IVAN CVITIĆ, Ph.D. Candidate[1]
(Corresponding author)
E-mail: ivan.cvitic@fpz.hr
DRAGAN PERAKOVIĆ, Ph.D.[1]
E-mail: dragan.perakovic@fpz.hr
MARKO PERIŠA, Ph.D.[1]
E-mail: marko.perisa@fpz.hr
SINIŠA HUSNJAK, Ph.D.[1]
E-mail: sinisa.husnjak@fpz.hr
[1] University of Zagreb,
Faculty of Transport and Traffic Sciences
Vukelićeva 4, 10000 Zagreb, Croatia

Information and Communication Technology
Review
Submitted: 27 Sep. 2018
Accepted: 12 June 2019

AN OVERVIEW OF DISTRIBUTED DENIAL OF SERVICE TRAFFIC DETECTION APPROACHES

ABSTRACT

The availability of information and communication (IC) resources is a growing problem caused by the increase in the number of users, IC services, and the capacity constraints. IC resources need to be available to legitimate users at the required time. The availability is of crucial importance in IC environments such as smart city, autonomous vehicle, or critical infrastructure management systems. In the mentioned and similar environments the unavailability of resources can also have negative consequences on people's safety. The distributed denial of service (DDoS) attacks and traffic that such attacks generate, represent a growing problem in the last decade. Their goal is to disable access to the resources for legitimate users. This paper analyses the trends of such traffic which indicates the importance of its detection methods research. The paper also provides an overview of the currently used approaches used in detection system and model development. Based on the analysis of the previous research, the disadvantages of the used approaches have been identified which opens the space and gives the direction for future research. Besides the mentioned this paper highlights a DDoS traffic generated through Internet of things (IoT) devices as an evolving threat that needs to be taken into consideration in the future studies.

KEY WORDS

network traffic anomaly; network-based attack; service availability; denial of service; network anomaly detection;

1. INTRODUCTION

The development of public, packet-oriented, communication network (Internet) followed by the increasing number of users and IC services has resulted in an increase of the amount of transferred data [1, 2]. Data that are stored, processed, and transmitted through the IC system are often the target of illegitimate users. Their goal and purpose is unauthorized access to sensitive data or disabling access to IC system resources for legitimate users [3]. The above results in an increased need for research in the field of security and protection of IC systems over the past decades.

The goal of IC system protection implies achieving and maintaining the required level of basic security principles. The basic principles of security are presented with a CIA model that includes confidentiality, integrity, and availability of IC resources [3]. According to [4], the availability principle is defined as a probability that the requested service (or other IC resource) will be available to a legitimate user in the required time. There are many factors that have the potential to negatively affect the availability of IC resources and can be classified according to the source activity (internal and external) and the agent (human, environment and technology) [5]. One of these factors whose trend has been steadily increasing over the last ten years is the network-oriented DDoS attack, or DDoS traffic as a means of attack implementation [6]. Traffic generated by DDoS attack is aimed at exploiting the disadvantages of IC systems responsible for processing and transmitting data such as communication links, active network equipment (routers, switches, firewalls, etc.) and devices intended for processing customer requests and delivery of services (servers). The primary disadvantage that a DDoS traffic exploits is the capacity limitations of communication link, network equipment, or servers [7]. Nowadays, DDoS traffic is causing a number of difficulties in electronic business, such as complete unavailability or degradation of service quality. Such state of service can have negative implication on the targeted organization reflected in the reputation loss, user loss and finally in economic loss.

The importance of DDoS traffic negative effects are widely recognized, and there are a number of studies whose goal is to successfully detect the mentioned traffic class. The aim of this research is to analyse the so far applied approaches for model and system of

DDoS traffic detection development through relevant and current scientific and technical literature. Based on the analysis, the exact disadvantages of observed approaches will be identified, and they will define the possibilities and the scope for future research of the problem area. Also, the possibilities for detection of DDoS traffic generated using IoT devices will be analysed as a new and emerging way of causing more intensive DDoS attacks which distinguishes this research from the similar ones.

2. NETWORK-BASED ATTACKS AIMED AT IC RESOURCE AVAILABILITY

According to [8] the network-based attacks are identified as anomalies of the network traffic. Anomalies are network traffic patterns that differ from the well-defined patterns of normal traffic. Denial of Services (DoS) implies a general class of network-based attacks targeting the availability of IC resources. According to the implementation method the DoS attacks can be divided into two general categories [9]: (1) single source denial of service (SSDoS) and (2) distributed denial of service (DDoS).

The source of SSDoS attack is one computer or device in the network. In DDoS attacks multiple devices are coordinated with the aim of generating large amounts of DDoS traffic to the target destination [7]. The DDoS attacks represent a growing problem in the recent years. The negative effects of such attacks on IC-based services and resources are reflected in the degradation of the service quality, disruption of service provider credibility, user loss and financial loss [10, 11].

2.1 Classification of DDoS attack methods

Since the first appearance of DDoS attack in 2000, many ways of implementing it have been developed and used. Numerous authors have suggested the taxonomy of DDoS implementation methods based on a variety of factors. The authors of research [12] differentiate DDoS implementation methods by the degree of attack automation, vulnerability utilization, impact, and attack speed dynamics. According to the dynamics of the attack speed, it is possible to classify them as high intensity of DDoS traffic (high rate) and low intensity of DDoS traffic (low rate). The goal of high rate DDoS attack is to flood the destination or communication link using a large number of network packets. Contrary to the high rate DDoS attack, the goal of low rate attacks is to generate traffic that is similar to normal traffic. It makes it more difficult to detect this kind of attack and has the potential to degrade the quality of service [13].

The methods for implementing DDoS attacks can also be classified depending on the TCP/IP (Transmission control protocol/Internet protocol) layer they are targeting. Accordingly, the infrastructure and application of DDoS attacks can differ [14]. The infrastructure of DDoS attacks are focused on resource flooding on the network and transport TCP/IP layer by exploiting vulnerabilities or shortcomings in communication protocols such as TCP, UDP (User datagram protocol) or ICMP (Internet control message protocol). The infrastructure attacks aim to exploit the capacity of a communication link or the capacity of server resources [6]. The application of DDoS attacks pose a continuing threat to services available over the Internet by using legitimate HTTP (Hypertext transfer protocol) protocols for exploiting the capacity of the destination web server [15]. There are often attempts to imitate flash crowd events on popular web sites, making it difficult to detect this attack method [16].

2.2 Trends of DDoS traffic

For the purposes of research and development of DDoS traffic detection methods it is necessary to continuously analyse the trends of the used protocols and the traffic intensity with the aim of timely reaction to future attacks.

