A Survey of DDOS Attacks Using Machine Learning Techniques

Arshi M[1,*], *Nasreen* MD[2], and *Karanam* Madhavi[3]

[1]Assistant Professor, CSE Department, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India
[2]Assistant Professor, CSE Department, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India
[3]Professor, CSE Department, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India

Abstract The DDoS attacks are the most destructive attacks that interrupt the safe operation of essential services delivered by the internet community's different organizations. DDOS stands for Distributed Denial Of Service attacks. These attacks are becoming more complex and expected to expand in number day after day, rendering detecting and combating these threats challenging. Hence, an advanced intrusion detection system (IDS) is required to identify and recognize an- anomalous internet traffic behaviour. Within this article the process is supported on the latest dataset containing the current form of DDoS attacks including (HTTP flood, SIDDoS). This study combines well-known grouping methods such as Naïve Bayes, Multilayer Perceptron (MLP), and SVM, Decision trees.

1. Introduction

Numerous kinds of network assaults arrive with expansion of computing networks, particularly the internet. International ransom ware virus called Wannacry has newly stopped internet services in around 156 countries. As per Kaspersky lab results throughout the fourth quarter, Botnet aided attacks were aimed at assets in nearly 69 countries. The final quarter also experienced the largest DDoS-based Botnet attack that lasts roughly 15.5 days 371 hours Crackers or dark hackers are constantly creating new forms of multilayered DDoS attacksthat happen mainly on a OSI network and application layer. Such attacks have used the spoofed IP addresses to confound source detection and conduct a huge-scale attack. These attacks are quite huge, as the attack traffic absolutely consumes the network spectrum at the peak, thus reducing the legal packets. Ironically, the victims are government entities, finance companies, defense forces and military agencies. Famous sites such as facebook, twitter, wiki leaks etc, had become victims of DDoS that also observed interruptions in routine maintenance resulting in financial failures, depletion of service and lack of access.

This article discusses the different methods of machine learning recognition such asSVM naïve bayesand decision trees for detecting and analyzing different forms of this attacks including, Smurf, UDP flood, HTTP flood. Herethe work has been performed on the novel dataset containing the new kinds of DDoS attacks since no specific data sets containing the current DDoS attacks can be on various layers, including SI-DDoS, HTTP flood[1]. A comparison analysis of the various classification methods is taken out - it's clear from empirical data that MPL has reached the best precision rate.

2. Types of Attacks

DDOS ATTACK

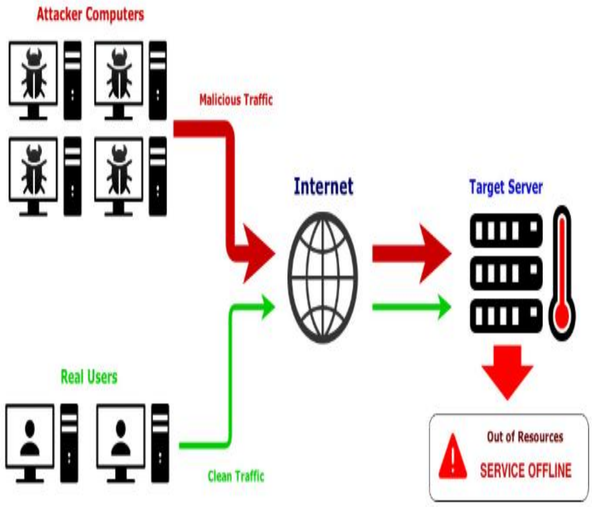

Fig 1: ddos attack

A distributed denial of service (DDoS) attack takes place, if several applications, typically one maybe more application server, flood the capacity or infrastructure of a targeted network. In the below figure1 shows an attack frequently results from several infected systems (e.g. a botnet) that flood the targeted network with traffic.

A.UDP FLOOD

UDP flood is a kind of Denial-of-Service (DoS) volumetric assault in which the attacker attacks and overcomes the host's random ports using IP packets consisting of User Datagram Protocol (UDP) packets. The below figure2 states the hosts look for applications related with certain datagrams throughout this form of attack.If none is detected, the host sends a "Unreachable Destination" packet returns to the sender. The result of this flood bombarding would be that the network is flooded and thus irresponsive to legitimate traffic.

