Detection of Distributed Denial of Service Attack via Machine Learning

THESIS

FOR THE DEGREE OF

MASTER OF TECHNOLOGY
IN
INFORMATION TECHNOLOGY

By

Mudit Rathore: ICM2015502

UNDER THE SUPERVISION OF

Dr. Abhishek Vaish
Assistant Professor
INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, ALLAHABAD

Acknowledgement

This project consumed a tremendous amount of research and dedication, which would not have been possible without the support of specific individuals. First and foremost, I would like to express my gratitude to Dr. Abhishek Vaish for his valuable feedback and unending inspiration throughout this project. His professional guidance and constructive criticism during my briefings and discussions have helped me pursue this research more effectively. I thank him for accepting me to his research group and encouraging me to pursue the topic of my liking. I am also very thankful to him for never saying no to any of my request, however frequent they were!

I feel thankful to my brother, Ankit Rathore, who has always been an inspiration for me with his sense of discipline and maturity. He would always listen to my research ideas and plans without ever losing patience or interest. I would also like to thank Ayushi Asthana for her advice and expertise with debugging my dirty code. I am grateful to all my lab mates who started as colleagues but soon became excellent friends. They helped in maintaining a healthy atmosphere conducive to both research and fun.

None of this would have been possible without my family so, at last, I would like to thank my parents for their guidance and the confidence they have instilled in me.

Abstract

A distributed denial of service (DDoS) attack is one of the most potent attacks on the internet. It aims to overwhelm the target with more traffic than it can accommodate. The goal of the attack is to overload the target leading to service denial and hindered operations. It makes the target incapable of serving requests from legitimate users. The traffic of this attack can consist of incoming messages, request for connections, or fake packets. It is a form of many-to-one attack, in which multiple systems attack a single target. The systems sending and receiving DDoS packets are both victims of the attack.

New records for DDoS attack are set every year, the average attack size has increased, and the attackers are finding better ways to amplify the impact. According to the NETSCOUT Threat Intelligence Report 2019, there was a 39% increase in the attack frequency for DDoS attacks in the first half of 2019 as compared to the 1H 2018. Bad actors feasted on the juicy middle of attack sizes, resulting in a 776% growth in attacks between 100 Gbps and 400 Gbps. Attackers have increasingly targeted satellite communications and wireless telecommunications, which experienced a 255 percent and 193 percent increase in attack frequency, respectively.

Another contributor to the increasing DDoS attacks is the booming number of IoT devices added to the network YoY. On average, 7.7 million IoT devices are connected to the internet every day, many of them having known security issues or no security at all. As per the report, It takes just five days from new attack vector discovery to weaponization, making these powerful attacks available to anyone with a grudge.

Detection of a DDoS attack is the process to analyze all of the information provided by each data packet to assess its legitimacy. Our

system evaluates each packet and renders a programmatic decision as to whether the packet is legitimate or part of an attack. This system is developed to accurately predict the legitimacy of a particular data packet based on various parameters.

Timeline

This thesis is divided into following 9 phases:

- Phase 1: Understanding the problem.
- Phase 2: Study of Related work in this field.
- Phase 3: Understanding the data.
- Phase 4: Exploring different techniques and models.
- Phase 5(Proposed): Data Analysis.
- Phase 6(Proposed): Developing the detection model.
- Phase 7(Proposed): Simulation and Testing of the model on our dataset.
- Phase 8(Proposed): Comparing the performance of different techniques.
- Phase 9(Proposed): Testing the model on various secondary dataset.

Table of Content

Section 1: Introduction

Section 1.1: Background

With the Internet becoming an indispensable part of human life, providing security of data passed over the Internet is getting more and more crucial. The Internet design supports scalability and openness without any concern for security, making it too easy a target for malicious users. The widespread usage of the Internet has made cyber attacks a global issue.

