Real-time Visualization and Classification of DDoS Attack using Supervised Learning Algorithms

Subbulakshmi T[1], Arun Santhosh R A[2], Mohith G K[3], Suganya R[4], Girish Subramanian[5],Shubh K Patel[6] , Baraiya Manit Rameshkumar[7]

research.subbulakshmi@gmail.com[1], arunsanthosh.ra2021@vitstudent.ac.in[2], mohith.gk2022@vitstudent.ac.in[3]

School of Computer Science and Engineering,Vellore Institute of Technology,Chennai[1,2,3]

Abstract. : With increased reliance on the Internet for various services, Distributed Denial of Service (DDoS) attacks are a major concern for organizations. DDoS attacks causes significant damage to the target system, leading to service downtime, data loss, and financial losses. To mitigate the impact of DDoS attacks, effective detection mechanisms are necessary. In this paper, a machine learning-based approach for DDoS attack detection is proposed. NSL-KDD Dataset of network traffic data containing both normal and attack traffic were used and various visualization techniques and machine learning models, including Random Forest Classifier, K- Nearest Neighbours Classifier and Logistic Regression were applied. The performance of these models was evaluated using cross-validation and their accuracies were measured. The experimental results show that Random Forest Classifier outperforms other models with an accuracy of 99.33% before and after applying post pruning method. The experiment also focused on the false positives and false negatives and the implications of the results were discussed.

Keywords: DDoS attacks, Machine learning, Intrusion detection, Random Forest, Logistic Regression, K-nearest neighbors, Confusion matrix.

1 Introduction

In a dynamic landscape marked by continually advancing threats, the most recent DDoS attack, employing a fusion of ACK, PUSH, RESET, and SYN flooding attack vectors at a remarkable rate of 55.1 million [1] packets per second, serves as a testament to the evolving landscape of cybersecurity challenges. In recent years, the threat of DDoS attack is a major concern for organizations and governments around the world. DDoS attacks involve flooding a targeted server or network with requests or traffic, effectively rendering it unusable. These attacks are often launched by cybercriminals, hacktivists, or nation-state actors with various motives, such as extortion, political or social activism, or disruption of services.

DDoS attacks can have serious consequences, including loss of revenue, reputation damage, and even service disruptions that can affect the operation of critical infrastructure. In recent years, DDoS attacks have become increasingly sophisticated, making them more difficult to

IACIDS 2023, November 23-25, Lavasa, India

DOI 10.4108/eai.23-11-2023.2343234

detect and defend against. Machine learning enables computer systems to learn and ameliorate from experience. To combat DDoS attacks, many organizations use a variety of defense mechanisms, including firewalls, intrusion detection/prevention systems, and content delivery networks. However, these mechanisms are often not enough to prevent or mitigate the impact of a large-scale DDoS attack. This research paper uses machine learning (ML) models to detect DDoS attacks. Machine Learning models analyzes large amounts of data and recognize past trends that may be indicative of a DDoS attack. By using Machine Learning models to detect DDoS attacks, organizations can better protect their networks and online services. The goal of this research paper is to evaluate the effectiveness of Machine Learning models in detecting DDoS attacks. To achieve this goal, an experimental study using a public dataset of network traffic was conducted. Several Machine Learning models using this dataset were trained, tested and evaluated their performance in terms of accuracy. Section 2 provides related work on DDoS attack detection and the utilization of Machine Learning models in cybersecurity. Section 3 describes the dataset used in this experimental study and the pre-processing steps that was taken to prepare the data for analysis. Section 4 presents the experimental methodology, including the Machine Learning models that were used and the performance metrics we evaluated. Section 5 presents the experimental results and discusses the performance of the Machine Learning models that were tested. Finally, Section 6 wraps up the paper and discusses the possible advancements. To summarize, this research paper contributes to the field of cybersecurity by evaluating the effectiveness of Machine Learning models in detecting DDoS attacks. By identifying the Machine Learning models that perform best in detecting DDoS attacks, organizations can better protect their networks and online services from this type of cyber-attack..

