Date of publication DEC-31, 2022, date of current version NOV-19, 2022
www.computingonline.net/ computing@computingonline.net

Print ISSN 1727-6209
Online ISSN 2312-5381
DOI 10.47839/ijc.21.4.2779

Estimates of the Complexity of Detecting Types of DDOS Attacks

NIKOLAY A. IGNATEV, ERKIN R. NAVRUZOV
National University of Uzbekistan, Vuzgorodok 4, Tashkent, 100174, Uzbekistan
(E-mail: n_ignatev@rambler.ru, erkinbek0989@gmail.com)
Corresponding author: Nikolay A. Ignatev. (e-mail: n_ignatev@rambler.ru).

ABSTRACT The problem of substantiating decisions made in the field of information security through estimates of the complexity of detecting types of DDOS attacks is considered. Estimates are a quantitative measure of a particular type of attack relative to normal network operation traffic data in its own feature space. Own space is represented by a set of informative features. To assess the complexity of detecting types of DDOS attacks, a measure of compactness by latent features on the numerical axis was used. The values of this measure were calculated as the product of intraclass similarity and interclass difference. It is shown that compactness in terms of latent features in its own space is higher than in the entire space. The values of latent features were calculated using the method of generalized estimates. According to this method, objects of normal traffic and a specific type of attack are considered as opposition to each other. An informative feature set is the result of an algorithm that uses the rules of hierarchical agglomerative grouping. At the first step, the feature with the maximum weight value is included in the set. The grouping rules apply the feature invariance property to the scales of their measurements. An analysis of the complexity of detection for 12 types of DDOS attacks is given. The power of sets of informative features ranged from 3 to 16.

KEYWORDS DDOS attacks; structure of object relations; latent features; Big Data; hierarchical agglomerative grouping.

I. INTRODUCTION

IN the computer science theory, the concept of complexity is most often associated with the complexity of algorithms or computational complexity [1]. Evaluation of the complexity of algorithms is usually determined by the time spent or by the size of the memory used. The value of computational complexity depends on the output data [2], [3]. The specifics of the DDOS attack detection process is taking into account that the attacker masquerades as much as possible under the parameters of normal traffic [4], [5]. There is a need to justify the decisions made in the field of information security through the ordering of types of DDOS attacks according to the complexity of their detection [6].
The specificity of calculating the estimates of the complexity of detecting attacks is [7]:

– the detection process is implemented through algorithms, not through instrumentation;
– the algorithms are adapted to the use of big data;
– there is a system of numerical indicators that allows representing the complexity of detecting attacks in an ordinal measurement scale.

An important component of the technology (methods) for calculating estimates of the complexity of detecting attacks is the ability to interpret the numerical results by experts and users in the field of information security [8]. It is necessary to show that whether the realization of this goal requires (not required) the development of special methods and computational algorithms based on them or not [1]. Associated with the computation of an estimate of the complexity of detecting attacks is the verification of the truth of such statements [9], [10].

- To calculate the complexity score, there is dedicated equation with a fixed number of arguments.
- There is no equation for calculating, for this reason, the value of the complexity estimate is determined algorithmically, according to the original set of raw

features for each type of attack.

The derivation of the equation for the first statement assumes the presence of a regression dependence of informative traffic indicators on the complexity of detection in the form of target attribute values. A practically unrealizable way to obtain these values is to collect data based on expert opinions [7]. When proving the truth of the second statement, it is proposed to use several criteria to calculate the integral indicator, which is considered as an estimate of the complexity of detecting attacks. An example of the development of a multimethod technology for the multicriteria choice of solutions based on integral indicators is described in several lireratures [11], [12].

When calculating integral indicators, it is necessary to take into account the diversity of features, the scale of their measurements, the absence (presence) of gaps in the data, the availability of numerical methods designed to work with Big Data [13], [14].

II. RELEATED WORKS

One of the real ways to reduce the cost of resources for detecting attacks is the selection of informative feature sets [10]. The main goals pursued in feature selection are to achieve high accuracy and prevent retraining of algorithms [15]. When identifying selection methods, they are divided into 3 main types [16]:

1) Filtration methods;
2) Wrappers methods;
3) Embedded methods.

A detailed review of the methods for selecting informative features and choosing recognition algorithms for the problems of detecting DDOS attacks is given by (Author) [16], [17]. For comparative analysis, four selection methods were used [18], [19], [20], the results of which are given in Table 1.

