Detecting Network Intrusion based on Machine Learning Algorithms

[1]S.Kavitha,[2]D.Subiksha,[3]J.Aarthi,[4]M.Priyanga,[5]M.Malathi

[1]Assistant Professor,[2,3,4,5]UG Students
Department of Computer Science and Engineering,
Velammal College of Engineering and Technology, Madurai, India

ABSTRACT:
An IDS is a hardware or software application that monitors network traffic data on a system or a network .To provide security for network only firewall and antivirus is not sufficient so, there is need to give more security to the network so, Intrusion Detection System is used. DDoS defense mechanism named CoFence which facilitates a domain-helps-domain collaboration network among NFV-based domain networks. CoFence through resource sharing helps to handle large volume of DDoS attacks. Specifically, it designs a dynamic resource allocation mechanism for domains so that the resource allocation is fair, efficient, and incentive compatible.Current Static Detection Techniques only detect the known malicious attacks, but it also intends to provide the NIDS the capability to analyse and classify the malicious contents along with the accuracy. Honeypot is used to detect intruders and to identify all the malicious activities performed over the internet. Naïve Bayes algorithm is used for classification of the data into normal and abnormal activities along with the accuracy.

Keyword Index: *DDoS Attack, Network Function Virtualization (NFV), Network intrusion detection systems; Machine learning.*

I. INTRODUCTION

Over the past decades, Artificial intelligence (AI) stream has become the broad and exciting field in computer science as it prepare the machines to perform the tasks that human being may do. and it aims to train the computers to solve real world problems with the maximum success rate. As perceiving scientific growth and advancement in technology AI systems are now capable to learn and improve through past experiences without explicitly assistance code if they exposed to new data. Eventually it leads to technology of Machine learning (ML) which uses learning algorithms to learn from the data available [1]. Machine Learning uses data mining techniques to extract the information from the huge size datasets. DDOS attacks can cause severe damage to ISPs and online services, especially for small and medium sized organizations who lack sufficient resources to withstand a high volume of DDoS traffic. In this paper, introducing a novel approach for DDoS mitigation using collaborative networks and Network Function Virtualization (NFV) technology.

NFV is an emerging technology where network functions are implemented and provided in software, which runs on commodity hardware. The usage of NFV technology makes device upgrading and create fast and low cost which brings a great opportunity for DDoS defence. Designing a fair resource allocation method which provides an effective collaborative DDoS defence. The experimental results demonstrate that the proposed solution can effectively reduce the DDoS attack flow to the targeted server, and the resource allocation is fare and provides incentive for domains to maximally help other domains in need. The contributions include: 1) proposes a collaborative DDoS defence network based on NFV technology. 2) Proposing a dynamicre source allocation mechanism for domains so that the system is fair, efficient, and incentive-compatible. 3)Verify that the proposed solution is effective to DDoS defence. Implementing the NVF security level and firewall security efficiency so the security may be highly implemented. The process gives the complete SYN flood attacks and DOS attacks details and attack information's. The data node transmission speed is high so the data packets may easily pass from source to the destination. The NVF efficiency is high. Naïve Bayes algorithm‘s efficiency gives better accuracy.

II. LITERATURE SURVEY

[1]Aditya Phadke, Mohit Kulkarni, Pranav Bhawalkar and Rashmi Bhattad proposed, “A Review of Machine Learning Methodologies for Network Intrusion Detection “in 2019. The accuracies for Network Intrusion Detection could be improved by optimized feature selection, optimizing learning algorithms by creating multiple weak classifiers for determining whether the network access is malicious, normalization of data and optimizing the neural network design by modifying the architecture of the neural network and using regularization to prevent overfitting.

[2]Bahman Rashidi and Carol Fung proposed, “A Collaborative DDoS Defence Using Network

Function Virtualization “ in 2016, describes CoFence, a collaborative network to defend against DDoS attacks based on network virtualization technology, where domain networks under DDoS attack can redirect excessive traffic to other collaborating domains for filtering.

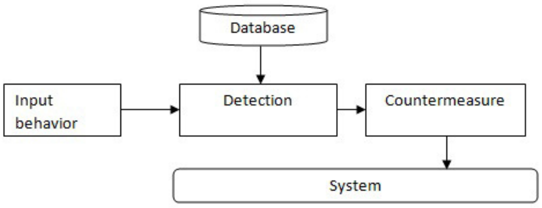
[3]Iman Sharafaldin, Arash Habibi Lashkari and Ali A. Ghorbani proposed, “Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization” in 2018.It extracts the 80 traffic features from the dataset and clarify the best short feature set to detect each attack family using Random Forest Regressor algorithm. In the future, can increase number of PCs as well as conducting more up to date attacks.