2 Literature Survey

DDoS attacks are a type of cyber-attacks that brings down websites, servers, and other online services by overwhelming them with traffic. There is a growing need for more advanced and proactive detection methods, such as machine learning (ML) models. ML models can be trained on large datasets of network traffic to automatically identify patterns and abnormalities that may indicate a DDoS attack. These models can also be used to classify traffic as either normal or malicious, based on learned features.

Several studies have explored the use of ML models for DDoS attack detection. For example, In the previous study [2] on different types of DDoS attacks, and a novel approach was developed to detect them using five new features derived from heterogeneous packets. These features include the rate of entropy of IP source flow, packet size and number of unreachable packets of ICMP destination. Another analysis [3] devises new techniques for causing DDoS attacks and its mitigation along with the classification of those DDoS attacks

In [4] the analysis of five popular machine learning algorithms, the decision tree algorithm performed better than the others in terms of overall performance.

In a payload-inspection-based intrusion detection system (PI-IDS) [5] intrusion attempts were identified by comparing packet payloads with previously known attack signatures. To overcome this limitation, the paper proposed a traffic-based intrusion detection system (T-IDS) that analyzed the packet headers instead of payload. The proposed model outperformed other

machine-learning models achieving 99.984% accuracy and a training time of 21.38 seconds on a benchmark botnet dataset. The study [6] proposed a machine learning based botnet detection that was efficient in detecting the botnets. The optimal timing for the sliding process in traffic data segmentation and obtained botnet activity information was analyzed [7]. The results were successful to the extent of detection accuracy of 97.93%.

The evaluation [8] of six machine learning algorithms that checked for their ability in detecting DDoS attacks helped users diagnose potential DDoS threats and improved production.

Overall, these studies demonstrate the potential of ML models for detecting and mitigating the impact of DDoS attacks. However, there exist various hurdles and constraints that require attention. These include the need for large and diverse training datasets, the trade-off between detection accuracy and false positives, and the dynamic nature of DDoS attacks and the network environment. This study focuses to visualize and classify ML models for DDoS attack detection using a large dataset of network traffic, and to assess their performance based on the accuracy and limitations through false positives and negatives.

3 Proposed Method of Visualization and Classification of attacks

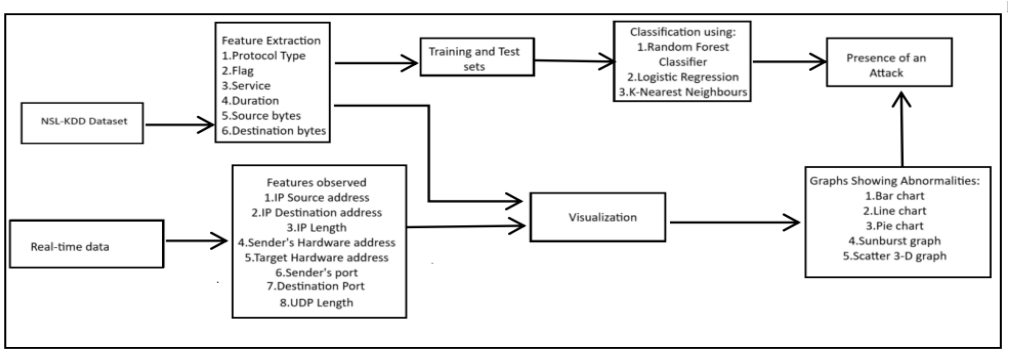


Fig.1. Proposed architecture for the implementation

3.1 Dataset Description

The NSL-KDD dataset is a variant of the KDD Cup 99 dataset, which is widely used as a benchmark dataset for intrusion detection systems (IDS). The establishment of the NSL-KDD dataset occurred to address the limitations of the KDD Cup 99 dataset, which included duplicate and irrelevant records, as well as unbalanced data.

The NSL-KDD dataset includes a pre-processed version of the KDD Cup 99 dataset, with 41 network features and a total of 148,517 records. The dataset is split into training and testing sets, with a ratio of approximately 4:1

In this research paper, a machine learning based method is proposed for detecting DDoS attacks in a network. The method involves two main steps: feature extraction and model training.