The largest number of infrastructure layer attacks in 2013 and 2014 was performed using the TCP protocol with exploitation of the SYN flag (31.22% and 25.73%). The SYN flag represents one of six possible TCP header flags (ACK, SYN, URG, FIN, RST, and PSH) whose function is to synchronize sequential packet numbers when initiating a TCP session, and it is often used for the implementation of DDoS attack. Except SYN and other TCP header flags *Figures 1* and *2* show other protocols or protocol parameters that were used in DDoS attacks based on infrastructure and application layer.

After 2014, changes in the frequency of certain infrastructure layer protocols use have been noticed. Since the third quarter of 2014, the share of SYN-based attacks has been in decline, and the use of other protocols such as UDP, NTP (Network time protocol) and DNS (Domain name system) is rising.

Figures 1 and *2* show the frequency of certain protocols use in the implementation of DDoS attacks, quarterly for the time period from the first quarter of 2013 to the fourth quarter of 2017. The analysed data were taken from the company Aakmai Inc., one of the leading firms for the protection against DDoS attacks worldwide. From *Figure 3*, the infrastructure layer attacks are more frequent in all 20 analysed quarters and have a continuous growth trend (76.54% - 99.43%), unlike application layer attacks (23.46% - 0.57%) whose trend is declining. From the

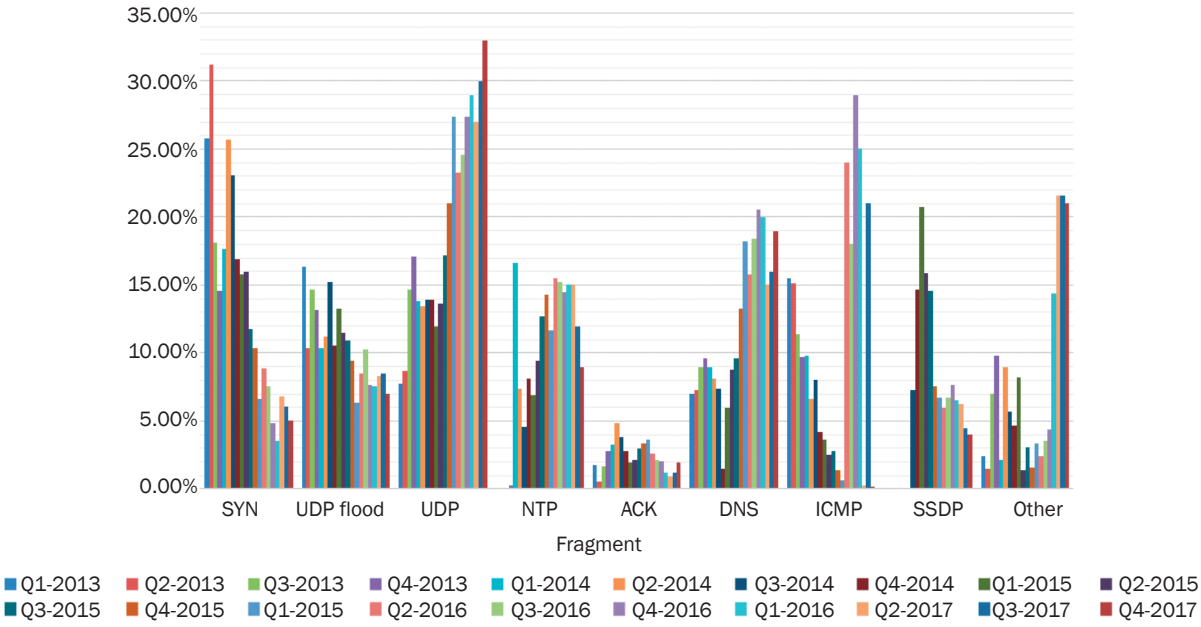

Figure 1 – The frequency of infrastructure layer protocols application in implementing DDoS attacks [17-35]

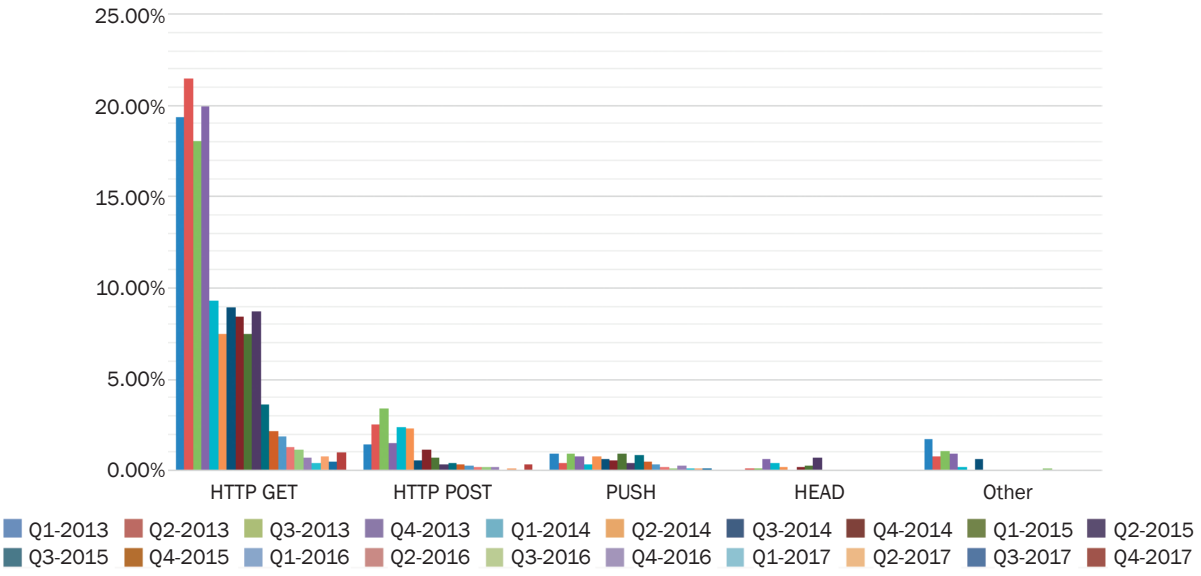

Figure 2 – The frequency of application layer protocols in implementing DDoS attacks [17-35]

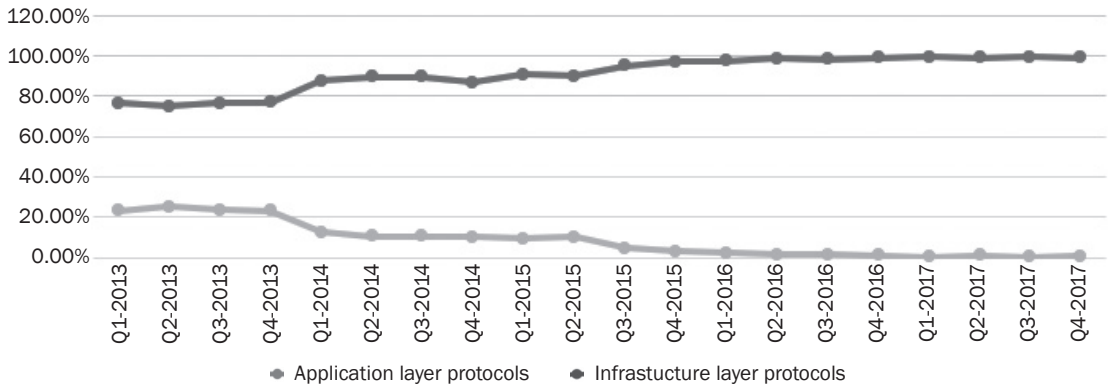

Figure 3 – Share of infrastructure and application level protocols in DDoS attack [17–35]

