APPLICATION OF ARTIFICIAL INTELLIGENCE IN DETECTION OF DDOS ATTACKS

Igor Vuković[1]
Ministry of the Interior of the Republic of Serbia

Brankica Popović, PhD
University of Criminal Investigation and Police Studies, Belgrade, Serbia

Petar Čisar, PhD
University of Criminal Investigation and Police Studies, Belgrade, Serbia

Abstract: Services distributed over the Internet are ranging from entertaining and informative to those whose availability must not be interrupted because it affects the quality of life, but also safety and health. Due to its importance, the global computer network is a desirable target, attacks are continually taking place, and the damage is more than considerable. Among the many types of attacks, one of the most effective, given the relationship between the damage done and the challenge to be prevented, detect and control, are DDoS attacks. This paper discusses the phases, components, categories, and types of DDoS attacks and emphasizes detection approaches. The standout approach and one that can answer the complexity of detecting DDoS attacks is the classification with artificial intelligence techniques. This work shows why artificial intelligence represents the starting point for further research in information security.

Keywords: distributed denial of service, intrusion detection systems, artificial intelligence, classifiers

INTRODUCTION

Vital information of business organizations and authorities, as well as other parts of society, are stored online, which is why the world today is facing an increase in the number of attempts to disrupt the security of such essential data. Business, government, and private data of individuals are stored in the cloud and use the services provided in this way (infrastructure as a service, platform as a service,

1 igor.vukovic@mup.gov.rs

database as a service, etc.). The lives of individuals revolve around data and information that can be found on the Internet, while communication over the Internet plays a crucial role in everyday life (Napanda, Shah & Kurup, 2015). Modern trends include governments of states that use web applications intensively to solve communication and management challenges as efficiently as possible. Still, these applications have numerous entry points that endanger systems and information security (Kadhem, Amagasa & Kitagawa, 2009). The number of reported security breaches of systems containing private, security, business, and other information at the level of government, organizations, and enterprises has reached worrying values (Kadhem et al., 2009), so maintaining confidentiality, accessibility, and integrity has become necessary.

For this reason, intrusions are among the most critical problems for researchers when it comes to business and personal computer networks (Tsai, Hsu, Lin & Lin, 2009). Intrusion can be defined as a set of actions (usually different types of attacks on the system) that compromise security goals such as availability, integrity, confidentiality, and accountability. Intrusion detection systems or IDS help the computer network resist attacks from outside, mainly when conventional systems for prevention of improper or unwanted data input (firewalls) cannot perform the task (Tsai et al., 2009). Attack detection is based on the assumption that an attacker's behavior differs from a legitimate user's behavior that can be quantified (Stalings, 2011).

This paper aims to present the problem and make an overview of protection against a particular type of attack, which regardless of the level of technological development of IDS systems and other defense mechanisms, remains a challenge because it exploits the basic principles of computer networks. These are Distributed Denial of Service (DDoS) attacks that can endanger individuals' normal functioning, the work of companies and state administrations, damage, and even endanger lives.

The first part of this paper presents the problem and challenges of DDoS attacks. The second part analyzes the detection approaches of this specific type of attack. The third part focuses on the application of artificial intelligence in the detection of DDoS attacks.

DISTRIBUTED DENIAL OF SERVICE

DDoS has become a common threat to companies that promote themselves or do business using the Internet and beyond (Srivastava, Gupta, Tyagi, Sharma & Mishra, 2011). The Internet Protocol or IP is not designed to verify that a packet source is authorized to access the service. Still, packets are delivered to a server that represents their destination, and the server must decide whether to accept packets and respond to the client. A situation where the main challenge is to separate a legitimate client request from a malicious one and when it is easier for a source to generate a service request than a server to validate it is when an opportunity for a DDoS attack arises (Peng, Leckie & Ramamohanarao, 2007).

The importance of DDoS attacks as a threat increased when the security vulnerabilities of individual computers on the Internet began to be exploited. This way attacker is forming a network with a large number of machines over which control is established. It is a network of zombie computers "robots" (botnets (robots + network)), which is used for a coordinated attack and enables the generation of traffic whose value exceeds the usual bandwidth and up to several tens of thousands of times (Srivastava et al., 2011; Saied, Overill & Radzik, 2016). DDoS attacks are aimed at consuming critical network service resources like server processor capacity, stack space, and the capacity of the internet connection (Peng et al., 2007). Distributed system services offered by operating systems such as resource sharing

and directory sharing are often abused, allowing the creation of large botnet systems that place DDoS attacks second only to viruses as threats to Internet users (Srivastava et al., 2011).