3.2 Feature Extraction

Feature extraction involves the identification and conversion of raw data into a collection of attributes suitable for input into a ML model. A set of features that was selected, which is a widely used benchmark dataset for intrusion detection research. These six features include information about the protocol type, service, flag, duration, source bytes, and destination bytes of network packets. One-hot encoding was used to convert categorical features (protocol type, service, and flag) into numeric features that can be used by the machine learning model.

3.3 Visualization

After feature extraction, the dataset was divided into training and validation subsets using the train_test_split function in scikit-learn to extract the required features. The visualization methods used in this research are helpful in distinguishing the normal and abnormal categories of the attacks present in the dataset. Before doing the classification of the dataset with ML algorithms, the visualization methods used in this research are adding an additional line of attack classification for this research which increased the possibility of emphasizing the attacks in the dataset with their categories. The scikit, seaborne, matplotlib, pandas, numpy libraries of python are used for this purpose.

3.4 Classification

The three classification algorithms used in this research are: logistic regression, random forest and k-nearest neighbors. Each algorithm was evaluated using threefold cross-validation on the basis of its performance and the best-performing model based on accuracy was selected. Once the models are selected, the top-performing models were trained using the complete training set and evaluated on the test set.

The performance of the models was analysed using standard performance metrics such as accuracy. To understand the types of errors made by the models, confusion matrix was analysed. To summarize, the proposed method involves feature extraction and model training using ML algorithms to detect DDoS attacks in a network.

This method can be applied to any network traffic dataset and can be used to develop a robust and accurate DDoS detection system.

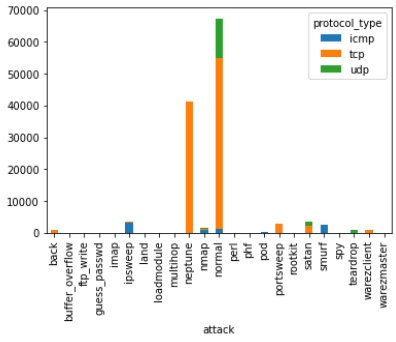


Fig.2. Attacks and Protocols targeted

The above figure shows the major protocols targeted by various DDoS attacks.

4 Implementation and Results

A few insights about the dataset are shown in this section. The bar charts below are displaying the distribution of attacks for a specific protocol. The size of each slice in the pie represents the proportion of attacks attributed to that particular protocol.

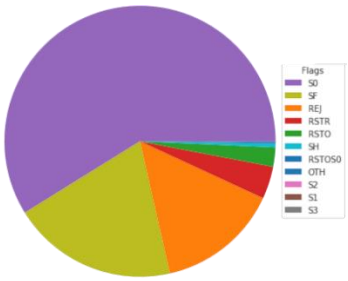
protocol_type	icmp	tcp	udp
attack			
back	0	956	0
buffer_overflow	0	30	0
ftp_write	0	8	0
guess_passwd	0	53	0
imap	0	11	0
ipsweep	3117	482	0
land	0	18	0
loadmodule	0	9	0
multihop	0	7	0
neptune	0	41214	0
nmap	981	265	247
normal	1309	53599	12434

protocol_type	icmp	tcp	udp
attack			
phf	0	4	0
pod	201	0	0
portsweep	5	2926	0
rootkit	0	7	3
satan	32	2184	1417
smurf	2646	0	0
spy	0	2	0
teardrop	0	0	892
warezclient	0	890	0
warezmaster	0	20	0
perl	0	3	0

Fig. 3. Distribution of attacks in ICMP, TCP and UDP

This can provide useful insights into the types of protocols commonly targeted by attackers and can help inform network administrators in taking appropriate measures to prevent attacks on their systems.

The two bar charts below are created to compare the distribution of flag values in normal network traffic and network traffic that contains attacks.