Cyberattacks use malicious code to alter computer code, logic, or data, resulting in disruptive consequences that lead to cybercrimes. Cybercrimes are relentless, undiminished, and unlikely to stop because it is just too comfortable and too rewarding, and the risks of being discovered and punished regarded as being too low. Center for Strategic and International Studies (CSIS), in partnership with McAfee, published a global report discussing the impact of cybercrime on economies worldwide. It concludes that close to $600 billion, nearly one percent of global GDP, is lost to cybercrime each year, which is up from a 2014 study that put global losses at about $445 billion. According to Hardik Modi, Senior Director for Threat Intelligence, NETSCOUT, "It's hard to express the scale of today's cyber threat landscape, let alone its global impact." According to the interview, there were nearly four million DDoS attacks around the world in the last six months(2019), the attack frequency grew by 39 percent in 2019. The NETSCOUT ASERT team saw 20,000 different samples per month from a single class of IoT malware, and it takes only five days from the discovery of a new attack vector to the availability of tools for the script-kiddie designed to exploit that vulnerability. This information is more than enough to describe the threats that cyberattacks pose and how open and easy it is to exploit vulnerabilities in a system.

Because of the increased interconnection among information systems and networks, the consequences of successful attacks by malicious individuals can have far-reaching implications. The results of malicious attacks can include financial loss, loss of reputation, a drop in the value of a company's stock, and many legal issues.

Distributed Network attacks are one of the major types of cyberattacks. Distributed Network Attacks are also known as Distributed Denial of Service (DDoS) attacks. That is because this type of attack takes advantage of the specific capacity limits that apply to any network resources. The DDoS attack sends multiple requests to the attacked web resource and overwhelms its resources so that it is unable to provide the required services. It generally happens in the distributed mode, messages to a target computer come from large numbers of hosts where a software has been planted to activate at a particular time or upon receiving a particular command.

Section 1.2: Scope

This study focuses specifically on DDoS attacks. Threats such as viruses, trojans, worms have taken distributed Attack to the next level, producing the so-called Denial of service attack. Because of the many-to-one dimension of this type of attacks, they are generally compelling and devastating and not so easy to detect and stop. Scientific community forecasts that the disruptive power of DDoS attacks, their sophistication, and damage capacity tends to increase at a very high rate, becoming a very critical threat for new and emergent internet services.

A DDoS sends undesired packets, which can be huge packets with heaps of data, little packets very rapidly, or packets that demand additional processing by the system. It can also make the targeted device waste time waiting for a response that never comes. The target is kept busy dealing with malicious packets and improper communication methods that little or no time is left to respond to regular incoming requests – so legitimate users are denied service.

DoS attack creates an outage or slowdown of a website, web application, web API, or network. An attack can cause downtime for minutes, hours, or days – and prevent legitimate users from buying products, using a service, or getting information from the target.

One of many Internet-facing organizations may be affected by a single attack. An attack that disrupts a standard service relied on by many websites, such as a domain name system (DNS) service, can create a more extensive outage. As a result, DDoS attacks are a significant threat to enterprise and application security and business continuity.

They are extremely risky, not so easy to detect and cause significant damage to company infrastructure. Early detection of DDoS attacks is necessary because of the catastrophic consequences, both financial and strategic.

Section 1.3: Outline

We use machine learning techniques for the detection of the DDoS attacks because the Attacks vectors are enormous in numbers. There are many challenges as well, which includes feature extraction, feature detection, data imbalance, and preprocessing. Thus, motivation for this research comes with the challenges it offers, the magnitude of the effect, and scale of impact.

The data available in numerous public repositories are used to study the different characteristics of DDoS. It helps in exploring the information available in the data about the nature of attacks which can be used to make conjectures about the relationship between various features. Then using the concepts of Machine Learning, we try to classify the traffic as legitimate or attack. More specifically, we propose Ensemble Learning for the classification because of the advantages it presents as a statistical model. We evaluate and compare bagging, boosting, blending, and stacking techniques and analyze the features for different kinds of attack in our proposed system, including the most prevalent DDoS attacks: flooding attacks, spoofing attacks and brute-force attacks.

For the study, we have used the CIC-IDS-2017 dataset. CICIDS2017 dataset contains benign and the most up-to-date frequent attacks, which resembles the actual real-world data (PCAPs).

In the following sections, we talk about the previous works on the topic of DDoS attacks, Machine Learning capabilities, and some successful detection systems.