Table 1. The results of the selection of informative features

#	Feature selection methods	Feature selected
1	Features selected using RFE witht linearsvc and n features to select =15 and step=5 (pp StandardScaler)	11, 12, 13,15, 16, 33, 41, 47, 59, 60, 70, 72, 74, 75
2	Feature selected after using SelectPercentile(12 features percentile=15) and using standardscaler	3, 17, 19, 20, 23, 24, 25, 29 ,30, 61, 81, 83
3	Lasso cv(toi=.0001) with SelectFromModel(model,threshold=-05) and StandardScaler	3, 8, 9, 23, 24, 32, 34, 47, 54, 59, 60, 61
4	RFE (n features to select 15) with Lasso(alpha=.05) and StandardScaler(results reverified)	3, 5, 6, 9, 10, 11, 12, 13, 19, 24 , 54, 61, 76, 81

As can be seen from Table 1, the results of the selection of informative features differ significantly depending on the methods used and from the point of view of information security are poorly justified, largely due to the lack of consideration of the features of detecting types of attacks. For this reason, a qualitative analysis of the composition of informative sets of features does not make practical sense.

Random Forest and *k-nearest neighbors* (KNN) methods were named as the best machine learning models using informative features [16]. It was noted that the preliminary selection of informative features had a strong impact on the results of the Random Forest and did not significantly improve the results of the KNN method [18]. A distinctive feature of the Random Forest method is that its implementation is carried out within the framework of the metamodel of the ensemble of algorithms using boosting technology.

According to (Authors) [21], the existence of a set of elementary classifiers on a subset of features is proved to form a recognition metamodel (ensemble of algorithms) using stacking technology, the accuracy of which is not lower than when using classifiers for the entire set of features.

In the process of implementing the technology for calculating the value of the complexity of detecting attacks, the method of generalized estimates of objects is used within the framework of a two-class recognition problem. As opposition to each type of DDOS attack, data from normal network operation traffic is considered. To calculate the attack detection complexity score, the following are used [11]:

- own set of initial features for each type of DDOS attacks to form a recognition metamodel based on an ensemble of elementary algorithms using stacking technology;
- values of the latent feature of objects according to their own set of initial features;
- the measure of the compactness of sampling classes.

The calculation process is associated with data preprocessing to form a new feature space of the same dimension as the original one. The results of the implementation of the preprocessing procedure are [20], [22]:

- invariance of signs to the scales of their measurements;
- the transformation of the values of quantitative characteristics into nominal gradations;
- the nonlinear transformation of the values of nominal features in $\{1, 2\}$.

Based on the results of preprocessing, noise features are determined. Belonging to a set of noise features is the basis for excluding a feature from the process of forming an ensemble of algorithms using stacking technology.

The implementation of computational algorithms is associated with operations on large amounts of input data. The transformation of values of heterogeneous features into $\{1, 2\}$ is nonlinear, the implementation of which uses the stability indicator of the feature. The stability values are consistent estimates [23]. The consistency of the estimates is a guarantee that, in samples from the general population, the differences between the values of stability will be insignificant [21]. Methodologically, this property allows you to perform calculations on selected subsamples in Big

Data problems and adapt the results to all data [24]. There are prospects for using latent features for decision rules by pairs (normal traffic, type of DDOS attacks) [25].

When developing a model based on machine learning algorithms, it is of particular importance to decide on the composition of the feature set and objects used as input data [15], [26]. The goals pursued in the selection of informative features:

1) ordering the types of DDOS attacks according to the complexity of their detection;
2) selection and justification of rules for classifying the types of DDOS attacks.

III. PROBLEM DEFINITION

A set of objects $E_0 = \{S_1, \ldots, S_m\}$, is considered, divided into $l+1$ disjoint subsets (classes) $K_0, K_1, \ldots, K_l$, of which $K_1, \ldots K_l$ represent descriptions of l types of DDOS attacks, K_0- is normal traffic.

Sample objects are set by a set of n different types of features $X(n) = (x, \ldots, x_n)$, ξ of them are measured in interval scales, $n - \xi-$ in nominal.

Procedures are considered to be defined for:

1) partitions $X(n)$ into disjoint groups $X^j(n_1), ..., X^j(n_t)$ by a pair (K_0, K_j), $t \geq 1$, $j = 1, \ldots, l$, $n_1 + \cdots + n_t \leq n$, for each of which latent features from the set $Y^j(t) = (y_1^j, ..., y_t^j)$are calculated;
2) calculation of compactness measure $F(j, r)$ for each $y_r^j \in Y^j(t)$.

Required:

1) determine $d = \arg\max_r F(j, r)$, $Q_j = F(j, d)$ and select a set $X^j(n_d)$ for each pair of classes (K_0, K_j);
2) sort the types of DDOS attacks by Q_j, $j = 1, \ldots l$.