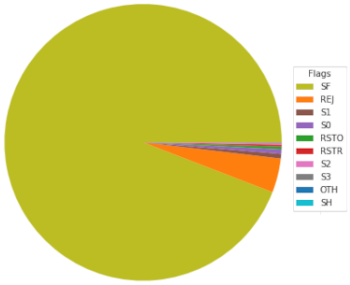
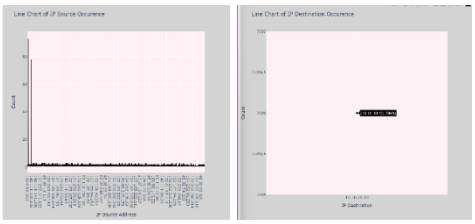


Fig.4. Flag values in a network traffic with attacks **Fig.5** Flag values in a normal network traffic

The legend shows the labels corresponding to each segment along with the percentage and absolute count values. By comparing the two charts, the differences in the flag distribution between normal and attacktraffic can be visually analysed.

5 Real-Time Visualizations



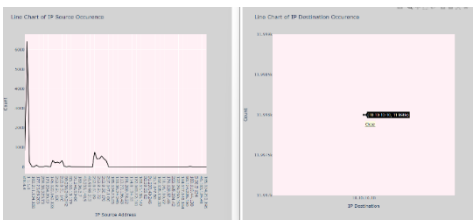


Fig.6. Line chart to show the attack traffic in TCP protocol **Fig.7**.Line chart to show the attack traffic in UDP protocol

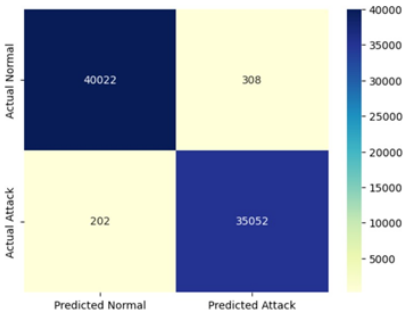
The above visualizations show the presence of a DDoS attack in TCP and UDP protocols which targets towards a single IP destination from various source IP addresses.

6 Classificatin using Machine Learning Algorithms

The performance of the three machine learning models, Random Forest Classifier, Logistic Regression, and K- Nearest Neighbors Classifier, were evaluated on the task of detecting DDoS attacks in network traffic. The models were trained on a dataset, with a split of 60% training and 40% validation data.

In the confusion matrices below, the horizontal axis shows the predicted values and the vertical axis shows the actual values. The diagonal cells represent the number of correct predictions, while the off-diagonal cells represent the incorrect predictions. The color of the cells indicates the number of predictions.

The heatmap provides insight into the performance of the model. The diagonal cells should have the highest values, indicating that the model has correctly predicted most of the data. The off-diagonal cells show the errors made by the model. The false positive cells (predicted attack but actual normal) show the instances where the model has incorrectly predicted an attack, which could result in unnecessary resources being allocated to protect against a non-existent threat. On the other hand, the false negative cells (predicted normal but actual attack) show the instances where the model has missed a real attack, which could result in serious consequences such as data theft, loss or damage.



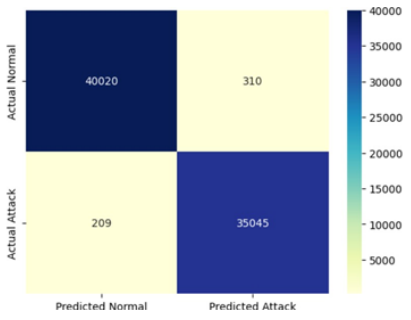


Fig. 8. Confusion matrix of the random forest classifier **Fig. 9.** Confusion matrix of the random forest classifier after applying post pruning

The confusion matrix generated from the predictions of Random Forest Classifier showed that there were 202 false negatives, indicating that 202 DDoS attacks in the validation data were incorrectly correctly identified. However, there were 308 false positives, meaning that the model incorrectly classified six instances as DDoS attacks when they were actually normal traffic.

The accuracy score of the Random Forest Classifier is 99.33 %, indicating the high accuracy at detecting DDoS attacks. The Logistic Regression model achieved a slightly lower accuracy score of only 80% while the K-Nearest Neighbors Classifier achieved the lowest accuracy score of around 98%. After Appling post pruning in Random Forest Classifier there was not significant improvements in the model, the same accuracy score of 99.33% was achieved.