Section 2: Theory

Section 2.1: Machine Learning

Data imbalance is one of the major problems prevailing in real-time anomaly detection datasets. A dataset is considered to be imbalanced if one of its classes dominates over the rest of the classes. This property can be observed with most prominence in binary classification datasets because most binary classification datasets are implicitly imbalanced. The primary class entries are higher, while the minor class entries occupy a tiny space in the dataset. This unequal distribution can be of form 100:1, 1000:1, 10000:1, and so on.

Classification, being a supervised learning process, depends mainly on the training data. The level of training plays a significant role in the resultant accuracy of the classifier. Imbalanced nature of the data sets acts as a massive downside in this scenario. Due to the minimal occurrence of the minor classes, the classifier is biased towards the majority class and hence provides inaccurate predictions. Oversampling, under-sampling, and hybrid sampling can be used to counter the imbalance.

Sampling is using a predetermined number of observations from a larger population. The methodology used to sample from a larger population depends on the type of analysis and the dataset distribution.

Advantages of sampling include ease of use and accuracy of representation. There is no need to divide the population into subpopulations or take any steps further than plucking the number of research subjects needed at random from the larger group. Again, the only requirements are that randomness governs the selection process and that each member of the larger population has an equal probability of selection.

Random forest classifier creates a set of decision trees from a randomly selected subset of the training set. It then aggregates the votes from different decision trees to decide the final class of the test object. A random forest is a meta estimator that fits several decision tree classifiers on various sub-samples of the dataset and use averaging to improve the predictive accuracy and control over-fitting. The sub-sample size is always the same as the original input sample size, but the samples are drawn with a replacement if bootstrap=True (default).

Gradient boosting is a machine learning technique for regression and classification problems, which produces a prediction model in the form of an ensemble of weak prediction models, typically decision trees. It builds the model in a stage-wise fashion as other boosting methods do, and it generalizes them by allowing optimization of an arbitrary differentiable loss function.

In machine learning, support vector machine is a supervised learning model that is used for classification and regression analysis. An SVM model represents the data samples as points in space. We manipulate the vector space so that these data samples can be placed in separate categories, and the distance between these categories is considerable. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall.

Robust Scaler removes the median and scales the data according to the quantile range (defaults to IQR: Interquartile Range). The IQR is the range between the 1st quartile (25th quantile) and the 3rd quartile (75th quantile). Centering and scaling happen independently on each feature (or each sample, depending on the axis argument) by computing the relevant statistics on the samples in the training set. The median and interquartile range are then stored to be used on later data using the transform method.

Ensemble methods are meta-algorithms that combine several machine learning techniques into one predictive model to decrease variance (bagging), bias (boosting) or improve predictions (stacking). It relies on the assumption that each model would look at a different aspect of the data which yield to capturing part of the truth. Models that perform well independently generally perform better when combined. Therefore, this would result in more accurate predictions and lower generalization errors. The performance of the ensemble model generally improves on adding models. Try to combine models that are as much different as possible. Doing so reduces the correlation between the models and improves the performance of the ensemble model and outperforming the best model significantly. In the worst case where all models are perfectly correlated, the ensemble would have the same performance as the best model and sometimes even lower if some models have inferior performance. As a result, pick models that are as good as possible.

Section 3: Literature Review

Detection of Distributed denial of service attack using machine learning is a topic of interest and has been researched by the technical community a lot. Majority of the works in this area for detection have considered only one protocol and have researched and performed experiments relating to that particular protocol. The proposed algorithms for attack detection are broadly classified into four techniques, clustering, classification, statistical analysis, and a technique that uses statistical concepts for attribute selection and then uses machine learning for predicting DDoS Attack.

The statistical learning focuses on the process generating the data, nature of the data, modeling the relationship between data and the underlying population, validity of the model, and predictive analysis in different scenarios that guide the future actions. Whereas the Machine Learning techniques focus on building a system that improves its performance based on previously acquired knowledge and obtained results, and use computational techniques to "learn" information straight from data without relying on a predetermined equation as a model.

In this section, we present a study of existing literature on DDoS attack detection methods. We briefly summarize the recent trends in DDoS attack detection mechanisms.