There are four components involved in a DDoS attack (Srivastava et al., 2011). The first one is an attacker (botmaster) that scans a large number of computers for security vulnerabilities. Machines that become control masters (C & C servers) are the second component, under the direct control of the attacker. With the help of malicious code, which the attacker installs on the control master, infection of the other computers is possible. The next component are computers that become "zombies". They are under the direct control of their master controllers, and indirectly by the attacker who signals a coordinated attack on the target, which is the fourth component of DDoS attack.

Most DDoS designers manipulate TCP, UDP, and ICMP protocols (Saied et al., 2016; Bindraa & Sooda, 2019). Therefore, the dominant type of attack is the SYN flooding, followed by ICMP, UDP, TCP, HTTP, which change in the second, third, and fourth place in quarterly reports for several years[2].

DETECTION OF DDOS ATTACKS

Attack detection can be defined as the detection of an activity that attempts to compromise the confidentiality, integrity, or availability of resources (Napanda et al., 2015). It is necessary to monitor computer systems and networks, analyze their traffic, due to possible attacks from the outside, or malicious use of systems or attacks that come from within a particular organization. According to the logic of detection, we can distinguish three methodologies (Bindraa & Sooda, 2019):

- Detection based on the known specifics of the attack, i.e. on its "signature";

- Detection based on the significant anomaly detection, primarily in the profile of incoming traffic. In order to notice the anomaly, the normal behavior of the system being protected is first introduced. In other words, it is necessary to determine whether the deviation of the established scheme of normal use can be marked as an intrusion;

- Specification-based Detection or Stateful Protocol Analysis, which is similar to anomaly detection, but depends on generic profiles i.e. network protocol models based on international organizations' standards.

Detection of DDoS attacks by the first method is not efficient enough since a large number of compromised computers used during the attack do not have to change the traffic pattern for the attack to be effective. Detection of anomalies can identify an attack if the monitored traffic behavior does not correspond to the normal traffic profile (Peng et al., 2007) and is especially important as it provides protection against new hitherto unknown attacks (mainly zero-days attacks). Problems are that it also results in a higher rate of false positives, it is not available while rebuilding the behavior profile, and it is difficult to achieve timely alarm activation (Liao, Lin, Lin & Tung, 2013). Various artificial intelligence techniques are used for anomalies detection (Napanda et al., 2015).

The accuracy of detection can be improved by applying hybrid or ensemble detection or classifiers, and the ensemble stands out (Kumar, Kumar & Sachdeva, 2010). Hybrid systems combine techniques for different jobs, while an ensemble is made around the same job. Success of the ensemble method depends on various factors like the size of the training sample, choice of the base classifier, the exact way in which the training set was modified, and so on. (Kumar et al., 2010).

2 https://securelist.com/ (Last accessed on 22 May 2019)

According to the place where the data is collected and where the analysis is performed, the following types of detection systems are recognized: Host-based Intrusion Detection Systems or HIDS, Network-based Intrusion Detection Systems or NIDS (Kumar et al., 2010; Napanda et al., 2015) and hybrid systems. HIDS is a system that monitors and analyzes information received from a specific computer or server, such as network traffic, system logs, and applications, system calls, etc. It reports an attack if an anomaly in the behavior of the system is detected (Kumar et al., 2010; Modi et al., 2013). NIDS collects information from the network, primarily the IP headers and transport layer headers of each packet, and uses detection techniques based on signatures and anomalies (Kumar et al., 2010; Modi et al., 2013; Napanda et al., 2015). HIDS systems collect and analyze data using agents, while others use mostly sensors (Liao et al., 2013).

APPLICATION OF ARTIFICIAL INTELLIGENCE

In the development of IDS, the ultimate goal is to achieve the best possible detection accuracy. Learning techniques developed for the needs of artificial intelligence are applied to build better IDS (Napanda et al., 2015). Artificial intelligence is the development of intelligent machines through the branch of computer science. An intelligent machine is a system that observes its environment and takes over activities that increase its chances of success using computer models (Napanda et al., 2015). The advantages of applying artificial intelligence in relation to other conventional techniques are the ability to establish explicit and implicit models that categorize the schemes used in detection, flexibility, and adaptability in relation to precisely defining thresholds and rules, as well as the ability to learn (Kumar et al., 2010).

