*THINK INDIA JOURNAL*

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

# A Survey on Machine Learning Techniques Used For Detection of DDOS Attacks

**Monisha H M**
Mtech Student
Dept of Information Science & Engineering
BMS collegeof Engineering
Bangalore , India
monisha.scn17@bmsce.ac.in

**Dr R Ashok Kumar**
Associate Professor
Dept of Information Science& Engineering
BMScollege of Engineering
Bangalore, India
ashokkumar.ise@bmsce.ac.i

**ABSTRACT**

Distributed Denial of service (DDOS) attack is one of the most annihilating attack which adheres the functionality of basic administrations given by the different firms in the internet. These attacks have turned out to be progressively complex and keep on expanding in number step by step, in this way making it hard to distinguish and to counter such attacks. In this way there is a need of keen intrusion detection system (IDS) to recognize and arrange any irregular conduct of the system traffic. In a DDOS attack, the intruder takes the advantage of the known or unknown flaws and vulnerabilities by using zombies (innocent compromised PCs) and sends a large number of packets from these zombies to the server. This takes the network bandwidth of the user and halts the services which the user was using. Hence the detection of DDOS attack is very critical .In this paper, a survey is made on various machine learning techniques used to detect the DDOS attack.

**Keywords**—DDOS attack, Cyber Security, Decision tree, Naïve Bayes, Support Vector Machine, Neural Network and Fuzzy Logic, Machine Learning(ML).

## I. INTRODUCTION

Communication networks and information systems plays a prominent role in social and economic development and in day to day of our lives. Rapid revolution in World Wide Web and communication and telecommunication networks, information systems are now sensitive to many types of cyber-attacks. The attacks happening on the communication networks and the information systems are becoming an issue of concern for security. In general, Cyber security attacks are rapidly increasing. Hence, in order to protect the network from all types of attacks an intrusion detection and prevention system plays a prominent role for detection of security violation.

An intrusion attempt is an illegal attempt or threat to (i) to gain the information, (ii) to change or alter the content, (iii) make the system unusable. For example, Worms and viruses exploit other nodes in the network by disturbing the normal functionality through corrupting the software by a bug (virus) or corrupting the network by a bug. Thus disturbing the network traffic,

A. Phishing is the fraudulent attempt to gather user's information such as usernames, passwords and credit cards, debit card details pretending as a reliable entity in an electronic communications.

B. Denial of service (DoS) is an attack where the attacker intends to make the machine or the network resource unavailable for its legitimate users,

A DoS attack which involves more than one machine in a distributed fashion to target a victim is called Distributed Denial of Service (DDoS) attack. This paper focuses on using ML techniques to detect DDoS attack.A Signature is left behind every cyber-attack.A set of rules is called as a signature that an IDS

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

used to identify malicious intrusive activity in the network such as DDoS attacks. Some of the methods to identify signature are:

- Attempt made to connect from a reserved IP address. By verifying the source address fields in an IP header connection could be easily identified.
- Illegal TCP flag combination in a packet. Looking at the flag set in a TCP header, against the known positive or negative flag mixes could be discovered.
- Virus attack in an email server. This can be found out by checking the subjects of every mail to the subject having virus-loaded email.
- DOS attack on a POP3 server caused by sending the same command many number of times. One signature will be kept to keep track the number of counts the same command has been sent and to notify if the same command exceeds after some value.
- File access attack on the FTP server by providing directories and file commands without even signing into it. A state tracking system can be expanded to track the FTP traffic for correct sign-in. This could notify if some of the commands were provided before the user gets authenticated.

Several genuine intrusion detection system depends on the human investigation to recognize among the intrusive and non-intrusive attacks. Human involvement is needed to generate, debug and deploy the appendage on the examined datasets so it might take long time or may be many hours to find and create a new signature for an attack.