Marwane et al. [8] proposed a detection scheme using Decision Tree based on C4.5 algorithm to detect DDoS flooding attacks. They have focused on layer three and layer four of the seven-layer OSI Model. They have done a comparative study of DDoS attack detection results in different duration. The correct classification and the attack detection time, in terms of the running time 0.58s for C4.5, Naive Bayesian, on the other hand, takes 1.1s. For correct classification, C4.5 works better than Naive Bayesian; it has a score of 98.8 (%) while Naive Bayesian scored 91.4 (%).

Reyhaneh et al. [11] proposed an anomaly-based DDoS detection method. Their study uses various features of packets attack using thorough analysis using the Radial Basis Function neural networks on UCLA dataset. They used a seven feature vector to activate the RBF neural network at each time-window. This vector included average packet size, number of packets, time interval variance, packet size variance, number of bytes, packet rate, and bitrate on UCLA dataset and their proposed system. They divided the data into two different categories, i.e., attack packet and reasonable or legitimate packet. If the incoming traffic constitutes attack traffic, the source IP of the specific packets concerned is filtered. The results showed that the system proposed by them could make real-time detection accuracy better than 96% for DDoS attacks.

Cheng et al. [11,12] presented an approach based on the notion of flow interaction. It starts with the idea that for healthy communication flows, entities interact, whereas, during attacks, entities might not present interaction mainly because there are spoofing and false source IP addresses. They have proposed an IP Address Interaction Feature algorithm based on the addresses' interaction, abrupt traffic change, addresses' many-to-one dissymmetry, distributed source IP addresses, and concentrated target addresses. It is designed to describe the essential characteristics of the network flow states. SVM is used to classify whether flows correspond to attack or regular traffic according to the notion of interaction on IP Address Interaction Feature Time Series. This approach presents a high detection rate, around 100%, and meager false alarm rates, even with multiple network flows, with the ability to even perceive low rate DDoS attacks mixed with a high volume of legitimate traffic.

Bao et al [7] presents a mechanism for attack detection using data collected via Simple Network-Based Protocol (SNMP) MIB statistical data gathered from SNMP agents, instead of using raw communication packets. Via an ensemble of SVM, traffic classification is performed to initially detect whether an attack is occurring and then the specific attack type. A useful feature selection mechanism known as Correlation feature selection (CFS) is used to select the SNMP MIB variables, and the data is gathered expertly by the MIB update time prediction mechanism. This approach presents a performance of 99.27% of attack detection rate, with 1.9% false positives and 0.73% false negatives.

Bharathi et al. [13] presented a framework based on Principal Component Analysis (PCA) for the detection of application-layer or higher layer of the seven-layered OSI Model for DDoS attacks, specifically, those based on HTTP traffic. It works by building profiles from the browsing activity of users. The detection structure uses sequence order independence's two fundamental and subsidiary attributes rather than the sequence order of web page to construct a behavior-based attribute vector matrix. The framework proposed presents little attack detection time, in comparison to other existing methods; however, no information is available on false positives. The significant challenges faced by them to Detect App-DDoS were (1)App-DDoS uses higher layer protocols such as HTTP to pass through the detection system designed for a lower layer. (2)Along with flooding, App-DDoS also consumes resources of the target server and either track the average request flow of the legitimate user and replicate it for attacking the server or engages large-scale botnet to create low rate attack flows. This replication causes the detection systems detecting the DDoS attack more complex. The proposed framework accurately identifies 94.4% of App-DDoS attacks and requires 0.368 seconds as an average detection rate of the application layer DDoS.

Raj Kumar et al. [14] present an approach based on an ensemble of neural classifiers, on the idea that by combining classifiers, it is possible to reduce the overall error. Applied to the detection of TCP, UDP, ICMP and HTTP based attacks, it presented a 98% of accurate detections, and around 3% of false positives.

