is about 15-20% with additional M_H attribute for almost all considered metrics. More significant benefit from using additional attributes is decreasing training and testing time. These results are shown in Table VIII.

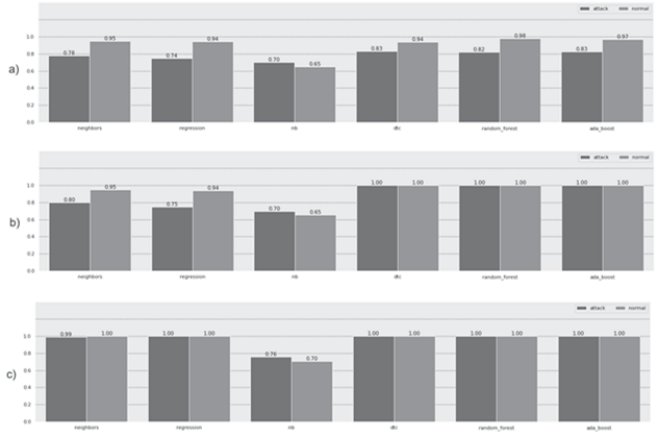

Fig. 14. Precision metric values of classification ***a)*** excluding FD; ***b)*** with parameter hurst_avg(50) only; ***c)*** with all statistical parameters of FD hurst_stat from Table V.

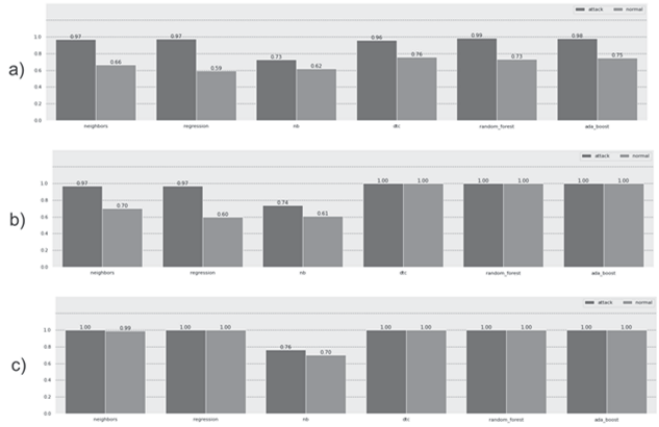

Fig. 15. Recall metric values of classification ***a)*** excluding FD; ***b)*** with parameter hurst_avg(50) only; ***c)*** with all statistical parameters of FD hurst_stat from Table V.

TABLE VII. PERFORMANCE OF CLASSIFICATION ALGORITHMS, SECONDS

Hurst exp.	**no**			**hurst_avg**			**hurst_stat**		
	train	**test**	**total**	**train**	**test**	**total**	**train**	**test**	**total**
K-nn	76	34.8	110.9	94.0	45.8	139.8	**67.6**	**27.4**	**95**
LR	6.8	0.0	6.9	7.2	0	7.2	**4.7**	**0**	**4.7**
NB	0.6	0.0	0.6	0.5	0	0.6	**0.5**	**0**	**0.5**
DTC	2.3	0.1	2.4	1.5	0.1	1.6	**0.6**	**0.1**	**0.7**
RF	16.5	0.4	16.8	8.2	0.3	8.5	**4.6**	**0.2**	**4.8**
AB	547	14.9	562	596.7	15.2	611.9	**469.2**	**13.5**	**482.7**

As can be seen, the "*Decision Tree*" and "*Random Forest*" algorithms are also the most effective here. In the case of the "*Decision Tree*" algorithm, usage of one additional parameter M_H value led to a reduction in training and testing time by more than 1.5 times, and for the "*Random Forest*" algorithm by 1.98 times. Using the four additional attributes **hurst_stat** (*hurst_avg*; *hurst_desp* ; *hurst_skew*; *hurst_kurtosis*) shown in Table V increased its importance so resulted in a 3.54 times reduction in training and testing time for "*Decision Tree*" algorithm and a 3.48 times reduction in the random forest algorithm. The absolute numbers were lower for the "*Decision Tree*" algorithm and amounted to 0.68 seconds, while for the "*Random Forest*" - 4.83 seconds.

VII. CONCLUSION

The introduction of additional statistical parameters of fractal dimension statistical characteristics of the average value of the Hurst index $\boldsymbol{M_H}$, the variance of the Hurst index $\boldsymbol{D_H}$, skewness $\boldsymbol{K_{sk}}$ and the $\boldsymbol{K_k}$ kurtosis characterizing FD values probability density function *w(H),* has a positive effect on the quality and speed of binary classification of attacks.