Due to the serious impact of compromised system on business and personal networks, Intrusion system has become a main research field for researchers, cyber-security admins and network admins. Intrusion detection system (IDS) can be used to discover different types of malicious network activities and malicious network intrusions, whereas the standard firewall cannot do these tasks. In general Intrusion detection systems can be classified as Signature based IDSs and Anomaly based IDSs.

A. Signature Based Detection

Signature based methods works on principle of previous observed signatures and tally the signatures that is stored on to the database. This database consist of list of signatures that are associated with the previous attacks. Signature of IDs that have some detailed information in computation and preprocessing so that it does not check for every action or network traffic on the environment is being monitored. Signature based approach is easy to utilize the information as it does not need to learn about the environment instead the known signatures are previously stored in the database. Signature based method is a valid method against the known attacks, but it can't be recognized as new attacks until it is activated with new signatures. Signature based IDS are anything but difficult to avoid as they are based on known attacks and are dependent on new signature which has to be used before that they can be identified as new attacks. Signature based methods are easier to modify and improve as its performance is based on the signature or rules deployed.

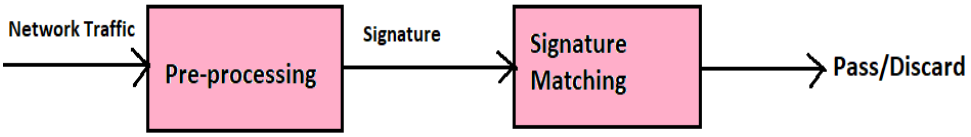


**Figure 1: Signature based detection**

Figure 1 shows the architecture of the signature based methodology. In this architecture, the network traffic is being pre-processed to gain the selected features and to discover the signature. The signature is then compared with action signature and it is checked in the signature database. If the signature matches, an alarm is raised and if there is no signature match, it does nothing. This type of Intrusion Detection Systems are used by many business by which the known attacks can be distinguished financially and with low false positive failures.

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

B. Anomaly Based Detection

Anomaly based detection is also called as "Behavior based detection" ,as the models are based on behavior of the network, most computer systems raises an alert messages when there is a normal detection in the behavior. This method detects the unwanted traffic and is best suited in finding a study on research towards network hardware.The general architecture includes an anomaly based methods as shown in Figure 2.This architecture includes the pre-processing function to collect the data and build the pattern to the connection if it differs from the normal behavior then an alert message is raised.

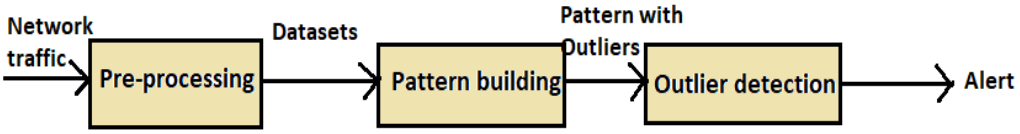


**Figure 2: Anomaly Based detection**

This paper focuses on survey of different machine learning technique to detect the DDOS attack. The rest of the paper is as follows: Section 2 gives the summary about the literature survey of the machine learning techniques used in detection of DDOS attacks. Section 3 gives the comparison about various ML techniques to detect the DDOS attacks along with its advantages and disadvantages. In Section 4, conclusion of the survey of the machine learning techniques are made.

## II. LITERATURE SURVEY

This section gives the literature survey of ML techniques used in the detection of DDOS attacks such as Decision tree, Naïve Bayes, Artificial Neural Networks, Support Vector Machine and Fuzzy Logic are made and their related works have been covered.

1. Decision Tree:

Many research were made on decision tree predictive model to detect the DDOS attack.Yi Chi Wu et al.[6], using a decision tree technique designed a DDoS-Detection system, the system traces back the attacker's location after an attack is detected using a traffic-flow pattern-matching technique. For detection of DDOS attacks a C.45 classifier is used. Author in [8] finds out a way in which the DDOS attack could be efficiently detected. Many ML techniques takes more time to detect the attack or produces less accuracy. In [8] C4.5 algorithm is used and it shows less accuracy and takes more time to generate the decision tree while C5.0 algorithm is proven to be efficient as it takes less time and memory compared to the C4.5.Bujlow, Khafhali and Saadi,[9] focuses on the classification of the network traffic, the results of the C5.0 algorithm was better than that of the C4.5 algorithm.Another work is carried out using ID3,C4.5 and C5.0 which generates an improved decision tree with reduced error pruning and feature selection[10].Based on the results obtained C5.0 has performed better with accuracy and usage of memory.