DDoS attacks require the classification of a large amount of different information, such as categorizing whether it is a regular traffic profile or increased due to an attack, and not for some legitimate reason. The need for a very complex classification imposes the application of artificial intelligence. We distinguish the following classification approaches in anomaly detection based on artificial intelligence: Supervised (parametric and nonparametric methods (k-nearest neighbor)), Unsupervised (clustering, outlier mining and association mining) and Probabilistic learning (hidden Markov model, Bayesian networks, naïve Bayes, Gaussian mixture model and expectation-maximization algorithm), Soft computing (for instance, artificial neural networks, genetic algorithms, rough sets and fuzzy logic), Knowledge based anomaly detection (rule based and expert system based, ontology and logic based) and Combination learners (ensemble based, fusion based and hybrid).

A classifier that applies the k-nearest neighbor (*k*-NN) method calculates the approximate distance between different points on the input vectors, and then assigns to the unmarked point the class of its K nearest neighbors (Tsai et al., 2009). K is an important parameter that determines how many nearest neighbors are required in order for something to be classified, but to be noted the higher the value of the parameter, the longer the classification time. An example of the application of this technique uses the frequency of system calls to detect malicious online activities (Napanda et al., 2015). The research (Polat, Polat & Çetin, 2020) has shown that in performance, i.e. classification accuracy, the k-NN is ahead of other machine learning techniques in DDoS attack detection.

Figure 1 shows an example of classification with K closest adjacent where x and y are the values monitored by the attack detection system, then the arrangement of vectors on the graph that are already classified as DDoS attacks, these are dark spots, and those that belong to the regular state of the system, represented by white dots. The new sample positioned on the graph, depending on the recorded values, will be classified by determining 5, which represents the previously selected K value of the closest

classified adjacent vectors. The result is 3 dark and 2 light dots, so the decision of the system is that it is a DDoS attack. The obvious problem is determining the K value. The only known thing is that it should be odd, to avoid the situation where classification is not possible because of the same number of neighbors from different classes. Similarly, K should not be divisible by the number of classes. Furthermore, the complexity of the search for the nearest neighbors with a large number of already classified vectors for each new sample creates the need for significant resources.

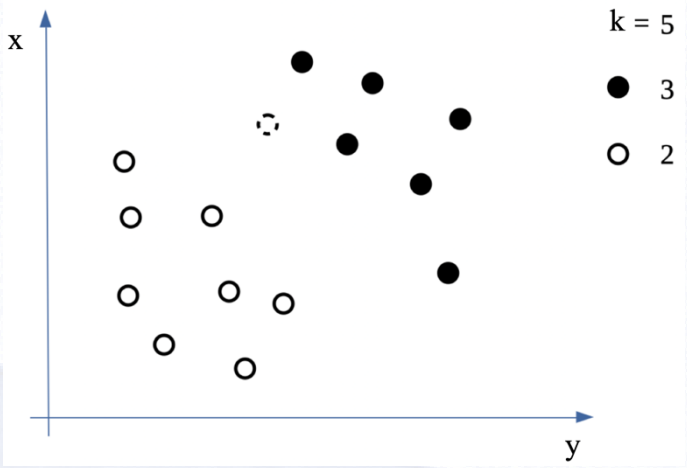

Figure 1 K-nearest neighbor classification

Detection systems based on Artificial Neural Networks (ANN), such as multilayer feedforward, multilayer perceptrons, and backward propagation neural networks, aim to generalize data based on incomplete information and to allow data to be classified as normal or harmful (Modi et al.,2013; Napanda et al., 2015). The neural network consists of artificial neurons, i.e. units that perform processing. The connection between them with a specific weight value depends on how one neuron will affect another. Different DDoS detection schemes have been proposed, mostly using a basic ANN with the Backpropagation algorithm and one hidden layer (Liang & Znati, 2019). Convolutional Neural Networks (CNN) have grown in popularity in recent times leading to significant computer vision innovations and Natural Language Processing. Most of all, this machine learning technique success in specific scenarios such as malware detection, code analysis, network traffic analysis, and intrusion detection motivate researchers to implement CNN in DDoS detection (Doriguzzi-Corin, Millar, Scott-Hayward, Martínez-del-Rincón & Siracusa, 2020).