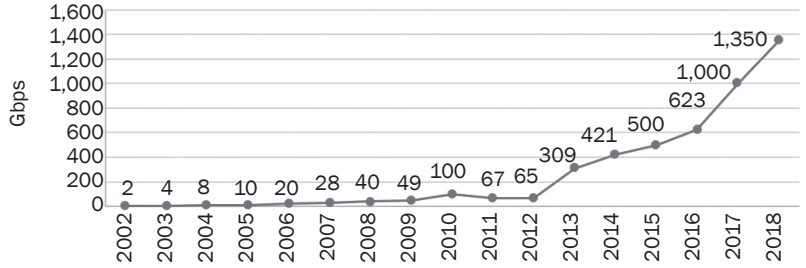

Figure 4 – Intensity of generated DDoS traffic over time period 2002-2018 [17, 27, 31]

data shown in *Figure 4*, there is a noticeable increase in the trend of traffic intensity generated by DDoS attacks since 2012. The most intense attack was recorded at the end of 2016 with the amount of 623 Gbps.

The reason for the current trend is the development of Cloud computing (CC) concept, which also implies the use of processing capacity considerably larger for the inbound traffic than is the case of traditional IC systems. Accordingly, successful disruption of IC resource availability in the CC environment requires generating a higher DDoS traffic intensity [36].

An additional cause is the technological development of new concepts such as IoT that enables exploitation of a large number of inadequately protected devices for generating high intensity DDoS traffic to the targeted destination [17, 37]. The concept of IoT is increasingly used in different economic sectors as well as for critical infrastructure management where the availability of IC resources is of key importance. Accordingly, DDoS attacks directed to IC resources in the IoT environment within critical infrastructure have the potential to cause significant damage but also endanger the end users' security [5].

3. THE APPROACHES USED IN DDoS TRAFFIC DETECTION

For the past two decades, numerous studies have been focused on the development of methods, models and systems that can detect DDoS traffic in real time. Despite the mentioned, the number of DDoS attacks and the intensity of DDoS traffic are steadily increasing, which is the reason for further research in the detection of this type of security threats [13].

Studies define several approaches to DDoS traffic detection. Generally, it is possible to divide them into two basic classes, based on the patterns and based on anomalies [38]. The research [10], among others, identifies the entropy-based approach and the research [12] identifies the possibilities of using a hybrid approach of DDoS traffic detection. The methods based on a pattern applied comparison of incoming traffic with predefined profiles and samples of known network anomalies [39]. The pattern-based DDoS traffic detection can be performed in three ways: based on the signature of the known attacks, based on the rules (if-then) and based on the state and transition [7]. The advantage of this detection approach is the high detection rate of the already known DDoS attacks with a small number of false positive and false negative results. The disadvantage is the inability to detect new and unknown attacks, that is, those attacks that are not in the database whose records are compared with the incoming traffic patterns. Given the dynamics of the problem area, it is important that the detection methods are able to detect unknown patterns of DDoS traffic [38].

Contrary to the above, an approach based on the detection of a network traffic anomaly uses predefined models of normal traffic which are then compared with the incoming traffic [12]. This detection approach has been developed to overcome the shortcomings of pattern detection approaches [10]. If the incoming traffic differs significantly from the defined normal traffic model, then the incoming traffic is identified as an anomaly or DDoS traffic [40]. The advantage of network traffic anomaly detection compared to patterns-based detection is in the ability to detect unknown attacks. The main disadvantage of anomalies-based detection is determining the threshold values between normal traffic and anomalies [12, 41]. The network traffic anomalies are detected when the values of the current traffic flow or other selected parameters exceed the predefined threshold value of the normal traffic model. A low-defined threshold value can cause many false positive results, and the highly-defined threshold value can lead to a large number of false negative results [42].

Numerous approaches based on scientific methods have been used to detect DDoS traffic. In current scientific literature, the most commonly used are approaches based on statistical and information theory methods, machine learning methods, and soft computing methods [43].

3.1 Detection of DDoS traffic based on statistical and information theory methods

Statistical traffic characteristics can be utilized to differentiate between normal and DDoS traffic. Statistical-based approaches are based on the use of statistical methods in determining the normal traffic model.

After that, it can be statistically determined whether a new traffic instance (flow, packet or package set) corresponds to a defined model [43]. The commonly used DDoS traffic detection methods from statistics and information theory domain are deviation, cumulative sum, correlation, entropy, and covariance [7]. The specificities and differences in the studies that used statistical and information theory methods can be seen in *Table 1*.

Self-similarity and long-range dependence (LRD) of network traffic are often used in statistical processing and DDoS traffic detection, as can be seen from numerous studies such as [13, 40, 44, 45]. Data traffic under normal conditions maintains an LRD property which implies loss or reduction of LRD property in the event of anomalies in the communications network such as the occurrence of DDoS traffic [40]. Therefore, by analysing LRD property of the incoming traffic it is possible to detect DDoS traffic. Self-similarity and LRD are expressed by the Hurst parameter (H), also called the long-term dependence index, and it is measured by statistical estimators such as autocorrelation, variance aggregation, wavelet, R/S method and similar [45]. The challenge in determining LRD property to determine the time period within traffic will be analysed [40]. If the time period is too short, the results of the analysis will not be valid due to the insufficient volume of traffic to determine the degree of LRD, while a too long time period will cause the inability to detect short-term anomalies [44]. In addition, the disadvantage of this detection is predefined static limit value of the Hurst parameter which results in the detection of DDoS traffic only when its intensity causes a change in the value of the Hurst parameter above a defined threshold.