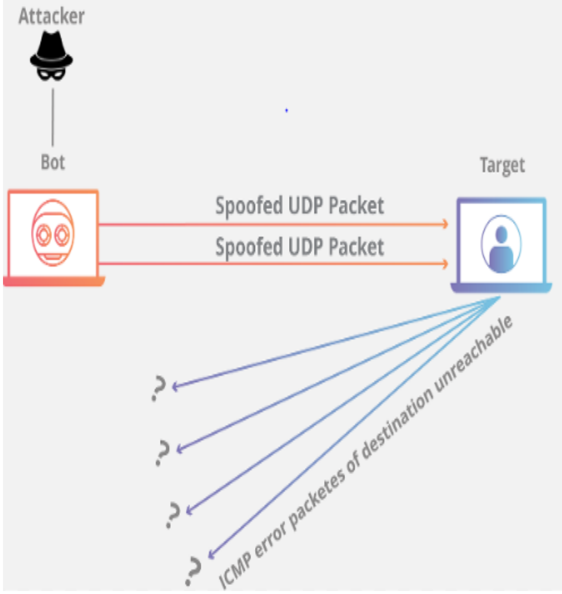

Fig 2: udp flood

B.ICMP(PING) Flood

Ping flood, is identified as ICMP flood, is a popular Denial of Service (DoS) attack where an attacker forces a victim's device down with flooding it with requests for ICMP echo, also called as pings.The figure3 will explains ICMP flood attack, thisattack includes overwhelming the victim's network by request packets, realizing the system will react with just as many reply packets as possible. File types to get a target down for ICMP requests also use custom software or code, like hping and scapy.

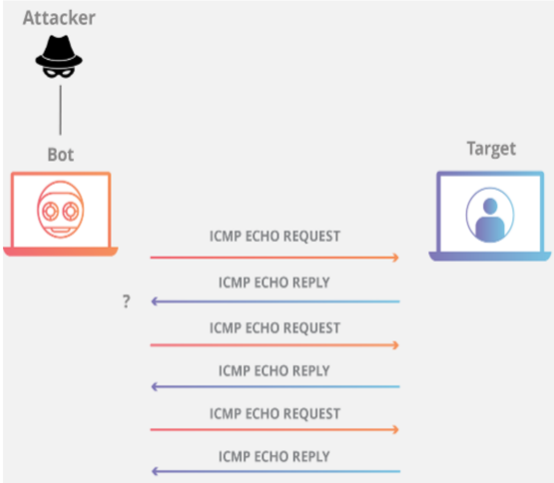

Fig 3: icmp(ping) flood

C.SMURF ATTACK

This is also a one of the ddos attack wherein massive groups of Internet Control Message Protocol (ICMP) packets mostly using spoofed source IP of the victim are broadcast over an IP broadcast address to a computer network. The below figure4 shows by default, many devices on a network will answer it by giving a response to the source IP address. If the quantity of systems over the network receiving then responding to such packet is quite high, so traffic can overwhelm the attacker's computer.

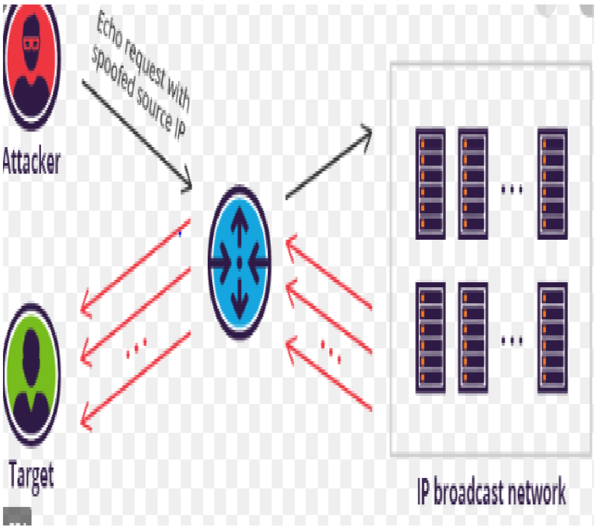

Fig4: smurf attack

D.HTTP FLOOD ATTACK

An HTTP flood is a denial-of-service distributed volumetric (DDoS) attack, it is shown in the figure5, it is built to overburden a selected server with HTTP requests. When the target has also been filled with queries and cannot react to regular traffic, there will be denial-of-service for specific requests from actual users.