The genetic algorithm is a class of computer models based on the concept of natural selection and evolution. Chromosome-analog data structure that evolves using selection (striving for an optimal solution from one generation to another), recombination, and mutation (mutation provides population diversity, while the other way is a large population, but it slows down the calculation) are used to convert the problem into a particular domain related to the model (Napanda et al., 2015). This type of algorithm can be used to define simple network access rules which will prevent the passage of known malicious attacks. The disadvantage of the algorithm is that it significantly consumes resources (Kumar et al., 2010).

The Bayesian network performs classification by answering the question of the probability that a particular type of attack can occur again based on previously recorded events related to that type of attack (Napanda et al., 2015). The technique connects previous knowledge and available data about the relationships between the variables based on which predictions are made. Still, the disadvantage is that the results are compared with statistical techniques, which requires significant resources for additional calculations to be performed (Kumar et al., 2010).

Figure 2 shows a segment of the Bayesian network, extremely simplified so that the focus remains on the essence. The system looks at both parameters and points out their probabilities of occurrence, and their existence affects the answer to whether it is a DDoS attack. More precisely, if both parameters exist, the possibility that it is an attack is 90 percent. At the same time, the probability is only 10 percent with the existence of both parameters that it is not an attack.

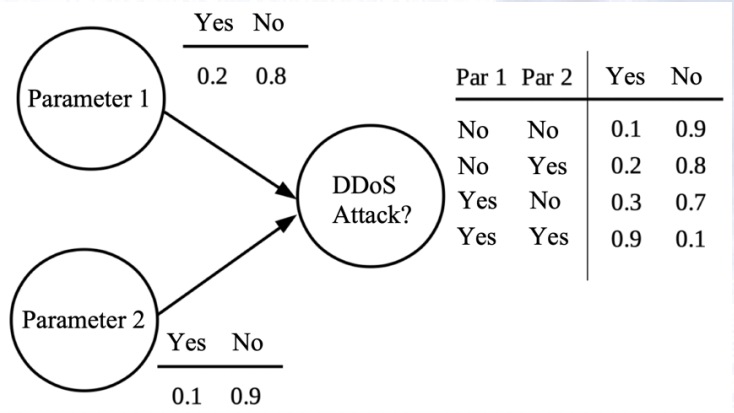

Figure 2 Bayesian network classification

The Decision Tree (DT) is a potent and popular tool for classification and prediction (Kumar et al., 2010). It is an algorithm for supervised learning that consists of a structure, a particular type of graph, in which decision-making at one point affects the decision that leads to the conclusion (Napanda et al., 2015). Each node has an attribute which is the most informative among the characteristics not yet considered on the way from the root. Each branch is marked by the value of the node attribute from which it came out. In contrast, nodes at leaf position are marked with class or category names. The path from the root node to the branches of the tree represents the rules of classification. These rules are defined according to anomalies recorded by the system being protected or encountered earlier (Napanda et al., 2015). DT model provides well-defined rules to accurately distinguish attack traffic from legitimate traffic (Liang & Znati, 2019).

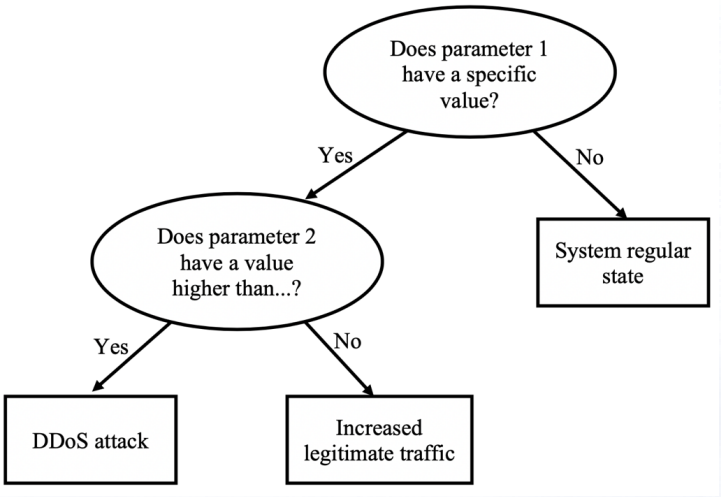

Figure 3 Decision tree classification

Figure 3 presents a decision tree application that classifies the state of a system based on arbitrary parameters measured by an intrusion detection system. The figure shows that one of the nodes at the

leaf position represents a class marked as a "DDoS attack." If you follow the path from the root to the subject node, you can see the rule based on which the classification was performed by monitoring two parameters by the attack detection system.