IV. PARTITIONING OF FEATURE VALUES INTO INTERVALS

The proposed technology for calculating estimates of the complexity of detecting types of DDOS attacks uses the results of splitting quantitative (raw and latent) features into non-overlapping intervals according to two criteria. Among the original properties of these criteria is invariance to measurement scales. This property allows us to consider the solution of the problem without any restrictions on the units used for the initial (raw) features.

Criteria are used to analyze the variety of relations between the values of quantitative features of objects on the numerical axis and data preprocessing. Features of data preprocessing are nonlinear transformations of heterogeneous (qualitative and nominal) features through the values of the membership function of objects to classes. The search for criteria extreme is preceded by the ordering of feature values in non-decreasing order.

Let for the values of the feature $x_c \in X(n)$ in the description of the objects $K_0 \cup K_j$, $j = 1, \ldots, l$, a non-decreasing ordered sequence

$$r_1, \ldots, r_i, \ldots, r_h, h = K_0 \cup K_j \tag{1}$$

is constructed. As boundaries of two disjoint intervals $[\pi_1; \pi_2]$, $(\pi_2; \pi_3]$, determined by, $\pi_1 = r_1$, $\pi_2 = r_i$, $1 < i < h$, $\pi_3 = r_h$ are used. The intervals $[\pi_1; \pi_2]$ and $(\pi_2; \pi_3]$ are identified as the first and second, respectively. The feature weight for class objects according to (1) is calculated [27] as the maximum of the product of intraclass similarity and interclass difference according to the

$$\left(\frac{\sum\limits_{d=1}^{2} \sum\limits_{i=1}^{2} \left(u_{a[i]}^d - 1 \right) u_{a[i]}^d}{\sum\limits_{i=1}^{2} \left| K_{a[i]} \right| \left(\left| K_{a[i]} \right| - 1 \right)} \right) \times \left(\frac{\sum\limits_{d=1}^{2} \sum\limits_{i=1}^{2} u_{a[i]}^d \left(\left| K_{a[i]} \right| - u_{a[3-i]}^d \right)}{2 \left| K_0 \right| \left| K_j \right|} \right) \rightarrow \max\limits_{\pi_1 < \pi_2 < \pi_3} \tag{2}$$

criterion, where $u_{a[i]}^d \left(u_{a[3-i]}^d \right)$ – is the number of x_c feature values for objects from the class $K_{a[i]}, (K_{a[3-i]})$ in the interval, $a[1] = 0$, $a[2] = j$. The set of admissible values of the criterion (2) belongs to (0;1] and is used to evaluate class objects on the numerical axis. If each interval contains all values of a feature of objects from the same class, then its weight is equal to 1.

The boundary between classes (threshold) for a quantitative trait x_a is calculated as

$$\Gamma_a = \frac{\pi_2 + \eta}{2}, \tag{3}$$

where η – is the closest value to π_2 from the interval $(\pi_2; \pi_3]$, determined by (2). It is believed that the nature of the data environment when calculating the threshold according to (3) is unknown. The value (2) is interpreted as a measure of the compactness of the sample objects from two classes on the real axis. In this paper, this measure is used to evaluate sets of initial features by the values of latent features formed from them. The estimate of compactness through the variety of structures of relations of class objects and the sample as a whole with a space dimension of 2 or more is the only one in contrast to (2).

To split (1) into a set of p_c $(p_c \geq 2)$ disjoint intervals $\{[r_u; r_v]^i\}$, $1 \leq u \leq v \leq h$, $i = 1, \ldots, p_c$ it is proposed to use the criterion from [23].

$$\left| \frac{d_{0c}(u, v)}{|K_0|} - \frac{d_{jc}(u, v)}{|K_j|} \right| \rightarrow \max, \tag{4}$$

where $d_{0c}(u, v)$, $d_{jc}(u, v)$ – is the number of representatives of the classes K_0, K_j in the interval $[r_u; r_v]^i$, $i \in \{1, \ldots, p_c\}$. Values within the interval $[r_u; r_v]^i$ in data analysis can be considered as a gradation of a nominal feature. It is believed that the set of numbers that identify p_c gradations of a nominal feature can always be one-to-one

mapped to the set $\{1,\dots,p_c\}$. If the value (4) is equal to 0, then the number of intervals is considered equal to the number of subsets of sample objects with equal values of the feature. The results of the implementation of the algorithm according to (4) is the coverage of all values (1) by disjoint intervals.