[4]JayshreeJha and Leena Ragha proposed, “Intrusion Detection System using Support Vector Machine” in 2013.It provides a review on current trends in intrusion detection using SVM together with a study on technologies implemented by some researchers in this research area. Second it proposes a novel approach to select best feature for detecting intrusion. The proposed approach is based on hybrid approach which combines filter and wrapper models for selecting relevant features. This reduced dataset will increase the performance and detection accuracy of SVM based detection model. Moreover the training and testing time will also be reduced with reduced set of feature.

[5]K. Swathi , D. Sree Lakshmi proposed, “Network Intrusion Detection Using Fast k-Nearest Neighbour Classifier” in 2014. It proposed Fast KNN Classifier algorithm for intrusion detection on large, mixed data set. Analysis of result gives a better prediction of result for different data set in KDD, but also suffered problem in Alarm generation. The processing speed of this algorithm is shown in terms of the number of iterations, and is compared with general KNN Classifier. Rough Set Theory and the Support Vector Machine is used as a tool to enhance the accuracy of the present intrusion detection algorithms.

III. INTRUSION DETECTIONSYSTEM

An IDS typically reports any policy violations or security breaches. An intrusion detection system has a static database of identified malicious behaviour. The input (i.e. the network traffic or system behaviour) is compared with the entries from this database. If the input is malicious, the severity of the threat is detected and a proper countermeasure is used. The countermeasures range from simple notifications to blocking the activity which is suspected to be a threat. The most prevalent types of IDS are Host-based and Network based.



A. Functionalities of IDS:

- User and system activities is monitored and analysed.
- Analysing system configurations and vulnerabilities.
- Assessing system and file integrity
- Ability of recognizing the patterns of attacks.
- Analysis of abnormal activity patterns.
- Tracking user policy violations.

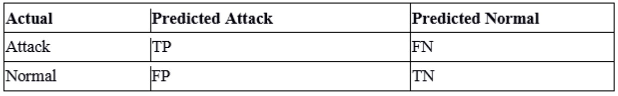
B. Evaluation Metrics for Intrusion Detection Systems:

To evaluate Intrusion Detection Systems (IDSs) for their efficiency and effectiveness various features of the IDSs can be considered, like performance and correctness to usability. The evaluation confusion matrix is used to represent classification results of the IDS. Following are the factors for the measurement of IDS;

- True Positive: A valid intrusion which triggers an IDS to produce an alarm (TP)
- False Positive: An event signalling of IDS gives an alarm when no intrusion has taken place (FP).
- False Negative: IDS is not able to detect an actual intrusion (FN).
- True Negative: When no attack has taken place places on an IDS based on past performance and analysis

To help determine its ability to effectively identify an attack Confusion (Evaluation) matrix Confusion matrix is an evaluation matrix that represents result of classification. It represents true and false classification results. The followings are the possibilities to classify events and depicted in Table 1: Confusion matrix

Table 1: Confusion matrix



Actual	Predicted Attack	Predicted Normal
Attack	TP	FN
Normal	FP	TN

Metrics from confusion matrix Different performance metrics are defined in terms of the confusion matrix variables. These metrics generate some numeric values that are easily comparable.

- Classification rate (CR)
- Detection rate (DR)
- False positive rate (FPR)
- Precision rate (PR)

C. Dataset for Intrusion Detection System:

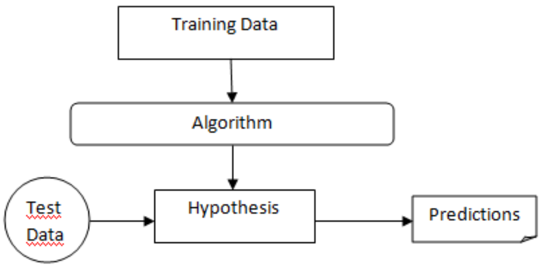
DARPA (Lincoln Laboratory): Dataset was built for network security analysis purposes. DARPA contains tasks like send and receive files using FTP, send and receive email using SMTP and POP3, browse websites, log into remote computers using Telnet and perform work, IRC messages are been sent and received, and the router is been monitored remotely using SNMP.

IV. NETWORKINTRUSIONDETECTION SYSTEM

A Network Intrusion Detection System (NIDS) is used to keep track of and provide analysis of internet traffic on the subnet. A NIDS reads all incoming data and looks out for suspicious behaviour. The system reacts to such behaviour based on the seriousness of the threat.