Kale et al. [15] evaluated the performance of a comprehensive set of machine learning algorithms for selecting the base classifier using the publicly available KDD Cup dataset and chose Resilient Back Propagation (RBP) as the base classifier for their research. Their proposed classification algorithm, RBPBoost, is achieved by combining ensemble of RBP classifier outputs and Neyman Pearson cost minimization strategy, for final classification decision The proposed algorithm achieved a high detection accuracy of 55.36 with fewer false alarms.

Jia et al. [16] proposed a DDoS attack detection method based on hybrid heterogeneous multi-classifier ensemble learning and designed a heuristic detection algorithm based on Singular Value Decomposition (SVD). They used Knowledge Discovery and Data mining (KDD) Cup 1999 dataset for verification of their detection model, which comprises about five million network traffic records. The comparison of heterogeneous detection model with Random Forest, k-NN, and Bagging classifiers, shows that when processed by SVD, the True Negative Rates (TNR) of their model are about 99.4% for different threshold values whereas the TNRs of Random Forest, k-NN, and Bagging are about 99.8%, 99.1%, and 17.8%, respectively. When Random Forest, k-NN, and Bagging are processed by SVD, respectively, for the comparison of accuracy, the accuracy of their model is 99.8% as compared to accuracies of Random Forest, k-NN, and Bagging which are about 99.9%, 21.2%, and 82.3%, respectively. The precision of the model, on the other hand, is 99.84% for our model and 99.9%, 0%, and 81.6% respectively for Random forest, k-NN, and Bagging. Therefore, the accuracy and precision of our model are very close to that of Random Forest, and it is superior to k-NN and Bagging.

Saied et al. [9] have developed their detection system and have integrated it with snort-AI, which is based on the Snort Signature Intrusion Detection System project. They have designed a model for detection and mitigation of the known DDoS attacks and unknown DDoS or zero-day DDoS attack in real-time environments. They chose an Artificial Neural Network (ANN) algorithm to detect DDoS attack based on specific characteristic features that separate DDoS attack traffic from regular traffic. They are coupling the output of their detection system with the destination IP address to command tables to mitigate forged packets while allowing genuine traffic to pass through. They have also used RSA encryption mechanism to encrypt the messages over TCP connection where each detector acts as a sender and a receiver. They have evaluated their solution based on accuracy, sensitivity, i.e., the ability to identify positive results and specificity, i.e., the ability to identify negative results. They have compared their detection system with other algorithms as well, and the table summarizes the results.

Wu et al. [17] in their study presented the usage of Neural Networks in the detection of DDoS attacks against DNS servers; the overall performance obtained in their work is 96.5%, with 0% false positives.

Wang et al. [18] propose detection of DDoS attacks on IP flow using neural network classification, specifically the backpropagation network. Although we don't mean to present too many details, graphics show low false positives and high reasonable classification rates.

PCA, Neural networks, and Random Forest Classifiers are good contenders for a base learner in our ensemble model. The statistical model of these three classifiers is significantly different. PCA is based on the orthogonal transformation of data space while Random Forest is a tree-based classification technique which works on feature subsets. Neural Network is yet another league of classifiers with iterative fine-tuning of perceptron layers to achieve accurate results. Because of the diverse methods and lower correlation between the parametric constraints, we can expect the ensemble model created with these base classifiers to cover a broader range of data points in the classification and give better results than any single model on the given dataset.

Section 4: Research Objective

The objective is to develop a machine learning-based distributed denial of service detection system. We will be developing a detection system using Ensemble Learning based on base learners and meta learners. Principal Component Analysis, Neural networks, and Random Forest Classifiers are good contenders for a base learner in our ensemble model.

We will evaluate and compare bagging, boosting, blending, and stacking techniques and will be analyzing features for different kinds of attack in our proposed system, including the most prevalent DDoS attacks: flooding attacks, spoofing attacks and brute-force attacks.

We will be using CIC-IDS-2017 dataset. CICIDS2017 dataset contains benign and the most up-to-date common attacks, which resembles the true real-world data. I want to be on top of you.