The essence of nonlinear transformations of attributes is reduced to replacing their initial values with the values of the membership function of objects to classes. In order to unify the notation, instead of $d_{0c}(u,v),(d_{jc}(u,v))$ for the interval $[r_u;r_v]^{\mu}$ in $x_c \in X(n)$, we will use $d_{0c}(t),(d_{jc}(t))$. The value of the membership function $f_c(t)$ to the class K_0 over the interval $[r_u;r_v]^t$ (gradations $t \in \{1,\dots,p_c\}$) is calculated as

$$f_c(\mu) = \frac{d_{0c}(\mu)/|K_0|}{d_{0c}(\mu)/|K_0| + d_{jc}(\mu)/|K_j|}. \tag{5}$$

Obviously, there is no order relation between gradations in the nominal measurement scale. To determine the order relation, it is proposed to use the replacement of feature gradations by the values of the function of objects belonging to classes. When calculating the membership function $f_c(\mu)$ to the class K_0 according to the gradation $\mu \in \{1,2,\dots,p_c\}$, the number of objects of the class $K_0(K_j)$ with the value μ is used as $d_{0c}(\mu)$ $(d_{jc}(\mu))$.

If for $x_c \in X(n)$ there is a gradation $\mu \in \{1,2,\dots,p_c\}$, for which

$$\frac{d_{0c}(\mu)}{|K_0|} = \frac{d_{jc}(\mu)}{|K_j|}, \tag{6}$$

holds, then the feature with $f_c(\mu) = 0.5$ cannot be used for nonlinear transformations. Denote by Z, $Z \subset X(n)$ – the set of features for which condition (6) is satisfied, by $D = \{i | x_i \in X(n)\backslash Z\}$ – is the set of feature indices that can be used for nonlinear transformations.

The boundary between class objects according to the elementary classifier for $x_c \in X(n)\backslash Z$, taking into account (5), is defined as

$$G_c = \frac{g1+g2}{2}, \tag{7}$$

where $q2 = max\{f_c(\mu)|0.5 - f_c(\mu) > 0\ ,\mu = 1,\dots,p_c\}$, $q1 = min\{f_c(\mu)|1 - f_c(\mu) < 0.5, \mu = 1,\dots,p_c\}$. Ideally, according to (7), one can obtain a correct (without errors) division of objects into classes. Correctness is guaranteed if the boundaries of each interval according to (4) contain representatives of one (K_0 or K_j) class. The feature weight x_c for objects whose values are represented through nonlinear transformations (5) can be calculated from (2) or through gradations from $\{1,2\}$ in the nominal scale. When calculating the gradation value b_{ic}, $c \in D$ for the object $S_i = \{x_{iu}\}_{u\in D}$ using (7) and taking into account the measurement scales, one of two conditions is considered: $x_{ic} \in [r_u;r_v]^{\mu}$ or $x_{ic} = \mu$. Checking the conditions is necessary to choose the values of the membership function $f_c(\mu)$ for calculating b_{ic} as

$$b_{ic} = \begin{cases} 1, f_c(\mu) < G_c, \\ 2, f_c(\mu) > G_c. \end{cases}$$

Let us denote by g^t_{0c}, g^t_{jc} the number of values of gradation $t \in \{1,2\}$ of feature $x_c \in X(n)\backslash Z$ in the description of objects of classes K_0 and K_j, respectively. The interclass difference on the basis of x_c is defined as the magnitude

$$\lambda_c = 1 - \frac{\sum\limits_{t=1}^{2} g^t_{0c} g^t_{jc}}{|K_0|\,|K_j|}. \tag{8}$$

The degree of homogeneity (measure of intra-class similarity) β_c of the values of the gradations of a feature by classes K_0, K_j is calculated by the formula:

$$\beta_c = \frac{\sum\limits_{t=1}^{2} g^t_{0c}(g^t_{0c}-1) + g^t_{jc}(g^t_{jc}-1)}{|K_0|(|K_0|-1) + |K_j|(|K_j|-1)}. \tag{9}$$

Using (8), (9), the weight of the feature $x_c \in X(n)\backslash Z$ in the nominal scale, similarly to (2), is determined as the product of intraclass similarity and interclass difference

$$w_c = \beta_c \lambda_c. \tag{10}$$

The set of admissible values of feature weights calculated by (10) belongs to the interval $(0;1]$.