V. MACHINE LEARNING

Machine learning is a class of algorithms that allows software applications to become more precise in estimating outcomes without being explicitly programmed. The algorithm applied to any data is jointly called a model. A machine learning algorithm learns from experience 'E' concerning some class of tasks 'T' and performance measure 'P' if it Performance at task 'T' enhances with experience 'E'



VI. COLLABORATIVEDDOSDEFENCE

Providing a platform for domain networks helps to enhance resistance against large-scale DDoS attacks. With the help of network function virtualization (NFV) technology, each NFV-enabled domain network can contribute their spare network resource to help other domains in the network when needed.

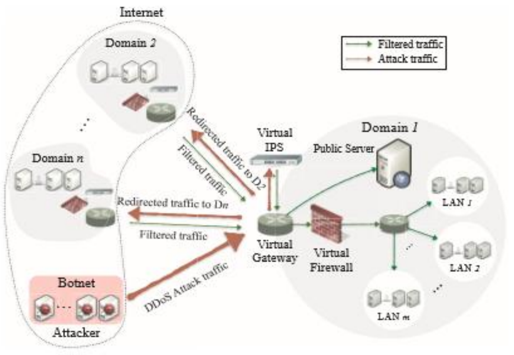


Figure 1 CoFence Diagram

In its network, the virtual gateway and a virtual IPS is been contained by each NFV-enabled domain. Virtual IPS detects and filters the DDoS attacks.

Due to the flexibility of NFV, a virtual IPS can be created and its capacity can be configured dynamically based on need. When a domains joins CoFence, the domain can choose whether to share its IPS with other trusted domains or not, and configure the maximum external traffic it is willing to handle for other domains. In a CoFence network trust is already established. Service agreement process addresses the trust.

Figure 1 shows a case study of CoFence. When the attacker launches a DDoS attack against the public server in domain 1 and the attack traffic volume exceed the maximum capacity of the local IPS,for filtering some incoming traffic may be redirected to its collaborator domains.

The SYN flood remotely filtered traffic is only forwarded back to domain 1. To be able to collaborate with other domains, a virtual IPS should contain the following functions: 1) Communication Component: This is used to communicate with other domains in the network. The communication in the collaboration network can be divided into three types: (a) request for help and offer to help. (b) Request to add as neighbours and respond to the neighbour adding request. (c) Request remove as neighbours and respond to the neighbour removal request. CoFence is set up as a “separate” defence network and usual traffic is transferred within another network. Ensuring that the requesting help is possible in case of the network link is saturated. 2) Resource Allocation:A domain needs to decide how much of its spare resource it should offer, after receiving a helping request from its neighbour. The decision problem becomes non-trivial when there are multiple requesters at the same time.

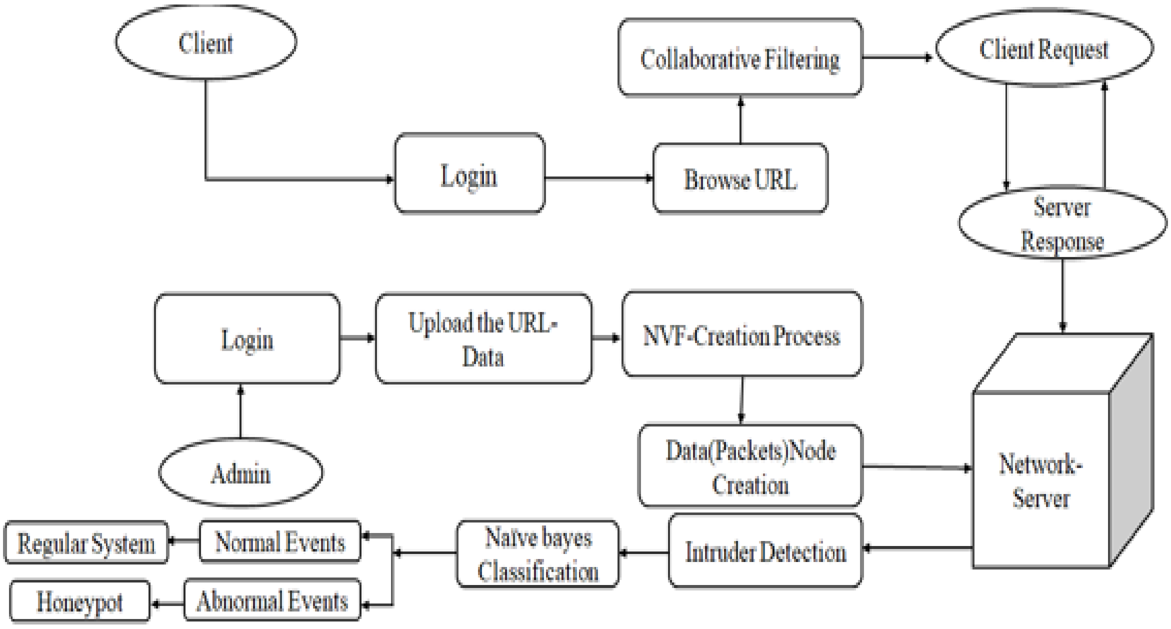
A resource allocation component computes the optimal way for the resource allocation decision.