Section 4.1 Research Challenges

- Detection of DDoS attack is a very challenging task because of the nature of a DDoS attack, it is very hard to mitigate as it penetrates right through the open ports on the firewall.
- Attributes such as source IP address, destination IP address, and packet rate are generally very good measures of detecting a DDoS attack, but, these attacks are often sourced from virtual machines in the cloud, rather than from the attacker’s own machine, to achieve anonymity and higher network bandwidth; thus traceability and identification becomes very difficult.
- Encryption of network data packets are also major challenges in this field as the defense can only access the metadata of the packet

i.e., the information provided in the header and not the information provided as the content of the data packets.

- The wide variety of network protocols means difference in the header profile and the information carried by packets, it becomes quite difficult to design a system to monitor all these packets with their different feature vectors and information analysis requirements.
    - This is especially critical because a DDoS attack does not necessarily have a single type of packet.

Section 4.2 Proposed Outcome

The primary outcome of this study is to be able to classify incoming network traffic packets into legitimate and DDoS attack packets and the secondary outcome of the study is to be able to identify the nature of DDoS attack, categorization of the network packets and, building a detection system which can identify a wide variety of attack packets from the incoming network traffic.

Section 5: References

[1]"Economic Impact of Cybercrime", Csis.org, 2019. [Online].

[2]"The Mainstreaming of Cybercrime | NETSCOUT", NETSCOUT, 2019. [Online].

[3]"DDoS Attacks | Akamai", Akamai.com, 2019. [Online].

[4]"What are the advantages of using a simple random sample to study a larger population?", Investopedia, 2019. [Online]

[5]"3.2.4.3.1. sklearn.ensemble.RandomForestClassifier — scikit-learn 0.21.3 documentation", Scikit-learn.org, 2019. [Online].

[6]"Introduction to Extreme Gradient Boosting in Exploratory", Medium, 2019. [Online].

[7]C.-M. Bao, “Intrusion detection based on one-class SVM and SNMP MIB data,” in 5th International Conference on Information Assurance and Security, IAS 2009, vol. 2, pp. 346–349, 2009.

[8]Zekri, Marwane & El Kafhali, Said & Aboutabit, Noureddine & Saadi, Youssef. (2017). DDoS attack detection using machine learning techniques in cloud computing environments. 1-7. 10.1109/CloudTech.2017.8284731.

[9]A. Saied, R. Overill and T. Radzik, "Detection of known and unknown DDoS attacks using Artificial Neural Networks", Neurocomputing, vol. 172, pp. 385-393, 2016.

[10] R. Karimazad and A. Faraahi, "An Anomaly-Based Method for DDoS Attacks Detection using RBF Neural Networks", 2011International Conference on Network and Electronics Engineering, vol. 11, 2011.

[11]J. Cheng, J. Yin, Y. Liu, Z. Cai, and C.Wu, “DDoS attack detection using IP address feature interaction,” in International Conference on Intelligent Networking and Collaborative Systems, INCoS 2009, pp. 113–118, 2009.

[12] J. Cheng, J. Yin, Y. Liu, Z. Cai, and M. Li, DDoS attack detection algorithm using ip address features, vol. 5598 LNCS. 2009.

[13]R. Bharathi and R. Sukanesh, “A PCA based framework for detection of application layer DDoS attacks,” WSEAS Transactions on Information Science and Applications, vol. 9, no. 12, pp. 389–398, 2012.

[14]P. Raj Kumar and S. Selvakumar, “Distributed denial of service attack detection using an ensemble of neural classifier,” Computer Communications, vol. 34, no. 11, pp. 1328–1341, 2011.

[15]Kale, M.K. (2014). DDOS Attack Detection Based on an Ensemble of Neural Classifier.

[16]Bin Jia, Xiaohong Huang, Rujun Liu, and Yan Ma, “A DDoS Attack Detection Method Based on Hybrid Heterogeneous Multiclassifier Ensemble Learning,” Journal of Electrical and Computer Engineering, vol. 2017, Article ID 4975343, 9 pages, 2017.

[17]J. Wu, X. Wang, X. Lee, and B. Yan, Detecting DDoS attack towards DNS server using a neural network classifier, vol. 6354 LNCS. 2010.

[18]D. Wang, G. Chang, X. Feng, and R. Guo, Research on the detection of distributed denial of service attacks based on the characteristics of IP flow, vol. 5245 LNCS. 2008.