2. Support Vector Machines(SVM):

For Machine learning tasks the Support Vector Machines are the most popular and common method. Vipin Das et al.[12]in 2010 using RTS and SVM conducted an experiment to detect DDOS attacks.Pre-processing of the packet data that was captured initially from the network was done by the RST. SVM model is fed with the feature set selected by the RST to learn and test respectively. RST and SMV could reduce false positives rates when results are compared with PCA(Principal component analysis),thus increasing the accuracy.

3. Naive Bayes:

This ML technique belongs to a simple probabilistic classifier [13].Carl Livadas [16] used Ml techniques to recognize the command and to control traffic of IRC based botnets .The author identifies IRC and non-IRC traffic by comparing the performance of J48,Naïve Bayes and Bayesian network. In paper[16],author finds the features that gives better accuracy. This classifier achieves both low false negative (2.49%) and false positive (15.04%) for real-life IRC/non-IRC flows and low false negative (7.89%) rates for botnets tested bed IRC flows proves Naïve Bayes to be an efficient classifier.Jasreena Kaur bains et al in [14] for detection rate of attacks proposed a hierarchical layered approach.

ISSN: 0971-1260
Vol-22-Issue-16-August-2019

4. Neural Networks

Neural Networks use processing elements to convert a set of input to a set of outputs whose functionality is like the biological nervous systems like a human brain .Gavrilis et al [15] used "RBF-NN detector" used 9 packet parameter and the respective frequencies of these parameters are calculated. Depending on the frequencies which were estimated, RBF-NN classifies traffic as normal class or wether it is an attack class**.** Distributed Time Delay Neural Network (DTDNN)[21] has high chances of detecting attacks with better accuracy. Classification of data with fast conversion rates and also with high speedis done by DTDNN.

5. Fuzzy Logic

Fuzzy techniques are mainly used by Anomaly detection. A detection and prediction mechanism was proposed by the author Vladimir[17] against DDoS attacks using Fuzzy logic in IEEE 802.15.4. Fuzzy based detection and prediction system (FBDPS) helped in detecting DDODS attack by comparing the energy consumption of sensor nodes.The nodes is identified as a malicious attacker by seeing the abnormal energy consumption.

## III. COMPARISION OF MACHINE LWARNING TECHNIQUES

All though machine learning based algorithms are used to detect intrusion as mentioned above to achieve high detection rate, they have their own advantages and disadvantages.

1. Fuzzy logic:

Ahmad Rizaain Yusof et al has proposed in [7] that Fuzzy c-means clustering shows better classification among other classifiers .Fuzzy c-means happened to be fast compared to other Machine learning algorithms. Fuzzy logic has reasoning which approximate rather than accurate. Suresh [4] implemented and evaluated Fuzzy c-means clustering on DDOS attacks with ML techniques and got better classification among other techniques. Reduced, relevant rule subset identification is a difficult task. Manjula and Anitha in [5], has evaluated the ML techniques to detect DDOS attack using CAIDA dataset which is based on chi-square and information gaining ranking for the selected features. The results shows better classification in Fuzzy-c means which proves to be faster than other algorithms. Fuzzy logic is effective against the port scans and probes.

2. Neural Networks

Jie-Hao et al.[2] used Artificial neural networks to detect DDOS attacks and comparative study made between ANN, decision tree, antropy and Bayesian. It has an ability to generalize from limited and incomplete data. Neural networks takes more time for training process and not suitable for real-time detection.