To calculate the generalized estimates of objects [3] on E_0, the contributions of feature gradations are used. The contribution of gradation $t \in \{1,2\}$ of feature $x_c \in X(n)\backslash Z$ is defined as

$$\delta_c(t) = w_c\left(\frac{\alpha^0_{ct}}{|K_0|} - \frac{\alpha^j_{ct}}{|K_j|}\right), \tag{11}$$

where $\alpha^0_{ct}, \alpha^j_{ct}$ is the number of values of gradation t of feature x_c, respectively, in classes K_0 and K_j, w_c is the weight of feature x_c according to (10). The generalized estimate of the object $S_r \in E_0$ according to the description in the nominal measurement scale $S_r = \{b_{ri}\}_{i\in D}$ on the set $X(n)\backslash Z$ and contributions (11) is calculated as

$$R(S_r) = \sum_{i\in D} \delta_i(b_{ri}). \tag{12}$$

The basic concept for the formation of informative feature sets is feature stability. To calculate the stability, the values of the membership function are used. Let in the description of the object $S_r \in K_0 \cup K_j$ the initial values of features from $X(n)\backslash Z$ be replaced by the values of the membership function $S_r = \{f_c(b_{ri})\}_{i\in D}$ according to (5). The stability of the feature $x_c \in X(n)\backslash Z$ is calculated as

$$\Omega(c) = \frac{1}{h}\sum_{r=1}^{h} \begin{array}{l} f_c(b_{rc}), f_c(b_{rc}) > 0.5, \\ 1 - f_c(b_{rc}), f_c(b_{rc}) < 0.5. \end{array} . \tag{13}$$

The set of admissible values (13) belongs to $(0.5;1]$.

V. GROUPING FEATURES ACCORDING TO THE HIERARCHICAL AGGLOMERATIVE ALGORITHM

The purpose of using a hierarchical agglomerative algorithm is to split the feature set $X_n \backslash Z$ into disjoint subsets according to the descriptions of objects from $K_0 \cup K_j$ for the synthesis of latent features. It is considered that for each $x_i \in X(n) \backslash Z$ the weight w_i is calculated according to (10) and the values of the contributions $\delta_i(t), t \in \{1, 2\}$ according to (11).

Features of the implementation of the process of formation of latent features by the algorithm of hierarchical agglomerative grouping are in:

– choosing the first feature in the group;

– a set of rules for including (not including) a feature in a group;

– calculating the values of generalized estimates of objects (12) by the contributions of features (11) from the group.

Denote by P, $ALOMAT$ the set of feature indices, respectively from $X(n) \backslash Z$ and the groups formed by the algorithm. The implementation of the algorithm step by step will be as follows.

Step 1. $P = \{i | x_i \in X(n) \backslash Z\}$. $son = 0$.
Step 2. Calculate $crit = 10$. $u = \arg \max_{i \in P} w_i$.
$ALOMAT = \{u\}$. $son = son + 1$.
Cycle according to $t \in \{1, \ldots, h\}$ $R(S_t) = \delta_u(b_{tu})$.
End of cycle; $cr1 = 10$. $P = P/\{u\}$.
Step 3. Cycle according to $u \in P$.
Cycle according to $t \in \{1, \ldots, h\}$ $d_t = R(S_t) + \delta_u(b_{tu})$.
End of cycle;
$M_1 = \sum_{S_t \in K_0} d_t$. $M_2 = \sum_{S_t \in K_j} d_t$.
$M_1 = M_1/|K_0|$. $M_2 = M_2/|K_j|$. θ=0. γ=0.
Cycle according to $t \in \{1, \ldots, h\}$ if $S_t \in K_0$,
then $\theta = \theta + |d_t - M_1|, \gamma = \gamma + |d_t - M_2|$.
Else $\theta = \theta + |d_t - M_2|, \gamma = \gamma + |b_t - M_1|$.
End of cycle**;**
if $\theta/\gamma < cr1$, then $cr1 = \theta/\gamma$, $q = u$.
End of cycle**;**
Step 4. if $cr1 < crit$, then $crit = cr1$. $P = P/\{q\}$.
$ALOMAT = ALOMAT \cup \{q\}$. $cr1 = 10$.
Cycle according to $t \in \{1, \ldots, h\}$
$R(S_t) = R(S_t) + \delta_q(b_{tq})$.
End of cycle**;**
else output $\{R(S_t)\}_{t \in \{1,\ldots,h\}}$, $ALOMAT$.
Step 5. if $|P| \geq 2$, then go 2; else output son.
Step 6. End.