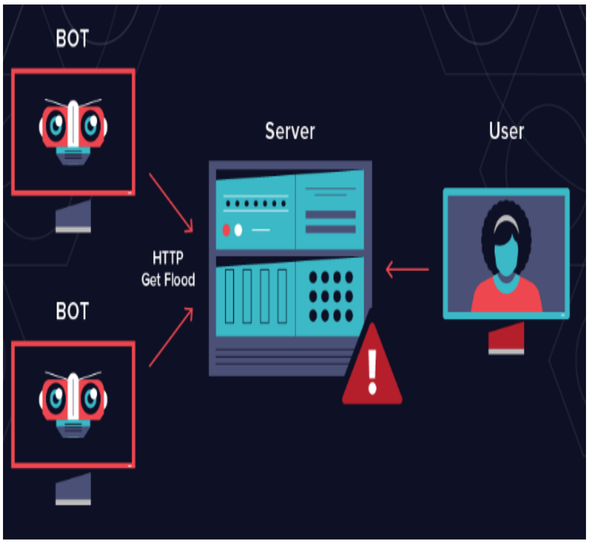

Fig 5: http flood attack

3. Machine learning methods related to ddos attack detection

Signature-based IDS is a human based operation, involving many hours of testing, developing and deploying the signature and creating new signature for unknown attacks too. So providing a less human based system becomes essential.Machine Learning languages derived anomaly-based IDS offers a solution to this issue, helping to incorporate a framework which can learn from data and predict unknown stats information on learned data.

A. Naïve bayes

Naive Bayes is focused on the Bayesian classification model. Establishing classifiers is an easy and simplest method: prototypes which gives class labels to issue cases, defined as the vectors of featuring values, in which the classes labels will be derived among certain finite set.

Kanagalakshmi. R et al.her paper indicated that the use of Secret Naïve Bayes (HNB) produces reliable results compared toStandard Naïve Bayes model. The Hidden Naive Bayes (HNB) technique could be used to anticipate intrusion problems such as DOS attacks benefiting from strongly associated dynamic characteristics and big network Data stream capabilities[13]. It is a paradigm of data mining that looses the naive techniques of Bayes Presumption of implicit impartiality. In his paper Mouhammad Alkasassbeh et al[1] collected a new dataset consisting of DDOS attacks in various layers of the network. DDoS detection is performed utilizing three Multilayer Perceptron (MLP) methodologies, Naïve Bayes, and Random Forest.

In [15] Jasreena Kaur Bains et al suggested a hierarchical layered method for the detection of attack rates. System used Naive Bayes classifier with K2 learning method between each attack class on reduced NSL KDD dataset. Each layer is trained in the research methodology to recognize a single form of attack. To raise the detecting rate, the output of one layer is moved on to another layer.

B.Support Vector Machine

Support Vector Machine (SVM) was at first introduced by Vapnik[7] and got significant attention in the research community of machine learning . SVM makes classification and regression using the supervised method of learning.Based on a group of trainied examples, each of which is marked as methods are divided into two classifications, an SVM algorithm creates a design which predicts that the new example tends to fall into one among the two.

In 2010 Vipin Das et al.[9] Work conducted using RST (rough set theory) and SVM (supporting vector machines) to classify DOS attacks; At first packets from the network were obtained, and the data is immediately processes by RST. The selected RST feature sets will be given to the SVM model for learning and testing, and so on. The results are then analyzed with PCA and show that RST and SMV are capable of doing so and the improving of efficiency is done by the false positive ratio.

T. Subbulakshmi et al[10] written an article aimed at tracking the online network and instantly activating a security techniquein the event of any suspicious behavior. This strategy allows for identification of both non-spoofed and spoofed IPs. Enhanced Support Vector Machines (ESVM) is used by the author to detect mechanisms for detecting spoofed IPs and Hop Count Filtering to find spoofed IPs These IPs are used to start the defense. The

Lanchester Rule is used to determine the attack force used to cause the defence mechanism.

Rung-Ching Chen et al[11] written a paperat where RST and SMV were used to identify Dos Attacks supplied to SVM by specific feature set (obtained from RST); The report has wrote by T.Subbulakshmi et al[10] Focused on creating and detecting the DDoS dataset and using Enhanced Support Vector Machines(ESVCM). The EMCSVM are used to detect attacks in various classes for a generated dataset, and SVM is used for EMCSVM evaluation.

C. Decision Trees

One of the basic techniques used in machine learning and data mining is the decision tree. It is also utilized as a predictive model where findings regarding an object are mapped to assumptions about the desired value of the item.A decision tree may be used in the decision data analysis to visually and explicitly indicate decision making. The data set is studied and constructed in this method. Consequently, if the new data element is given for classification, the prior dataset will classify it appropriately.