3. Support Vector Machine(SVM)

Li et al.[1] **,**proposedan intelligent module for network intrusion prevention system by combining SNORT and firewall which is configurable. False alarm rate is reduced by using SVM classifier with SNORT, thus improving the accuracy of Intrusion Prevention System. SVM are simple to interpret results and it is efficient. Only disadvantage is that SVM handles only binary classification. It fails to give additional information of detected attack type for binary classifiers.

4. Naïve Bayes

Mouhammed Alkasassbeh et al[13] collected new dataset that consists of DDOS attacks in different network layers. He used Multilayer perceptron (MLP), Naive Bayes and Random forest are three techniques used to detect DDoS attacks. The highest accuracy was shown by MLP (98.63%) compared to other techniques. Advantage of Naïve Bayes is implementation is easy. On larger data points Naïve Bayes could be employed but it increases the complexity. Disadvantage is need of probability data and assumptions of conditional independence.

5. Decision tree

Ismanto and Wardoyo [19]show that C4.5 decision tree algorithm is more stable than k-Nearest Neighbour algorithm. Experiment is conducted on tree intrusion detection which evaluates Multi-layer Perceptron(MLP), C4.5 and SVM classifiers [23] in which C4.5 shows that better in detection(99.05%) and minimum training time is C4.5.In [20] the authors have evaluated the C4.5, Naïve Bayes and C5.0 in detecting the DDOS attacks. The accuracy is high for C5.0 compared to other two algorithms. Many research has been made by people to infer that C4.5 has been performing well with higher accuracy but C5.0 has started performing even better than C4.5 [20] .Advantage of Decision tree is that high accuracy is achieved with increase in the size of the datasets .Disadvantage of Decision tree is that if the numeric dataset are complex, then decision tree algorithm are unstable.

## IV. CONCLUSION

After the thorough survey it is concluded that network attacks are harmful and appropriate intrusion detection systems has to be implied. As mentioned in the previous section, each machine learning techniques have performed well with one or the other feature that could be counted in accordance with the detection of the intrusion in a system. Some Machine learning based detection algorithm has better accuracy with increase in datasets while some algorithms has less false alarm rate and some algorithm will be good at port scans and probes. Thus it is concluded that the machine learning algorithms has its advantages and disadvantages in detecting the DDOS attacks. Appropriate algorithms should be selected depending on weather the DDOS attack is happening at the Ports or network or application layer and also the type of datasets available.