Studies like [10, 41, 46, 47] use entropy as the primary DDoS traffic detection method supported by other statistical methods. Studies [47] and [48] use entropy and Pearson's chi-square correlation test in the function of measuring the statistical properties of the packet header parameter values. As an example is the use of the above methods for analysing source IP (Internet protocol) addresses in the given incoming packet set [47]. Four datasets collected in different IC environments were used in the research. The detection accuracy varies significantly depending on the dataset. The lack of research is visible in defining the chi-square test threshold that can result in a large number of false negative or false positive results. An additional disadvantage, according to the authors of the research, is the choice of packet header parameters whose values will be analysed because it is necessary to have good knowledge of what parameters will affect DDoS traffic. In addition, according to [49] correlation methods such as Pearson, Spearman and Kendall are considered inadequate in DDoS traffic detection because they often exhibit a high correlation rate between different objects or instances of traffic. The detection of DDoS traffic based on entropy was used in the research [10]. The developed detection model is based on the traffic flow aggregation and the use of the fast entropy method. If the entropy value falls below the threshold value, the observed traffic flow is considered as DDoS traffic. The determination of the threshold value in this research is adaptive and its adjustment is based on the mean value and the standard deviation of the number of traffic flows in the observed time interval.

The frequently used statistical method in detecting DDoS traffic is multivariate correlation analysis (MCA). The examples of MCA methods use are visible in the research [49-51]. The MCA method is used because of its advantage over other statistical methods such as a small number of false positive results [50]. The disadvantage are the user-defined threshold values [49]. Research [49] uses two datasets, CAIDA DDoS 2007 and DARPA 2000, for the validation of the proposed detection model, and research [50] uses CAIDA DDoS 2007, KDD CUP 99 and TUIDS datasets. The detection accuracy in both models depends significantly on the correlation threshold value between the legitimate and DDoS traffic. High accuracy and a small number of false positive results for each dataset requires a different threshold value, where the problem of defining a threshold value on a new set of data arises. In addition, all the analysed research implies an increase in the number of false positive results depending on the number of accurately detected instances of DDoS traffic [49-51]. In order to detect DDoS traffic using MCA, but also using other statistical methods, great importance lies in the selection of parameters of traffic that will be analysed because not all parameters have equal importance in the analysis and classification of the network traffic [50]. A greater number of used parameters can increase the detection accuracy but requires more processing resources, which often prevents real-time detection.

3.2 Detection of DDoS traffic based on machine-learning methods

The use of machine-learning methods is one of the approaches to DDoS traffic detection. The reason for their use is the advantage over the pattern-based detection method because the human factor's impact is significantly reduced in the overall DDoS traffic detection process [52]. The machine-learning methods can be classified on supervised (existing knowledge is used to classify the future unknown instances) and unsupervised (attempts to determine the corresponding instance class without prior knowledge) [7]. Examples of supervised machine-learning methods commonly used in DDoS traffic detection are decision trees, k-nearest neighbour (kNN), support

Table 1 – Studies using statistical and information theory methods in DDoS traffic detection

Ref	Method	Estimators / performance measuring	DDoS type	Dataset	No. of used features	Window size	Accuracy
[10]	Fast Entropy	N/A	Unknown	CAIDA	6	N/A	Proof that DDoS traffic causes low entropy
[44]	Optimization method	Hurst index (SOSS), d (FARIMA)	Volumetric	KSU (King Saud University)	11	15,20, 30 min	Proof that LRD fails in volumetric DDoS traffic
[45]	Wavelet-based Multi-Resolution Analysis (MRA)	Hurst index	Pulsating DDoS (PDDoS); Low Rate DDoS (LRDDoS)	Simulated dataset	N/A	N/A	Proof that in case of PDDoS traffic self-similarity property is higher than in normal traffic
[46]	Entropy; Artificial Neural Network with Genetic Algorithm	N/A	Application layer DDoS	Simulated dataset; CAIDA 2007; DARPA 2009; BONESI	N/A	N/A	98.31%
[48]	Entropy/Chi-square	N/A	unknown	Yatsushiro National College of Technology	1	1,000 and 5,000 packets	Proof of variation of chi-square and entropy in DDoS traffic
[49]	Multivariate data analysis; FFSc	N/A	Ping ICMP flood; TCP SYN flood, and HTTP	CAIDA 2007, DARPA 2000	3	N/A	100% (FPR - 0%; FNR - 0%) on CAIDA 2007
[50]	Multivariate correlation analysis	N/A	unknown	CAIDA, TUIDS and KDD CUP 99	3	N/A	86%-98% for the TUIDS; 60% to 98.85% for the KDD CUP 99; 98.8% to 99.6% for the CAIDA
[51]	Multivariate Correlation Analysis	N/A	Teardrop, Smurf, Pod, Neptune, Land and Back attacks	KDD Cup 99	N/A	N/A	95.20%

vector machines (SVM) and naïve Bayes classifier. Unsupervised machine-learning methods commonly used in DDoS traffic detection are fuzzy C means and k-mean clustering [53]. The overview of the studies that used machine-learning methods is given in *Table 2*, which shows the differences in the used method, ways in performance measuring, DDoS type detection, used dataset, number of used features, and accuracy of the developed model.

The use of the decision-tree method and the naïve Bayesian classifier is visible in research [54]. The authors use the above method for detecting DDoS traffic in the CAIDA dataset. The results of the research show a high degree of efficiency in the application of these methods. The accuracy of decision-tree detection is 99%, and the naïve Bayes classifier 97%. The use of the same method over another dataset (NSL KDD) shows less detection accuracy of the naïve Bayesian classifier (<90%) while the accuracy of the tree-detection decision is the same as in the previous research [55].

Research [56] analyses the application of fuzzy C mean, SVM, kNN, k means, decision trees, and the naïve Bayesian classifier at the CAIDA dataset. All the analysed methods demonstrate high detection accuracy (>95%), where SVM, kNN and decision-tree methods have a high false positive result. Research [57] uses the singular value decomposition (SVD) method in the DDoS traffic detection model development. The developed model uses a total of 41 parameters based on which traffic classification is performed on normal and DDoS traffic. The results of the research show high detection accuracy on the KDD-CUP 1999 dataset compared to the use of machine-learning methods such as kNN, random forest, and bagging. Threshold values between normal and DDoS traffic are also user-defined as in the statistical approaches. Additionally, SVD method shows less change in the detection accuracy (99.4% - 99.8%) under the influence of the threshold values, in contrast to other methods used.

3.3 Detection of DDoS traffic based on soft computing methods

The advantages of soft computing methods compared to the previously described is tolerance on imprecision, uncertainty, incompleteness and partial authenticity of the input data. The robustness and efficiency of these methods have been proven in solving many complex problems like pattern matching. *Table 3* shows comparison of the studies that used soft computing methods in DDoS traffic detection.