Decision tree algorithm is used to detect the DOS attacks. In his article, HodaWaguih[2] suggested a data mining method for detecting DOS attacks, using classification methods. In the case of DoS attacks the approach above focuses its classification of "normal" traffic over "anomalous" traffic. The paper looks at the efficacy of the J48 a type of decision tree algorithm for DoS attack detection and then compares it with the other rule-based algorithms, such as Decision table and oneR.

Md Dewan. In their paper, Farid et al.[3] suggested an anomaly-based network intrusion detection learning algorithm which prevents attacks from regular activities and recognizes multiple kinds of intrusions utilizing decision tree algorithm. The data set used is KDD99 dataset for network intrusion detection.

D. Artificial Neural Network

Chandrika Palagiri demonstrated that, particularly for a specific attack, a modeling network can obtain a reasonable outcome to show a Neural Network. Scientists also concentrate on a Neural Network which can take fast decisions and identify them in real time.

Resilient back propagation (RBP) is selected as the basis classifier for the work in a paper wrote by Madhav Kale et al[21]. This paper focused on increasing the RBP classifier's efficiency through a fusion of classifier outputs and cost reduction approach from Neyman Pearson, for actual classification decision. The two factors evaluated to learn the RBP Boost classification algorithm's efficiency were Detection accuracy and cost per sample.

The purpose of this paper by Md Salem et al[22] was to decide whether a firewall can examine its traffic patterns to recognize targeted denial of service. In this paper, a baseline of the network was determined by carrying out the statistical analyses of firewall logs for a hugenetwork.. In this paper a network baseline was calculated by performing statistical analysis of firewall logs for a wide network.Estimatedtraffic rates were calculated for comparison with the baseline utilizing linear regression and the Holt-Winter approaches. Analysis results were good, with deviation from the predicted rejected packet rates suggesting a massive campaign in the network.

In [23] author Mohammad Masoud Javidi et al introduced IDS, which uses supervised neural network to identify malicious of DDOS in the NSLKDD database. The researcher also used signature-based methodology in the proposed IDS. IDSs are developed using a neural network capable of detecting various kinds of DoS attacks and having a different IDS for every one to identify the particular attack.

E.K-means clustering

It is a clustering technique[5] widely used to partition a collection of data in groups k automatically. The K-means clustering algorithm works by choosing k initial cluster centers in a data set and then refining them recursively as describes

1. Every example shall be allocated to its nearest cluster core.

2. It updates the mean of its component cases to each of the cluster centres. The algorithm converges when the allocation of instances to clusters does not alter further

Mangesh, D. Salunke et al [7] introduced a design which gathers packets, the packet is controlled by the specification like selection of features, and so on. Therefore k-means and naïve Bayes methods are being required to determine if the packet is usual or it is DOS attack.

4. Conclusion

It is concluded after a detailed analysis that web attacks are risky and that IDS / IPS may not tackle the new attacks that affect the networks. Machine learning approaches play a critical role in gaining exposure to the intensity of the assault and thereby making enterprises take suitable measures to limit certain attacks.

5. Future enhancement

A thorough study of the data sets containing the latest types of attacks such as HTTP flooding, SIDDoS, Smurf and UDP flooding etc., collected from the college network using deep learning techniques, will be carried out in future. It will allow the extent of attacks on the network connection or any organization to be assessed, so that the network is subject to correct firewall rules.