SVM (Support Vector Machine) attempts to solve an optimization problem that consists of finding decisions boundary in the feature space that separates data from different classes (Liang & Znati, 2019). Support vectors are subsets of training data that are used to determine the boundary between two classes. When SVM cannot divide two classes, the problem is solved by mapping the input data into a higher spatial dimension using one of the so-called kernel functions (for example, radial base function). In a higher dimension, it is possible to create a hyperplane that allows linear separation. There is a curved surface in the lower dimensional space, which is why the kernel function plays an important role in SVM. The ability to learn SVM is independent of the spatial dimension, so they generalize well when there are a large number of features (Kumar et al., 2010). SVM provides the parameter the so-called penalty factor, determined by the user and based on decision between a situation with more misclassified samples, but more stable decision-making, or from the other side a narrower space, but a smaller number of misclassified samples (Tsai et al., 2009).

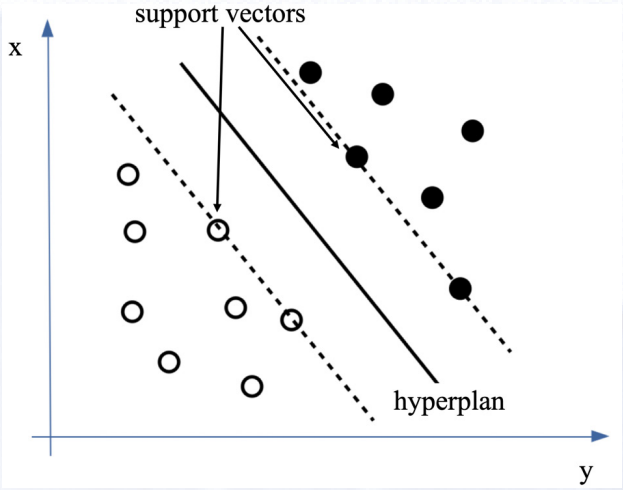

Figure 4 Support vector machines classification

Figure 4 shows a graphical representation of the SVM classification, where the x and y axes represent two arbitrary parameters that are monitored when detecting DDoS attacks. In contrast, the data used for SVM training are shown by light dots if classified as regular system behavior, and dark if classified in the class of attack carried out. The dashed lines represent the boundaries of the decision, while the points through which they pass are the support vectors. A solid line represents a hyperplane that divides two classes. Based on the parameters values, the newly recorded data is positioned on the graph and classified to which side the hyperplane is on telling the detection system whether it is normal traffic or an attack.

The efficiency and accuracy of any classifier depend primarily on the choice of parameters that the detection system monitors. The parameters used are the source and destination IP addresses, source ports and destination ports, connection length, a number of bytes exchanged between pages, connection status, TCP control bits. Besides the number of packets sent from any IP address, the average size of IP packets sent from each IP address and the standard deviation of the packet size sent from each IP address is monitored (Kong, Yang, Sun, Li & Shi, 2017).

Classifiers can be combined to achieve higher precision.

EXAMPLE OF IMPLEMENTATION OF IDS SYSTEM FOR DETECTION OF DDOS ATTACKS

Figure 5 shows the detailed architecture of the DDoS attack detection system presented in (Han, Bi, Liu & Jia, 2017).