Soft computing approach is effective in solving problems where information about the problem is incomplete, and the possible problem solution is not exact [7]. This is the reason for the frequent use of this group of methods in DDoS traffic detection, where artificial neural networks (ANN) and fuzzy logic are often used, as can be seen from several studies like [58-62]. The effectiveness of the use of artificial neural networks can be seen from the research [58, 60]. The authors of research [58] have developed a model of DDoS traffic detection SPUNNID (Statistical Pre-Processor & Unsupervised Neural Net based Intrusion Detector). The model uses eight parameters of network packets for detecting high intensity DDoS traffic (UDP, SYN, and ICMP flooding). The parameters used were selected due to the statistic changes of their values under the influence of DDoS traffic in relation to normal traffic. Based on the selected parameters, learning, testing and validation of artificial neural network on a dataset generated in a simulated environment were conducted. The validation results of the model show high accuracy (94.9%) and DDoS traffic detection rate (0.7 seconds). High detection accuracy also shows a model based on ANN in the research [60]. The authors use five packet header parameters and four publicly accessible datasets for learning, testing, and validation of detection model. The developed model uses the back-propagation learning method and the sigmoidal activation function, which also proved effective in the research [59]. The model detects and distinguishes three classes of DDoS traffic (DNS, UDP and CharGen) and normal traffic with a total accuracy of 95.6%. The research results show the lowest accuracy of UDP DDoS traffic detection of 82.1% due to the matching parameter values of such traffic with the parameter values of normal traffic.

Fuzzy logic in the function of DDoS traffic detection was used in research [61]. The authors of the research suggest a TCP SYN DDoS traffic detection model. The detection accuracy and the number of false positive and negative results are dependent on the defined traffic intensity threshold according to which the probability of DDoS traffic is determined. The authors of research [62] use fuzzy logic for DDoS traffic intensity detection because there is no clearly defined boundary between low and high intensity DDoS traffic. Fuzzy logic was used in combination with a wavelet-based estimation of the Hurst parameter to detect the change of the network traffic self-similarity level. The detected changes of the Hurst parameter value are input into the model based on fuzzy logic, which estimates the intensity of DDoS traffic according to the defined rules. An attempt is made to solve the problem of defining the threshold value of the Hurst parameter above which traffic is considered as DDoS by considering the degree of self-similarity of normal traffic.

3.4 Detection of DDoS traffic generated by IoT devices

Currently, DDoS detection methods are oriented to detecting network traffic anomalies generated by common terminal devices used by end users (humans) such as personal computers, laptops, smartphones,

Table 2 – Studies using machine-learning methods in DDoS traffic detection

Ref.	Method	Performance measuring	DDoS type	Dataset	Number of used features	Accuracy
[54]	CART Decision Tree; Naive Bayes	N/A	Unknown	CAIDA	12	99% Decision Tree; 97% Naive Bayes
[55]	J48, Random Forest; One R; Decision Tree; Bayes Net and Naïve Bayes	N/A	Unknown	NSL- KDD	14	J48 - 99.08%, Random Forest - 99.36%, One R - 91.45%, Decision Tree - 97.00%, Bayes Net - 92.33% and Naïve Bayes - 90.47%
[56]	SVM; K-NN; Naive Bayes; Decision Tree; K-means and Fuzzy c-means clustering	ROC curve and F-measure	Unknown	CAIDA; individually collected traffic	8	Fuzzy C Means - 98.7%, Naive Bayes - 97.2%, SVM - 96.4%, KNN - 96.6%, Decision Tree - 95.6%, K - Means 96.7%
[57]	SVD; majority voting method	TNR, accuracy, precision	Teardrop, Smurf, Pod, Neptune, Land and Back attacks	KDD Cup 99	41	99.4-99.8%

Table 3 – Studies using soft computing methods in DDoS traffic detection

Ref.	Method	Performance measuring	DDoS type	Dataset	Number of used features	Accuracy
[58]	Unsupervised Neural Adaptive Resonance Theory Network;	Exact True Type Detection Rate; True Detection Rate; False Positive Detection Rate; False Negative Detection Rate	UDP Flood, SYN Flood, ICMP Flood, ICMP SMURF	N/A	8	94.9%
[59]	Supervised Artificial Neural Network	Accuracy; Sensitivity; Specificity; Precision	Unknown	Individually generated	N/A	98%
[60]	Artificial Neural Network	Confusion Matrix; ROC curve; cross-entropy error	DNS; CharGen; UDP	CAIDA; UNB ISCX; BootersDS	5	95.6%
[61]	Fuzzy Logic	False Negative; False Positive; True Positive; Precision; Sensitivity	TCP SYN Flooding	N/A	N/A	N/A
[62]	Fuzzy Logic; Discrete wavelet transform; Schwarz information criterion	N/A	Flood	Individually generated	N/A	N/A

tablets, and others. With the development of IoT concept, new and emerging threats need to be considered. The security of IoT devices is a subject of numerous studies. Due to many limitations, IoT devices are a potential target or source of various cyber-attacks. The availability of IC resources in an IoT environment is a key security challenge and can often be hindered by DDoS attacks. In addition of being the target of attack, the devices in IoT environment are ever more frequent sources of DDoS attacks, or generators of illegal DDoS traffic through unprotected IoT devices associated with the botnet network. An example of such botnet through which many DDoS attacks are performed is the Mirai botnet. Mirai has controlled more than 100,000 inadequately protected IoT devices and thus generated illegitimate network traffic (DDoS traffic) to the desired destinations. The problem of DDoS attacks generated by inadequately protected IoT devices is currently an insufficiently researched problem.

IoT devices, unlike common terminal devices, generate specific type of traffic called MTC (Machine Type Communication) traffic. MTC traffic possesses specific characteristics in normal communication process that can be used in creating a normal traffic model. Currently, there are only few studies dealing with the problem of detection of DDoS traffic generated using IoT devices which opens a space for future research. One of the first studies of detection of DDoS traffic generated through the IoT device is [63]. The research is based on the differences between MTC and HTC (Human Type Communication) traffic. The IoT device that generates MTC traffic can receive a fixed number of states and accordingly, MTC traffic is deterministic and structured. Five methods of machine learning (KNN, SVM, Decision trees, Random forest and Artificial neural networks) were used to detect DDoS traffic, with detection accuracy from 91% to 99%. The lack of the present study is only reflected in the three IoT devices used and the 10-minute collection time. Research [64] developed a DDoS traffic detection model generated by using IoT devices. The model is based on the Deep Autoencoding method, and the experiment has been proven to detect 100% DDoS traffic instances. A downside of this research is traffic collected from only nine various IoT devices. There are also studies that try to implement novelty in anomaly detection approaches in IoT environment such as research [65] where device class-based anomaly detection is discussed.

4. DISCUSSION

Despite a large number of studies of DDoS traffic detection possibilities and the use of different approaches, the trends show a continual increase in the attacks that generate this type of traffic (in number and intensity). The analysis of recent research points of disadvantages that may affect the accuracy of DDoS traffic detection as well as the number of false positive and false negative results.