## REFERENCES

[1] H. Li and D. Liu, "Research on intelligent intrusion prevention system based on snort," in *International Conference on Computer, Mechatronics, Control and Electronic Engineering (CMCE)*, vol. 1. IEEE, 2010, pp.251–253
[2] J. Li, Y. Liu, and L. Gu, "Ddos attack detection based on neural network," in *2nd International Symposium on Aware Computing (ISAC)*,. IEEE, 2010, pp. 196–199
[3] Quinlan, J. R. (2014). *C4. 5: programs for machine learning*. Elsevier.
[4] Livadas, C., Walsh, R., Lapsley, D.E., & Strayer, W.T. (2006). Usilng Machine Learning Technliques to Identify Botnet Traffic. *Proceedings. 2006 31st IEEE Conference on Local Computer Networks*, 967-974.
[5] Suresh M., Anitha R. (2011) Evaluating Machine Learning Algorithms for Detecting DDoS Attacks. In: Wyld D.C., Wozniak M., Chaki N., Meghanathan N., Nagamalai D. (eds) Advances in Network Security and Applications. CNSA 2011. Communications in Computer and Information Science, vol 196. Springer, Berlin, Heidelberg
[6] Yi-Chi Wu, Huei-Ru Tseng, Wuu Yang and Rong-Hong Jan" DDoS detection and traceback with decision tree and grey relational analysis" Int. J. Ad Hoc and Ubiquitous Computing, Vol. 7, No. 2, 2011
[7] Yusof A.R., Udzir N.I., Selamat A. (2016) An Evaluation on KNN-SVM Algorithm for Detection and Prediction of DDoS Attack. In: Fujita H., Ali M., Selamat A., Sasaki J., Kurematsu M. (eds) Trends in Applied Knowledge-Based Systems and Data Science. IEA/AIE 2016. Lecture Notes in Computer Science, vol 9799. Springer, Cham
[8] Zekri, M., El Kafhali, S., Aboutabit, N., & Saadi, Y. (2017, October). DDoS attack detection using machine learning techniques in cloud computing environments. In *Cloud Computing Technologies and Applications (CloudTech), 2017 3rd International Conference of* (pp. 1-7). IEEE.
[9] Bujlow, T., Riaz, T., & Pedersen, J. M. (2012, January). A method for classification of network traffic based on C5. 0 Machine Learning Algorithm. In *Computing, Networking and Communications (ICNC), 2012 International Conference on*(pp. 237-241). IEEE.
[10] Rutvija Pandya and Jayati Pandya. Article: C5.0 Algorithm to Improved Decision Tree with Feature Selection and Reduced Error Pruning. *International Journal of Computer Applications* 117(16):18-21, May 2015.
[11] R. Revathy, and R. Lawrance. (2017). "Comparative Analysis of C4.5 and C5.0 Algorithms on Crop Pest Data," *International Journal of Innovative Research in Computer and Communication Engineering(IJIRCCE)* Vol. 5, Special issue 1, March 2017
[12] Vipin Das , Vijaya Pathak, Sattvik Sharma, Sreevathsan, MVVNS.Srikanth,Gireesh Kumar T," NETWORK INTRUSION DETECTION SYSTEM BASED ON MACHINE LEARNING ALGORITHMS", International Journal of Computer Science & Information Technology (IJCSIT), Vol 2, No 6, December 2010
[13] Mouhammd Alkasassbeh IT Department ,Ghazi Al-Naymat ,Ahmad B.A Hassanat ,Mohammad Almseidin," Detecting Distributed Denial of Service Attacks Using Data Mining Techniques," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 1, 2016
[14] Jasreena Kaur Bains ,Kiran Kumar Kaki ,Kapil Sharma," Intrusion Detection System with Multi Layer using Bayesian Networks", International Journal of Computer Applications (0975 – 8887) Volume 67– No.5, April 2013
[15] Gavrilis, Dimitris & Dermatas, Evangelos. (2005). Real-time detection of distributed denial-of-service attacks using RBF networks and statistical features.Computer Networks.48.235-245.10.1016/j.comnet.2004.08.014.
[16] Irfan Sofi, Amit Mahajan, Vibhakar Mansotra ,Machine Learning Techniques used for the Detection and Analysis of Modern Types of DDoS Attacks (IRJET), Vol.04,Issue,06,June2017
[17] C Balsrengadurali and Dr.S Saraswathi," Fuzzy Based Detection and Prediction of DDoS Attacks in IEEE 802.15.4 Low Rate Wireless Personal Area Network," IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 6, No 1, November 2013
[18] A. F. Sheta, A. Alamleh, "A professional comparison of c4. 5 mlp svm for network intrusion detection based feature analysis", *The International Congress for global Science and Technology*, vol. 47, pp. 15, 2015.
[19] H. Ismanto and R. Wardoyo, "Comparison of running time between c4.5 and k-nearest neighbor (k-nn) algorithm on deciding mainstay area clustering," *International Journal of Advances in Intelligent Informatics*,vol. 2, no. 1, pp. 1–6, 2016
[20] Hariharan. M, Abhishek H. K, B. G. Prasad, "DDoS Attack Detection Using C5.0 Machine Learning Algorithm", International Journal of Wireless and Microwave Technologies(IJWMT), Vol.9, No.1, pp. 52-59, 2019.DOI: 10.5815/ijwmt.2019.01.06
[21] L. M. Ibrahim, "Anomaly network intrusion detection system based on distributed time-delay neural network (dtdnn)," *Journal of Engineering Science and Technology*, vol. 5, no. 4, pp. 457–471, 2010.