The set of values $\{R(S_t)\}_{t \in \{1,\ldots,h\}}$, obtained at step 4 of the algorithm, form descriptions of objects K_0, K_j by the set $Y^j(son) = (y_1^j, ..., y_{son}^j)$ in the latent feature space, the dimension of which is $son < n$. Each $y_1^j \in Y^j(son)$ corresponds to a set of $X^j(n_i) \subset X(n) \backslash Z$. The maximum value (2) over $y_c^j \in Y^j(son)$ serves as a condition for choosing the set $X^j(n_c)$ as informative on $K_0 \cup K_j$. The informative set is further considered as the basis for the formation of an ensemble of algorithms (elementary classifiers) using stacking technology.

Reducing the combinatorial complexity of calculations when choosing a latent space is achieved by applying the principle of dynamic programming in the process of implementing a hierarchical agglomerative algorithm. The rules for selecting an informative set of raw features by the maximum value of the measure of compactness (2) by latent indicators are used.

VI. COMPUTATIONAL EXPERIMENT

For the experiment, the CICDDOS2019 [28] dataset was used, containing a description of 12 types of DDOS attacks and normal traffic through the TCP/UDP application layer protocols.

From the data set, 12 samples were formed in pairs of classes (normal traffic, type of DDOS attacks). For each sample, normal traffic was represented by the same objects. On 12 samples, a nonlinear transformation of the values of raw features of objects in {1, 2} was carried out and 12 sets of informative features were obtained using the hierarchical agglomerative algorithm from clause V. We checked the truth of the statement that the values of generalized estimates (of a latent feature) have the values of the measure of compactness (2) on the informative set of features higher than on the entire set. The results of the verification of the assertion are given in Table 2.

Table 2. Values of compactness measure by types of DDOS attacks.

№	Type of DDOS attacks	Number of informative features	Measure of compactness by (2) on features	
			all	informative
1	DNS	9	0.9535	0.9611
2	LDAP	3	0.9786	0.9804
3	MSSQL	3	0.9898	0.9922
4	NetBIOS	6	0.9304	0.9399
5	SNMP	4	0.9662	0.9729
6	WebDDOS	7	0.8035	0.9086
7	UDP-Lag	16	0.8493	0.9300
8	SYN	3	0.9160	0.9970
9	NTP	3	0.9861	0.9884
10	UDP	3	0.9951	0.9965
11	SSDP	3	0.9800	1.0000
12	TFTP	3	0.9910	0.9935

Table 2. shows that each type of DDOS attacks has its own set of informative features, the compactness of (2) on which is higher than for the entire set. The value of the compactness measure (2) is interpreted as an estimate of the complexity of attack detection. The lower the score value, the more difficult it is to detect attacks. Based on this rule, the WebDDOS type with a score of 0.9086 on 7 features should be considered the most difficult to detect. Demonstration of informative sets of features by types

and the possibility of analyzing them for enumeration is contained in Table 3.

Table 3. Sets of informative features by types of DDOS attacks.

№	Type of DDOS attacks	Sets of informative features
1	DNS	Bwd Packet Length Mean
		Bwd Packet Length Std
		Bwd IAT Mean
		Bwd IAT Min
		SYN Flag Count
		Active Min
		Idle Std
		Idle Max
2	LDAP	Bwd Packets/s
		FIN Flag Count
		Active Min
3	MSSQL	Bwd Packets/s
		FIN Flag Count
		Active Min
4	NetBIOS	Bwd IAT Min
		Bwd Packets/s
		SYN Flag Count
		Active Mean
		Active Min
		Idle Mean
5	SNMP	Bwd Packets/s
		Active Mean
		Active Min
		Idle Mean
6	WebDDOS	Fwd Packet Length Mean
		Flow IAT Max
		Fwd IAT Mean
		Bwd URG Flags
		ACK Flag Count
		Init Win bytes forward
		Idle Max
7	UDP-Lag	Frequency
		Fwd Packet Length Mean
		Bwd Packet Length Mean
		Flow Packets/s
		Flow IAT Min
		Bwd IAT Mean
		Bwd IAT Min
		Fwd Packets/s
		FIN Flag Count
		SYN Flag Count
		ACK Flag Count
		URG Flag Count
		Subflow Bwd Bytes
		Init Win bytes forward
		Act data pkt fwd

continued on next page

continued from previous page

#	Type of DDOS attacks	Sets of informative features
		SimillarHTTP
8	SYN	FIN Flag Count
		Subflow Bwd Bytes
		Idle Max
9	NTP	Total Length of Bwd Packets
		Active Min
		Idle Std
10	UDP	Bwd Packets/s
		Active Min
		Idle Mean
11	SSDP	Bwd Packets/s
		Active Min
		Idle Mean
12	TFTP	Total Length of Bwd Packets
		Min Packet Length
		FIN Flag Count

Analysis of the results from Table 2 and Table 3 shows that the easiest what? to detect are UDP and SSDP types with the same set of three informative features. An explanation for the stability of the compositions of informative feature sets by types of DDOS attacks is the use of the principle of dynamic programming in the hierarchical agglomerative algorithm when selecting.