The observed disadvantages in the research of using the statistical and information theory methods in DDoS traffic detection are the determination of the threshold value that represents the difference between normal and DDoS traffic. In most of the analysed research, the threshold value is user-defined and static. Exceptions are research [62] and [10] that use adaptive threshold values, which is crucial because of continuous changes in DDoS traffic characteristics. The challenge of future research is valid selection of the packet header or traffic flow parameters whose value needs to be analysed in the function of DDoS traffic detection [50]. The selection of the relevant packet header parameters and traffic flow characteristics is of great importance due to the reduction in the time required for the analysis and detection of DDoS traffic. It is important that the number of parameters is as small as possible [56]. The above-mentioned research challenge is present in all the analysed approaches of DDoS traffic detection.

DDoS traffic detection approach based on machine-learning methods shows deviations in the detection accuracy and the number of false positive and false negative results depending on the dataset over which methods are applied [54]. This implies the dependence of the detection efficiency on the characteristics of DDoS traffic in different scenarios [55]. This indicates the problem of datasets used in the validation of DDoS traffic detection models. The most commonly used datasets such as KDD-CUP 99, DARPA 2000, CAIDA DDoS 2007, NSL-KDD 2009 and TUIDS 2012 were generated in the laboratory environment or they are outdated and do not reflect the characteristics of today's traffic that are changing under the influence of technological development of new IC devices, concepts and services [44, 66]. Given the above-mentioned efficiency of the developed models in current and future real datasets are questionable. Equally as in the use of statistical methods, the accuracy of machine-learning detection depends on the threshold values of normal and DDoS traffic that are user-defined, which is evident from research [57].

Soft computing approach in detecting DDoS traffic generally shows high accuracy detection with few false positive and false negative results. Some disadvantages are, as with the statistical approach, the selection of packet or traffic flow parameters that will differentiate normal from DDoS traffic as seen in research [59]. In addition, the problem of determining the threshold values between normal and DDoS traffic is also noticeable, which is also observed in statistical and machine-learning approaches [61].

DDoS traffic generated by poorly protected IoT devices represents an emerging problem that needs to be solved. Considering a small number of studies of the mentioned problem there is room for further research. Deterministic characteristics of traffic that IoT devices generates in normal conditions, and discrete number of states of IoT devices can be a starting point in DDoS traffic detection and management of such traffic and malfunctioning IoT devices.

5. CONCLUSION

Distributed denial of service attacks and DDoS traffic generated through such attacks represent a continuous threat to business based on IC technology. The development of new IC concepts such as CC and IoT and applying them in a variety of environments such as autonomous vehicles, smart cities, and critical infrastructure management, significantly increase the potential negative impact of DDoS traffic. Progress and technological development of the IC system is causing development and increasing complexity of DDoS attacks. Despite a large number of studies, the intensity and volume of DDoS-generated traffic are continuously increasing, with the generated traffic being increasingly based on infrastructure layer protocols. Trends point to continuous changes in attacks which also cause changes in DDoS traffic characteristics. Accordingly, the detection based on the known DDoS traffic patterns is not suitable for solving this problem. In order to detect new and unknown instances of DDoS traffic, it is necessary to apply an approach based on network traffic anomalies detection. The current research is most often based on three basic approaches of anomalies detection: using statistical methods, machine-learning methods, and soft computing methods.

Based on the analysis of the previous studies, several disadvantages have been identified that open room for future research. The first identified disadvantage is related to the determination of the threshold value according to which DDoS differs from normal traffic. Because of the specificity of each IC system, but also because of the DDoS attack dynamic, the determination of the threshold value must be adaptive to achieve as few false positive and false negative results as possible. Another identified disadvantage relates to the selection of the packet header or traffic flow parameters whose values are analysed for the DDoS traffic detection. A large number of selected parameters increase detection accuracy and reduce the number of false positive and false negative results but at the same time require more processing resources that affect the possibility of real-time DDoS traffic detection. Therefore, it is necessary to optimize the number of parameters used to minimize the resources required for traffic processing and maximize the detection accuracy. The last identified disadvantage relates to datasets used for testing and validating DDoS traffic detection models and systems. The analysis conducted in this research shows the obsolescence of the used datasets. Given that the characteristics of network traffic change under the influence of technological development, the validation of the DDoS traffic detection model using obsolete datasets may have a significant effect on their use on current datasets. The detection systems must be able to detect unknown DDoS traffic instances due to the continuous development and increasing complexity of this threat type. As a key direction in future research is the detection and management of DDoS traffic and malfunctioning IoT devices as a new and emerging threat.

IVAN CVITIĆ, mag. ing. traff., doktorand[1]
E-mail: ivan.cvitic@fpz.hr
Prof. dr. sc. **DRAGAN PERAKOVIĆ**[1]
E-mail: dragan.perakovic@fpz.hr
Doc. dr. sc. **MARKO PERIŠA**[1]
E-mail: marko.perisa@fpz.hr
Dr. sc. **SINIŠA HUSNJAK**[1]
E-mail: sinisa.husnjak@fpz.hr
[1] Sveučilište u Zagrebu, Fakultet prometnih znanosti
Vukelićeva 4, 10000 Zagreb, Hrvatska

PREGLED PRISTUPA DETEKCIJE PROMETA GENERIRANOG DISTRIBUIRANIM NAPADIMA USKRAĆIVANJA USLUGE

SAŽETAK

Dostupnost informacijsko-komunikacijskih (IK) resursa predstavlja rastući problem uzrokovan porastom broja korisnika, IK usluga i ograničenjima kapaciteta. IK resursi moraju biti dostupni legitimnim korisnicima u traženo vrijeme. Prema tome dostupnost postaje ključni zahtjev u IK okruženjima kao što su pametni gradovi, autonomna vozila ili kritična infrastruktura. U spomenutim i slični okruženjima nedostupnost resursa može rezultirati negativnim posljedicama na fizičku sigurnost ljudi. Distribuirani napadi uskraćivanje usluge (DDoS) i promet koji takvi napadi generiraju predstavljaju rastući problem posljednje desetljeće. Njihov je onemogućiti pristup IK resursima legitimnim korisnicima. Ovim istraživanjem analizirani su trendovi DDoS prometa čime se ukazuje na važnost istraživanja metoda njegove detekcije. Istraživanje pruža i pregled trenutno korištenih pristupa korištenih pri razvoju modela i sustava detekcije. Temeljem analize trenutnih istraživanja identificirani su nedostaci do sada korištenih pristupa što otvara prostora i daje smjer za buduća istraživanja. Uz navedeno, istraživanjem je naglašen i problem DDoS prometa generiranog korištenjem uređaja u okruženju Interneta stvari (IoT) kao nova i rastuća prijetnja koju je potrebno uvažiti u nadolazećim istraživanjima.

KLJUČNE RIJEČI

anomalije mrežnog prometa; mrežno temeljeni napadi; dostupnost usluge; uskraćivanje usluge; detekcija mrežnih anomalija;

REFERENCES

[1] Bhattacharyya DK, Kalita JK. *Network Anomaly Detection: A Machine Learning Perspective.* Boca Raton, USA: CRC Press; 2014.