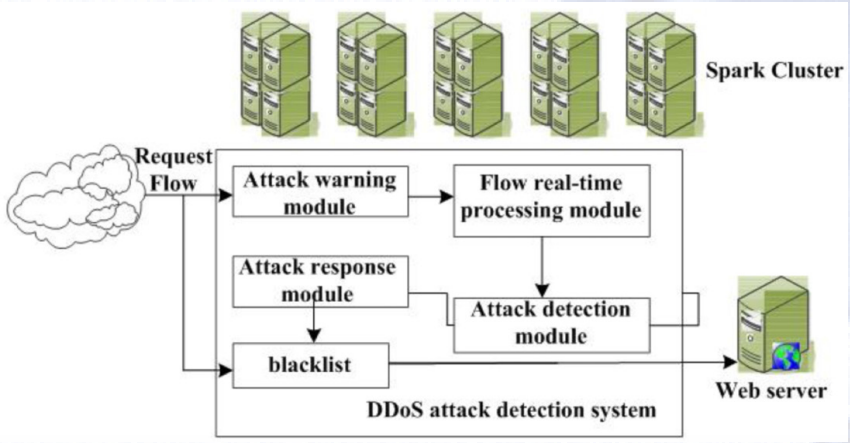

Figure 5 An example of an attack prevention system architecture

An early warning of a DDoS attack based on abnormal changes in the entropy of information about the source and destination IP data was implemented within the attack detection system shown in Figure 5 (Han et al., 2017). Entropy is used as a measure of the unpredictability or uncertainty of an order, it is highest for entirely random data from an information source, and lowest when the information source provides predictable data (Petkovic, Basicevic, Kukolj & Popovic, 2018; Wang, Luo & Zhong, 2019). The utility of entropy in detecting DDoS attacks is that the entropy of regular network traffic varies within a narrow range, while many anomalies caused by DDoS attacks change the distribution of addresses and ports, as well as other traffic characteristics (Petkovic et al., 2018). After launching the attack, the amount of IP addresses of the source will increase drastically as the distribution will be more scattered. On the other hand, the distribution of destination IP addresses and ports will be more concentrated (Han et al., 2017).

The method used in example is the K-Means method, which in addition to a similar name, differs from the K-nearest neighbor classification method (Han et al., 2017). The clustering process begins with a random isoform of K objects representing the center or average value of the cluster. Then, based on the minimum distance from the selected cluster centers, the other sample elements are assigned to each cluster. Determining the new center of gravity is based on the average value of each cluster instances. Sorting begins again based on the distance from the newly calculated center of gravity of the cluster. The process stops when further iterations do not lead to significant changes, i.e. the algorithm converges. During a DDoS attack, the K-mean method requires processing a large amount of data mixed with the flow of data related to the attack making it difficult to determine the initial centers. Therefore, dynamic sampling of K-mean clusters is applied to improve the algorithm and meet the requirements of the DDoS attack detection system (Han et al., 2017).

The system in Figure 5 consists of Attack warning module, Flow real-time processing module, Attack detection module, Attack response module and blacklist. Attack warning module uses an algorithm based on the entropy of data flow information. Processing module collects the data flow characteristics from the warning module and forwards them to the attack detection module. Detection module

uses the data obtained from the processing module for the clustering process based on when the DDoS attack is recognized.

CONCLUSION

Although the DDoS attack has been a problem for the security of information systems for decades, it remains a challenge to detect such attacks. Researchers in their work are intensively looking for solutions or combining existing ones, but the application of artificial intelligence stands out among other branches of computer science. There are many approaches and techniques used to build IDS to detect DDoS attacks. Each approach and technique has its advantages and disadvantages, which is why they cannot be the final solution separately. Signature-based approaches cannot identify completely new attacks, rule-based methods can detect new attacks, but creating and updating a knowledge base is a demanding job, while heuristic methods, such as applying artificial intelligence and classifiers, are not suitable to work in real-time because they require very complex calculations.

DDoS attacks are a critical problem on the Internet, and there are no signs that this will change any time soon. More attackers are continually looking for new targets and new ways to deplete computer networks, and attacks become more complex and sophisticated.

To protect the systems from attacks, firstly they must be detected in a timely and precise manner. However, even after successful detection, the problems do not diminish. One of the responses to a DDoS attack after detection is blocking the source, but the use of botnets makes it difficult to identify legitimate users when blocking. Filtering malicious traffic as another response is also a problem because it is challenging to distinguish between legitimate and malicious packets. Some solutions offered commercially are mostly systems whose size has the role of absorbing the attack, and not to detect it nor stop it[3].