Several design goals include how to make the resource usage efficient, fair to new neighbours, and incentive compatible to encourage more generous sharing. The focus is to design a resource allocation mechanism to meet the above goals. It is worth’s that CoFence is a distributed model and it can be applied to networks with different scales.

VII. DATASET COMPARISION WITH ATTACKS

Dataset	Browser Attack	Brute Force	DOS	SCAN	Back door	DNS	Other Attacks
DARPA	y	Y	Y	Y	N	N	Y
DEFCON	N	N	-	Y	N	N	Y
CAIDA	N	N	Y	Y	N	Y	Y
CDX	N	N	Y	Y	N	Y	-
TWENTE	N	Y	Y	Y	N	Y	Y
UMASS	N	N	-	Y	N	N	Y

VIII. SYSTEM ARCHITECTURE



IX. CONCLUSION

Proposing CoFenceto defend against DDoS attacks based on network virtualization filtering. Focusing the resource allocation mechanism that determines how much resource one domain should offer to the requesters so that the resource is distributed efficiently, fairly, and with incentives.Navies Bayes algorithm classifies the data with the better accuracy of upto 95%.In future work, can implement this concept with the neural network algorithm based on deep learning to get higher accuracy of upto 98% on classification of data

X. REFERENCE

[1] Aditya Phadke et.al proposed, "*A Review of Machine Learning Methodologies for Network Intrusion Detection* "in 2019.

[2] Bahman Rashidi and Carol Fung proposed, "*A Collaborative DDoS Defence Using Network Function Virtualization* " in 2016

[3] Iman Sharafaldin, Arash Habibi Lashkari and Ali A. Ghorbani proposed, "*Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization*" in 2018

[4] Jayshree Jha and Leena Ragh proposed, "*Intrusion Detection System using Support Vector Machine*" in 2013.

[5] K. Swathi , D. Sree Lakshmi proposed, *"Network Intrusion Detection Using Fast k-Nearest Neighbour Classifier"* in 2014.

[6] P. Ransack, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "*Group lens: An open architecture for collaborative filtering of Netnews,"* in *Proc CSCW*, 1994, pp. 175–186.

[7] Md Nasimuzzaman Chowdhury & Ken Ferens, Mike Ferens (2016). "*Network Intrusion Detection Using Machine Learning*".2016 Int'l Conf. Security and Management, SAM'16.

[8] Eslamnezhad, Mohsen & Varjani, A. (2014). "*Intrusion detection based on Min Max K-means clustering*". 2014 7th International Symposiumon Telecommunications, IST 2014. 804-808.

[9] http://www.statsoft.com/textbook/support-vector-machines/.

[10] C.J.Fung and B. McCormick. Vguard: A distributed denial of service attack mitigation method using network function virtualization. In Network and Service Management (CNSM), 2015 11th International Conference on, pages 64–70. IEEE, 2015.

[11] Arborddos detection and protection. http://security.arbornetworks.com/protection/?gclid=CNrToYHYqM0CFRY7gQodjvcN_w.

[12] Atlas q2 2015 update. http://www.slideshare.net/Arbor Networks/atlas-q2-2015final.

[13] Biggest internet attack in history threatens critical systems. http://www.ibtimes.co.uk/biggest-internet-attack-history-threatens-critical-infrastructure-450969.

[14] Ghosal, Amrita & Halder, Subir. (2017)."A survey on energy efficient intrusion detection in wireless sensor networks". Journal of Ambient Intelligenceand Smart Environments. 9. 239-261. 10.3233/AIS-170426.

[15] Arthur, David & Vassilvitskii, Sergei. (2007). "K-Means++: The Advantages of Careful Seeding". Proc. of the Annu. ACM-SIAM Symp. on Discrete Algorithms.