VII. CONCLUSIONS

For analyzing the structure of relations between types of DDOS attacks, a unique methodology has been developed –using the values of the stability of signs for pairs of classes “normal traffic” and “type of DDOS attacks”.

The rationale for stability is the small variability of its values in samples from the general population, the property of invariance to the scale of measurements of features.

A technology for selecting and analyzing informative sets of features using nonlinear transformations based on membership functions and calculating generalized estimates of objects (latent features) for pairs of non-overlapping classes is proposed. The effectiveness of applying the rules of the hierarchical grouping algorithm for the formation of latent features based on sets of initial ones is proved. The maximum measure of compactness from the set of latent features in the description of objects of two classes “normal traffic” and “type of DDOS attacks” serves as an estimate of the detection complexity. The values of the estimate in (0;1] simplify its interpretation for practical use. It is shown that each type of DDOS attacks has its own set of informative features that determine the complexity of its detection. Estimates of the complexity of detection are recommended to justify the choice of anti-malware tools.

References

[1] V. Datla Anurag, A. Ravi, S. Venkata, B. Venkatesh, and R. Kannadasan, “Detection of ddos attacks using machine learning techniques: A hybrid

approach," ICT Systems and Sustainability, p. 439–446, 2020. [Online]. Available: https://doi.org/10.1007/978-981-15-8289-9_42
[2] S. Rezaei and X. Liu, "Deep learning for encrypted traffic classification: An overview," IEEE Communications Magazine, vol. 57, pp. 76–81, 2019. [Online]. Available: https://doi.org/10.1109/MCOM.2019.1800819
[3] A. Finamore, M. Mellia, M. Meo, and D. Rossi, "Kiss: Stochastic packet inspection classifier for udp traffic," IEEE/ACM Transactions on Networking, vol. 18, pp. 1505–1515, 2010. [Online]. Available: https://doi.org/10.1109/TNET.2010.2044046
[4] L. Vu, C. Bui, Q. Nguyen, and D. Rossi, "A deep learning based method for handling imbalanced problem in network traffic classification." December 2017, pp. 333–339. [Online]. Available: https://doi.org/10.1145/3155133.3155175
[5] G. Aceto, D. Ciuonzo, A. Montieri, and P. A, "Multi-classification approaches for classifying mobile app traffic," Journal of Network and Computer Applications, vol. 57, pp. 131–145, 2018. [Online]. Available: https://doi.org/10.1016/j.jnca.2017.11.007
[6] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," 2018, pp. 108–116. [Online]. Available: https://doi.org/10.5220/0006639801080116
[7] A. Agarwal, M. Khari, and R. Singh, "Detection of ddos attack using deep learning model in cloud storage application," Wireless Personal Communications, 2021. [Online]. Available: https://doi.org/10.1007/s11277-021-08271-z
[8] D. Jisa and T. Ciza, "Detection of distributed denial of service attacks based on information theoretic approach in time series models journal of information security and applications," Journal of Information Security and Applications, vol. 55, 2020. [Online]. Available: https://doi.org/10.1016/j.jisa.2020.102621
[9] S. Gómez, B. Martínez, J. Antonio, and H. Luis, "Ensemble network traffic classification: Algorithm comparison and novel ensemble scheme proposal," Computer Networks, vol. 127, pp. 131–145, 2017. [Online]. Available: https://doi.org/10.1016/j.comnet.2017.07.018
[10] P. Wang, C. Xuejiao, Y. Feng, and S. Zhixin, "A survey of techniques for mobile service encrypted traffic classification using deep learning," IEEE Access, vol. 7, pp. 54 024–54 033, 2019. [Online]. Available: https://doi.org/10.1109/ACCESS.2019.2912896
[11] A. B. Petrovsky and V. N. Lobanov, "Multi-criteria choice in the space of high-dimensional features : Paks-m multi-method technology," Artificial intelligence and decision making, no. 3, pp. 92–104, 2014.
[12] A. B. Petrovsky, "Indicators of similarity and differences of multi-attribute objects in metric spaces of sets and multisets," Artificial intelligence and decision making, no. 4, pp. 78–94, 2017.