To estimate these parameters, traditional methods of estimating the Hurst index can be applied, using the normalized span analysis (R/S method), analysis of the variance change graph, and wavelet analysis. The sample size that allows evaluating the specified parameters *n* should allow evaluating the specified parameters with a given error both at the training stage and at the testing stage.

Comparative analysis of additional attributes has shown that the most significant attributes on the "*Decision Tree*" algorithm are the $\boldsymbol{M_H}$ and $\boldsymbol{D_H}$. When using the "*Random Forest*" algorithm, the $\boldsymbol{D_H}$ attribute has the highest significance. However, the value of both the $\boldsymbol{K_{sk}}$ attributes and the $\boldsymbol{K_k}$ that characterize the form of the *w(H)* distribution is high.

Thus, the introduction of additional statistical parameters of the Hurst exponent has a positive effect on the quality and performance of binary classification. The greatest effect is noticeable for the classification algorithms "*K-nearest neighbors* "and "*Logistic regression".*

The use of the average value, variance, skewness and kurtosis coefficients as additional information features that characterize the shape and parameters of the distribution of statistical characteristics of the FD distribution makes it possible to increase the binary classification efficiency by an average of 10%.

The greatest effect of additional FD statistical parameters is noticeable for the classification algorithms "*K-nearest neighbors* "and "*Logistic regression"*

For the "*Decision Tree*" and "*Random Forest*" algorithms, the greatest effect of using four statical characteristics of FD ($\boldsymbol{M_H}$, $\boldsymbol{D_H}$, $\boldsymbol{K_{sk}}$, $\boldsymbol{K_k}$) as additional attributes is to reduce the training and testing time in about 3.5 times for each algorithm.

REFERENCES

[1] O. Sheluhin, S. Smolskiy, and A. Osin, "Self-similar processes in telecommunications," John Wiley & Sons, 2007.

[2] A. Atayero and O. Sheluhin, “Integrated model for information communication systems and networks,” Desigtn and Development. IGI Global. USA, 2013. p. 462.

[3] K. Park and W. Willinger, “Self -similar network traffic and performance evaluation,” John Wiley & Sons. - 2000.

[4] H. Monowar, D. Bhattacharyya and J. Kalita, “Network anomaly detection: methods, systems and tools ,” IEEE Communications surveys & tutorials, 2013, vol. 60(1). — pp. 303–336.

[5] X. Wang and B. Fang, “An exploratory development on the Hurst parameter variety of network traffic abnormity signal,” J. Harbin Inst. Technol., 37: 1046-1049, 2005.

[6] A. Mohiuddin, M. Abdun Naser and H. Jiankun, “A survey of network anomaly detection techniques,” J. Network and Comp. App., no 60, p. 21, 2005.

[7] Z. Sheng, Z. Qifei, P. Xuezeng and Z. Xuhui, “Detection of low-rate ddos attack based on self similarity,” in 2010 Second International Workshop on Education Technology and Computer Science, vol. 1 , pp. 333-336, 2010.

[8] G. Kaur, V. Saxena and J.P. Gupta, “Study of self-similarity for detection of rate-based network anomalies,” International Journal of Security and Its Applications Vol. 11, No. 8 (2017), pp.27-44.

[9] O. Sheluhin and I. Lukin, “Network traffic anomalies detection using fixing method of jumps of multifractal dimension in the real-time mode,” Automatic Control and Computer Sciences, Sep. 2018, Volume 52, Issue 5, pp 421–430. DOI 10.3103/S0146411618050115

[10] O.I. Sheluhin. and M.A. Kazhemskiy, “Influence of fractal dimension on network anomalies binary classification quality using machine learning methods,”_Automatic Control and Computer Sciences, 2020, Vol. 54, No. 3, pp. 216–228 DOI: 10.3103/S0146411620030074

[11] KDD Cup 1999 Data https://kdd.ics.uci.edu/databases/kddcup99/kddcup99

[12] NSL-KDD Dataset https://www.unb.ca/cic/datasets/nsl.html

[13] Australian Center for Cyber Security (ACCS). (2014). Retrieved from http://www.accs.unsw.adfa.edu.au/

[14] N. Moustafa and J. Slay, “UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set),” Military Communications and Information Systems Conference (MilCIS), 2015, At Canberra, Australia, DOI: 10.1109/MilCIS.2015.7348942

[15] O. Sheluhin, S. Erokhin and A. Vaniushina, “Ip-traffic classification by machine learning methods,” Moscow: Hotline-Telekom, 2018