[2] Husnjak S, Peraković D, Cvitić I. Relevant affect factors of smartphone mobile data traffic. *Promet – Traffic & Transportation*. 2016;28(4): 435-44.

[3] Bidgoli H. *Handbook of Information Security.* 3rd ed. New Jersey: John Wiley & Sons Inc.; 2006.

[4] Tulloch M. *Encyclopedia of Security*. Redmond, USA: Microsoft Press; 2003.

[5] Cvitić I, Peraković D, Periša M, Jerneić B. Availability Protection of IoT Concept Based Telematics System in Transport. In: Mikulski J, editor. *Challenge of Transport Telematics, Katowice, Poland*. Springer International Publishing; 2016. p. 109-21.

[6] Hoque N, Bhuyan MH, Baishya RC, Bhattacharyya DK, Kalita JK. Network attacks: Taxonomy, tools and systems. *Journal of Network and Computer Applications*. 2014;40(1): 307-24.

[7] Bhattacharyya DK, Kalita JK. *DDoS Attacks: Evolution, Detection, Prevention, Reaction and Tolerance.* Boca Raton, USA: CRC Press; 2016.

[8] Chandola V, Banerjee A, Kumar V. Anomaly detection. *ACM Computing Surveys*. 2009;41(3): 1-58.

[9] Hussain A, Heidemann J, Papadopoulos C. A framework for classifying denial of service attacks. In: *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications - SIGCOMM '03, Karlsruhe, Germany.* ACM Press; 2003. p. 99-110.

[10] David J, Thomas C. DDoS Attack Detection Using Fast Entropy Approach on Flow- Based Network Traffic. *Procedia Computer Science*. 2015;50: 30-6.

[11] Somal LK, Virk KS. Classification of Distributed Denial of Service Attacks – Architecture , Taxonomy and Tools. *International Journal of Advanced Research in Computer Science & Technology (IJARCST 2014).* 2014;2(2): 118-22.

[12] Mirkovic J, Reiher P. A taxonomy of DDoS attack and DDoS defense mechanisms. *SIGCOMM Computer Communication Review*. 2004;34(2): 39-53.

[13] Deka RK, Bhattacharyya DK. Self-similarity based DDoS attack detection using Hurst parameter. *Security and Communication Networks*. 2016;9(17): 4468-81. Available from: doi: 10.1002/sec.1639 [Accessed 2017 Jun 16].

[14] Alomari E, Manickam S, Gupta B, Karuppayah S, Alfaris R. Botnet-based Distributed Denial of Service (DDoS) Attacks on Web Servers: Classification and Art. *International Journal of Computer Applications*. 2012;49(7): 24-32.

[15] Yi Xie, Shun-Zheng Yu. Monitoring the Application-Layer DDoS Attacks for Popular Websites. *IEEE/ACM Transactions on Networking*. 2009;17(1): 15-25.

[16] Zhou W, Jia W, Wen S, Xiang Y, Zhou W. Detection and defense of application-layer DDoS attacks in backbone web traffic. *Future Generation Computer Systems*. 2014;38: 36-46.

[17] Peraković D, Periša M, Cvitić I. Analysis of the IoT impact on volume of DDoS attacks. In: Bakmaz M, Bojović N, Marković D, Marković G, Radojičić V. (eds.) *XXXIII Simpozijum o novim tehnologijama u poštanskom i telekomunikacionom saobraćaju – PosTel 2015. Beograd, Serbia;* 2015. p. 295-304.

[18] Prolexic. *Prolexic Quarterly Global DDoS Attack Report (Q2-2013).* Prolexic Technologies, Inc.; 2013.

[19] Prolexic. *Prolexic Quarterly Global DDoS Attack Report (Q3-2013).* Prolexic Technologies, Inc.; 2014.

[20] Prolexic. *Prolexic Quarterly Global DDoS Attack Report (Q4-2013).* Prolexic Technologies, Inc.; 2014.

[21] Prolexic. *Prolexic Attack Report (Q1-2014).* Prolexic Technologies, Inc.; 2014.

[22] Akamai. *Faster Forward to the Latest Global Broadband Trends (Q2-2014).* Akamai Technologies Inc; 2014.

[23] Akamai. *Akamai's State of the Internet - Security (Q3-2014).* Akamai Technologies Inc.; 2014.

[24] Akamai. *Akamai's State of the Internet - Security (Q4-2014).* Akamai Technologies Inc.; 2014.

[25] Akamai. *Akamai's State of the Internet - Security (Q2-2015).* Akamai Technologies Inc.; 2015.

[26] Akamai. *Akamai's State of the Internet - Security (Q3-2015).* Akamai Technologies Inc.; 2015.

[27] Akamai. *Akamai's State of the Internet - Security (Q4-2015).* Akamai Technologies Inc.; 2015.

[28] Akamai. *Akamai's State of the Internet - Security (Q1-2016).* Akamai Technologies Inc.; 2016.

[29] Akamai. *Akamai's State of the Internet - Security (Q2-2016).* Akamai Technologies Inc.; 2016.

[30] Akamai. *Akamai's State of the Internet - Security (Q3-2016).* Akamai Technologies Inc.; 2016.

[31] Akamai. *Akamai's State of the Internet - Security (Q4-2016).* Akamai Technologies Inc.; 2016.

[32] Akamai. *Akamai's State of the Internet - Security (Q1-2017).* Akamai Technologies Inc.; 2017.

[33] Akamai. *Akamai's State of the Internet - Security (Q2-2017).* Akamai Technologies Inc.; 2017.

[34] Akamai. *Akamai's State of the Internet - Security (Q3-2017).* Akamai Technologies Inc.; 2017.

[35] Akamai. *Akamai's State of the Internet - Security (Q4-2017).* Akamai Technologies Inc.; 2017.

[36] Somani G, Gaur MS, Sanghi D, Conti M, Buyya R. DDoS attacks in cloud computing: Issues, taxonomy, and future directions. *Computer Communications*. 2017;107: 30-48.

[37] Cvitić I, Vujić M, Husnjak S. Classification of Security Risks in the IoT Environment. In: Katalinic B. (ed.) *Annals of DAAAM and Proceedings of the International DAAAM Symposium, 21 – 24 September 2015, Zadar, Croatia.* 2016. p. 0731-40.

[38] Tan Z, Jamdagni A, He X, Member S, Nanda P, Member S, et al. Detection of Denial-of-Service Attacks Based on Computer Vision Techniques. *IEEE Transactions on Computers.* 2015;64(9): 1-14.

[39] Bhuyan MH, Bhattacharyya DK, Kalita JK. Network Anomaly Detection: Methods, Systems and Tools. *IEEE Communications Surveys & Tutorials*. 2014;16(1): 303-36.