[13] N. Miloslavskaya, A. Tolstoy, and S. Zapechnikov, "Taxonomy for unsecure big data processing in security operations centers," Aug.22-24 2016, pp. 154–159. [Online]. Available: https://doi.org/10.1109/W-FiCloud.2016.42
[14] N. Miloslavskaya and A. Makhmudova, "Survey of big data information security," vol. 8, Aug.22-24 2016, pp. 133–138. [Online]. Available: https://doi.org/10.1109/W-FiCloud.2016.38
[15] S. F. Madrakhimov, K. T. Makharov, and M. Y. Lolayev, "Data preprocessing on input," AIP Conference Proceedings, vol. 1, no. 16, pp. 29–41, 2021. [Online]. Available: https://doi.org/10.1063/5.0058132
[16] B. Naveen and S. Manu, "Evaluating the impact of feature selection methods on the performance of the machine learning models in detecting ddos attacks," Romanian journal of information science and technology, vol. 23, no. 3, p. 250 – 261, 2020.
[17] I. Sharafaldin, A. H. Lashkari, H. Saqib, and A. Ghorban, "Developing realistic distributed denial of service (ddos) attack dataset and taxonomy," in In Proceedings of the 2019 International Carnahan Conference on Security Technology (ICCST). IEEE, Oct. 1-3, pp. 1–8. [Online]. Available: https://doi.org/10.1109/CCST.2019.8888419
[18] S. E. Mahmoud, L. Nhien-An, D. Soumyabrata, and D. J. Anca, "Ddosnet: A deep-learning model for detecting network attacks," in 2020 IEEE 21st International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM). IEEE, 31 Aug.-3 Sept. 2020, pp. 1–8. [Online]. Available: https://doi.org/10.1109/WoWMoM49955.2020.00072
[19] M. S. Yin, P. A. Pye, and S. H. Aye, "A slow ddos attack detection mechanism using feature weighing and rankingn," Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management Singapore, pp. 4500–4509, March. 7-11, 2021.
[20] A. H. Lashkari, D. G. Gerard, M. M. Mamun, and A. A. Ghorbani, "Characterization of tor traffic using time based features," 2017, pp. 253–262. [Online]. Available: https://doi.org/10.5220/0006105602530262
[21] N. A. Ignatiev, "On nonlinear transformations of features based on the functions of objects belonging to classes," Pattern Recognition and Image Analysist, vol. 2, no. 31, pp. 197–204, June 30 2021. [Online]. Available: http://dx.doi.org/10.1134/S1054661821020085
[22] P. N. Matheus, F. C. Luiz, L. Jaime, and L. P. Mario, "Long short-term memory and fuzzy logic for anomaly detection and mitigation in software-defined network environment," 2020, pp. 83 765–83 781. [Online]. Available: https://doi.org/10.1109/ACCESS.2020.2992044
[23] E. N. Zguralskaya, "Sustainability of dividing data in intervals in the problems of recognition and searching for hidden laws," Proceedings of the Samara Scientific Center Russian Academy of Sciences, vol. 3, no. 4, pp. 451–455, 2018.
[24] N. Miloslavskaya, "Information security management in socs and sics," Journal of Intelligent Fussy Systems. - IOS Press (Netherlands), vol. 35, no. 3, pp. 2637–2647, 2018. [Online]. Available: https://doi.org/10.3233/JIFS-169615
[25] N. A. Ignatyev and M. A. Rakhimova, "Formation and analysis of sets of informative features of objects by pairs of classes," Artificial intelligence and decision making, no. 4, pp. 18 – 26, 2021. [Online]. Available: http://dx.doi.org/10.14357/20718594210402
[26] N. G. Zagoruiko, I. A. Borisova, and O. A. Kutnenko, "Constructing a concise description of data using the competitive similarity function," Siberian Journal of Industrial Mathematics, vol. 1, no. 16, pp. 29–41, 2013.
[27] N. A. Ignatiev, "Structure choice for relations between objects in metric classification algorithms," Pattern Recognition and Image Analysist, vol. 28, no. 4, pp. 695–702, 2018. [Online]. Available: https://doi.org/10.1134/S1054661818040132
[28] "Ddos evaluation dataset (cic-ddos2019)," 2018. [Online]. Available: https://www.unb.ca/cic/datasets/ddos-2019.html

NIKOLAI A. IGNATEV professor of the Department of Algorithms and Programming Technologies of the National University of Uzbekistan. Research interests: mathematical modeling and data mining methods. Published over 90 scientific articles and three monographs.

ERKIN R. NAVRUZOV pursuing a PhD degree at the National University of Uzbekistan. His research areas are:intelligent Machine learning and security of information technologies.