The fact is that the number and scope of DDoS attacks are increasing and that the losses per hour of DDoS attacks can be estimated at tens of thousands of dollars (Han et al., 2017), which is why their detection is necessary. Finding new and improving existing detection methods and selecting appropriate parameters that can be monitored and based, on which correct and efficient decisions can be made, is a broad area for new scientific research. At the same time, the importance of the problem does not diminish over time.

REFERENCES

1. Bindraa, N. & Sooda, M. (2019). Detecting DDoS Attacks Using Machine Learning Techniques and Contemporary Intrusion Detection Dataset. Automatic Control and Computer Sciences, 53 (5), 419-428.
2. Doriguzzi-Corin, R., Millar, S., Scott-Hayward, S., Martínez-del-Rincón, J. & Siracusa, D. (2020). Lucid: A Practical, Lightweight Deep Learning Solution for DDoS Attack Detection. *IEEE Transactions on Network and Service Management*, 17 (2), 876-889.
3. Han, D., Bi, K., Liu, H. & Jia J. (2017). A DDoS Attack Detection System Based on Spark Framework. *Computer Science and Information Systems*, 14 (3), 769-788.

3 https://www.cloudflare.com/ddos/ (Last accessed on 25 September 2019)

4. Kadhem, H., Amagasa, T. & Kitagawa, H. (2009). A Novel Framework for Database Security Based on Mixed Cryptography. *Fourth International Conference on Internet and Web Applications and Services*, Venice/Mestre, Italy, 163-170.
5. Kong, B., Yang, K., Sun, D., Li, M. & Shi, Z. (2017). Distinguishing Flooding Distributed Denial of Service from Flash Crowds Using Four Data Mining Approaches. *Computer Science and Information Systems*, 14 (3), 839-856.
6. Kumar, G., Kumar, K. & Sachdeva, M. (2010). The use of artificial intelligence-based techniques for intrusion detection: a review. *Artificial Intelligence Review*, 34 (4), 369-387.
7. Liao, H. J., Lin, C. H. R., Lin, Y. C. & Tung, K. Y. (2013). Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications*, 36, 16-24.
8. Liang, X. & Znati, T. (2019). On the performance of intelligent techniques for intensive and stealthy DDos detection. *Computer Networks*, 164, 106906.
9. Modi, C., Patel, D., Patel, H., Borisaniya, B., Patel, A. & Rajarajan, M. (2013). A survey of intrusion detection techniques in Cloud. *Journal of Network and Computer Applications*, 36 (1), 42-57.
10. Napanda, K., Shah, H. & Kurup, L. (2015). Artificial Intelligence Techniques for Network Intrusion Detection. *International Journal of Engineering Research & Technology*, 4 (11), 357-361.
11. Peng, T., Leckie, C. & Ramamohanarao, K. (2007). Survey of network-based defense mechanisms countering the DoS and DDoS problems. *ACM Computing Surveys*, 39 (1), 1-42.
12. Petkovic, M., Basicevic, I., Kukolj D. & Popovic, M. (2018). Evaluation of Takagi-Sugeno-Kang fuzzy method in entropy-based detection of DDoS attacks. *Computer Science and Information Systems*, 15 (1), 139-162.
13. Polat, H., Polat, O. & Çetin, A. (2020). Detecting DDoS Attacks in Software-Defined Networks through Feature Selection Methods and Machine Learning Models. *Sustainability*, 12, 1035.
14. Saied, A., Overill, R. E. & Radzik, T. (2016). Detection of known and unknown DDoS attacks using Artificial Neural Networks. *Neurocomputing*, 172, 385–393.
15. Srivastava, A., Gupta, B. B., Tyagi, A., Sharma, A. & Mishra, A. (2011). A Recent Survey on *DDoS* Attacks and Defense Mechanisms. *International Conference on Parallel Distributed Computing Technologies and Applications, Advances in Parallel Distributed Computing*, 570-580.
16. Stalings, W. (2011). Cryptography and network security principles and practices. Prentice Hall, New York City, USA.
17. Tsai, C. F., Hsu, Y. F., Lin, C. Y. & Lin, W. Y. (2009). Intrusion detection by machine learning: A review. *Expert Systems with Applications*, 36, 11994-12000.
18. Wang, M., Luo, Y. & Zhong, H. (2019). DDoS detection and defense mechanism based on cognitive-inspired computing in SDN. *Future Generation Computer Systems* 97, 275–283.