[40] Zeb K, AsSadhan B, Al-Muhtadi J, Alshebeili S. Anomaly detection using Wavelet-based estimation of LRD in packet and byte count of control traffic. In: *2016 7th International Conference on Information and Communication Systems (ICICS)*; 2016. p. 316-21.

[41] Xiang Y, Li K, Zhou W. Low-rate DDoS attacks detection and traceback by using new information metrics. *IEEE*

Transactions on Information Forensics and Security. 2011;6(2): 426-37.

[42] Zargar ST, Joshi J, Tipper D. A survey of defense mechanisms against distributed denial of service (DDOS) flooding attacks. *IEEE Communications Surveys and Tutorials*. 2013;15(4): 2046-69.

[43] Bhuyan MH, Kashyap HJ, Bhattacharyya DK, Kalita JK. Detecting Distributed Denial of Service Attacks: Methods, Tools and Future Directions. *The Computer Journal.* 2014;57(4): 537-56.

[44] Zeb K, AsSadhan B, Al-Muhtadi J, Alshebeili S, Bashaiwth A. Volume based anomaly detection using LRD analysis of decomposed network traffic. In: *Fourth edition of the International Conference on the Innovative Computing Technology (INTECH 2014).* IEEE; 2014. p. 52-7.

[45] Kaur G, Saxena V, Gupta JP. Detection of TCP targeted high bandwidth attacks using self-similarity. *Journal of King Saud University - Computer and Information Sciences.* 2017; Available from: http://linkinghub.elsevier.com/retrieve/pii/S1319157817300617

[46] Johnson Singh K, Thongam K, De T. Entropy-Based Application Layer DDoS Attack Detection Using Artificial Neural Networks. *Entropy.* 2016;18(10): 350.

[47] Feinstein L, Schnackenberg D, Balupari R, Kindred D. Statistical approaches to DDoS attack detection and response. In: *Proceedings DARPA Information Survivability Conference and Exposition.* IEEE Comput. Soc; 2003. p. 303-14.

[48] Oshima S, Nakashima T, Sueyoshi T. A Statistical DoS/DDoS Detection Method Using the Window of the Constant Packet Number. In: *2009 2nd International Conference on Computer Science and its Applications.* IEEE; 2009. p. 1-6.

[49] Hoque N, Bhattacharyya DK, Kalita JK. FFSc: a novel measure for low-rate and high-rate DDoS attack detection using multivariate data analysis. *Security and Communication Networks.* 2016;9(22): 2032-41.

[50] Hoque N, Bhattacharyya DK, Kalita JK. Denial of Service Attack Detection using Multivariate Correlation Analysis. In: *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies - ICTCS '16.* New York, USA: ACM Press; 2016. p. 1-6.

[51] Arjun H, Maknur SG. A System for Denial-of-Service Attack Detection Based on Multivariate Correlation Analysis. *International Journal of Innovative Research in Computer and Communication Engineering.* 2015;3(4): 447-56.

[52] Sharma N, Mahajan A, Mansotra V. Machine Learning Techniques Used in Detection of DOS Attacks: A Literature Review. 2016;6(3): 100-5.

[53] Hamid Y, Sugumaran M, Journaux L. Machine Learning Techniques for Intrusion Detection. In: *Proceedings of the International Conference on Informatics and Analytics - ICIA-16.* New York, New York, USA: ACM Press; 2016. p. 1-6.

[54] Balkanli E, Alves J, Zincir-Heywood AN. Supervised learning to detect DDoS attacks. In: *2014 IEEE Symposium on Computational Intelligence in Cyber Security (CICS).* IEEE; 2014. p. 1-8.

[55] Osanaiye O, Choo K-KR, Dlodlo M. Analysing Feature Selection and Classification Techniques for DDoS Detection in Cloud. In: *Southern Africa Telecommunication Networks and Applications Conference (SATNAC) 2016. Western Cape, South Africa*; 2016. p. 198-203.

[56] Singh M, Jain SK. Evaluating Machine Learning Algorithms for Detecting DDoS Attacks. In: Wyld DC, Wozniak M, Chaki N, Meghanathan N, Nagamalai D. (eds.) *Communications in Computer and Information Science.* Berlin, Heidelberg: Springer; 2011. p. 608-621.

[57] Jia B, Huang X, Liu R, Ma Y. A DDoS Attack Detection Method Based on Hybrid Heterogeneous Multiclassifier Ensemble Learning. *Journal of Electrical and Computer Engineering.* 2017:1-9.

[58] Jalili R, Imani-Mehr F, Amini M, Shahriari HR. Detection of Distributed Denial of Service Attacks Using Statistical Pre-processor and Unsupervised Neural Networks. In: *International Conference on Information Security Practice and Experience, Singapore*; 2005. p. 192-203.

[59] Saied A, Overill RE, Radzik T. Artificial Neural Networks in the Detection of Known and Unknown DDoS Attacks. In: Corchado MJ, Bajo J, Kozlak J, Pawlewski P, Molina JM, Gaudou B, Julian V, Unland R, Lopes F, Hallenborg K, García P. (eds.) Proof-of-Concept. In: *PAAMS 2014: Highlights of Practical Applications of Heterogeneous Multi-Agent Systems.* Springer; 2014. p. 309-320.

[60] Peraković D, Periša M, Cvitić I, Husnjak S. Model for detection and classification of DDoS traffic based on artificial neural network. *Telfor Journal.* 2017;9(1).

[61] Tuncer T, Tatar Y. Detection SYN Flooding Attacks Using Fuzzy Logic. In*: 2008 International Conference on Information Security and Assurance (ISA 2008)*. IEEE; 2008. p. 321-5.

[62] Xia Z, Lu S, Li J, Tang J. Enhancing DDoS flood attack detection via intelligent fuzzy logic. *Informatica.* 2010;34(4): 497-507.

[63] Doshi R, Apthorpe N, Feamster N. *Machine Learning DDoS Detection for Consumer Internet of Things Devices.* CoRR, abs/180404159. 2018; Available from: http://arxiv.org/abs/1804.04159

[64] Meidan Y, Bohadana M, Mathov Y, Mirsky Y, Breitenbacher D, Shabtai A, et al. N-BaIoT: Network-based Detection of IoT Botnet Attacks Using Deep Autoencoders. *IEEE Pervasive Computing.* 2018;13(9): 1-8.

[65] Cvitić I, Peraković D, Periša M, Botica M. Novel approach for detection of IoT generated DDoS traffic. *Wireless Networks* [Internet]. 2019; Available from: doi:10.1007/s11276-019-02043-1

[66] Bhuyan MH, Bhattacharyya DK, Kalita JK. Towards generating real-life datasets for network intrusion detection. *International Journal of Network Security.* 2015;17(6): 683-701.