References

1. M. Alkasassbeh, G. Al-Naymat et.al," Detecting Distributed Denial of Service Attacks Using Data Mining Technique," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, pp. 436-445, 2016. Science and Information Technologies, Vol. 6 (2), pp. 1096-1099, 2015.
2. HodaWaguih, "A Data Mining Approach for the Detection of Denial of Service Attack", International Journal of Artificial Intelligence, vol. 2 pp. 99106(2013).
3. Dewan Md. Farid, Nouria Harbi, EmnaBahri, Mohammad Zahid ur Rahman, Chowdhury Mofizur Rahman,‟ Attacks Classification in Adaptive Intrusion Detection using Decision Tree "International Journal of Computer, Electrical, Automation, Control and Information Engineering, Vol:4, No:3, 2010.
4. Kiri Wagsta,ClaireCardie ‚Seth Rogers ‚Stefan Schroedl,‟ Constrained K-means Clustering with Background Knowledge‟ Proceedings of the Eighteenth International Conference on Machine Learning, 2001, p. 577-584.
5. Singh, S.K., Gupta, A.K. Application of support vector regression in predicting thickness strains in hydro-mechanical deep drawing and comparison with ANN and FEM (2010) CIRP Journal of Manufacturing Science and Technology, 3 (1), pp. 66-72
6. Ramesh.G, Madhavi, K. "Summarizing Product Reviews using NLP based Text Summarization", International Journal of Scientific & Technology Research, September 2019. (Scopus).
7. Mangesh Salunke, RuhiKabra, Ashish Kumar.‟ Layered architecture for DoS attack detection system by combine approach of Naive Bayes and Improved Kmeans Clustering Algorithm‟, International Research Journal of Engineering and Technology (IRJET), Volume: 02 Issue: 03, June-2015.
8. Ramesh G, Madhavi K., "Best keyword set recommendations for building service-based systems" International Journal of Scientific and Technology Research, October, 2019.
9. T. Subbulakshmi et.al, ‟A Unified Approach for Detection and Prevention of DDoS Attacks Using Enhanced Support Vector Machine and Filtering Mechanisms‟, ICTACT Journal on Communication Technology, June 2013.
10. Yogeswara Reddy B, Srinivas Rao J, Suresh Kumar T, Nagarjuna A, International Journal of Innovative Technology and Exploring Engineering, Vol.8, No. 11, 2019, pp: 1194-1198.
11. T. Subbulakshmi, K. BalaKrishnan; S.M.Shalinie ; D.Anand Kumar ; V.Ganapathi Subramanian ; K. Kannathal.‟Detection of DDoS attacks using Enhanced Support Vector Machines with real time generated dataset‟, ICTACT Journal on Communication Technology, Volume: 04, Issue: 02 , June 2013.
12. Kanagalakshmi.R, V. Naveenantony Raj,‟ Network Intrusion Detection Using Hidden Naïve Bayes Multiclass Classifier Model,‟ International Journal of Science, Technology & Management ,Volume No.03, Issue No. 12, December 2014.
13. A. Bivens, C. Palagiri, R. Smith, B. Szymanski, M. Embrechts, et al, "Network-based intrusion detection using neural networks," Intelligent Engineering Systems through Artificial Neural Networks, vol. 12, no. 1 , pp. 579–584, 2002.

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. 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.
16. Ch. Mallikarjuna Rao, G. Ramesh, Madhavi, K., "Feature Selection Based Supervised Learning Method for Network Intrusion Detection", International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8, Issue-1, May 2019.
17. R Vijayasarathy, Balaraman Ravindran,S.V Raghavan,‟A System Approach to Network Modeling for DDoS Detection using a Naive Bayesian Classifier,‟ Department of Computer Science and Engineering IIT Madras, India.
18. V. Hema and C. EmilinShyni, ‟ DoS Attack Detection Based on Naive Bayes Classifier, ‟ Middle-East Journal of Scientific Research 23 (Sensing, Signal Processing and Security): 398-405, 2015.
19. Afrah Nazir, " A Comparative Study of different Artificial Neural Networks based Intrusion Detection Systems" International Journal of Scientific and Research Publications, Volume 3, Issue 7, July 2013.
20. SamanehRastegari, M. Iqbal Saripan and MohdFadlee A. Rasid,‟ Detection of Denial of Service Attacks against Domain Name System Using Neural Networks‟, IJCSI International Journal of Computer Science Issues, Vol. 6, No. 1, 2009.
21. Madhav Kale and D.M. Choudhari, ‟ DDOS Attack Detection Based on an Ensemble of Neural Classifier, ‟ IJCSNS International Journal of Computer Science and Network Security, VOL.14 No.7, July 2014.
22. Mohammed Salem, Helen Armstrong,‟ Identifying DOS Attacks Using Data Pattern Analysis,‟ Australian Information Security Management Conference Security Research Institute Conferences,2008.
23. Mohammad Masoud Javidi, Mohammad Hassan Nattaj, ‟ Journal of mathematics and computer Science 6 (2013), 85-96.
24. Thirupathi, N., Madhavi K., Ramesh G., Sowmya Priya, K. "Data Storage in Cloud Using Key-Policy Attribute-Based Temporary Keyword Search Scheme" (KP-ABTKS), Lecture Notes in Networks and Systems, 2020.
25. Madhavi.K., G. Ramesh, G. Lavanya "Load effectiveness on coverage-technique for test case prioritization in regression testing", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-7 May, 2019.