Fig. 10 represents boxplot to compare the accuracy of different machine learning models used in the analysis. The boxplot provides a visual representation of the distribution of accuracy values for each model. The boxes represent the interquartile range (IQR) of the data, with the whiskers extending to the furthest data point within 1.5 times the IQR. Any data points beyond this range are plotted as individual points. The horizontal line within each box represents the median of the data. By comparing the boxplots for each model, idea of the variation in accuracy and identify any outliers or trends was observed that Random Forest algorithm has the highest accuracy in detecting DDoS attacks.

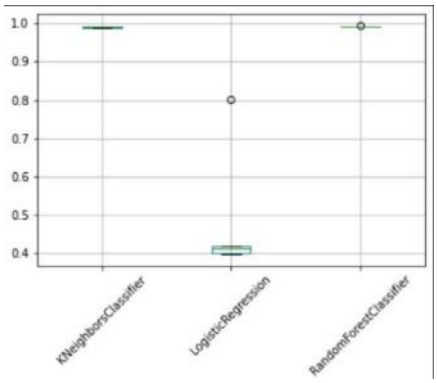


Fig.10. Boxplot to compare accuracies of different machine learning models used in the analysis

7 Uniqueness

The key observations of this paper are,

- Visualization of NSL-KDD dataset and classification of attacks in various protocols through Machine learning algorithms.

Real-time observations of data transfer and visualization of TCP and UDP attacks to check for the presence of a DDoS attack.

8 Conclusion

This research paper presents a comprehensive approach for detecting DDoS attacks using machine learningmodels. The proposed system uses the NSL-KDD dataset of network traffic to

train and evaluate the models' performance. The algorithms considered are Random Forest, Logistic Regression, and K-Nearest Neighbors. The results show that the Random Forest model outperformed the other models, achieving an accuracy of 99.3%, both before and after applying post pruning. These results demonstrate the efficacy of using machine learning models for DDoS attack detection. The proposed system's architecture, which involves pre- processing the data, feature extraction, and model training, provides a framework for future research on DDoS attack detection. Overall, this research paper contributes to the ongoing effort to improve network security by presenting an effective method for detecting DDoS attacks using machine learning algorithms. Future work may include investigating additional machine learning models or exploring the use of deeplearning models to further improve detection accuracy.

References

[1] (No date) Akamai prevents the largest ddos attack on a U.S. financial company ... Available at: https://www.akamai.com/blog/security/akamai-prevents-the-largest-ddos-attack-on-a-us-financial-company (Accessed: 15 September 2023).

[2] L. Zhou, Y. Zhu, Y. Xiang, and T. Zong, “A novel feature-based framework enabling multi-type DDoS attacks detection,” World Wide Web, Apr. 2022, doi: https://doi.org/10.1007/s11280-022-01040-3.

[3] K. S. Vanitha, S. V. UMA and S. K. Mahidhar, "Distributed denial of service: Attack techniques and mitigation," *2017 International Conference on Circuits, Controls, and Communications (CCUBE)*, Bangalore, India, 2017, pp. 226-231, doi: 10.1109/CCUBE.2017.8394146.

[4] Gulia, N., Solanki, K. and Dalal, S. (2022) ‘Comparative analysis to identify the effective machine learning method for prediction of DDOS attack’, 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) [Preprint]. doi:10.1109/icrito56286.2022.9965126.

[5] Al-Jarrah, O.Y. et al. (2016) ‘Data randomization and Cluster-based partitioning for botnet intrusion detection’, IEEE Transactions on Cybernetics, 46(8), pp. 1796–1806. doi:10.1109/tcyb.2015.2490802.

[6] S.-C. Chen, Y. -R. Chen and W. -G. Tzeng, "Effective Botnet Detection Through Neural Networks on Convolutional Features," 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), New York, NY, USA, 2018, pp. 372-378, doi: 10.1109/TrustCom/BigDataSE.2018.00062.

[7] D. P. Hostiadi and T. Ahmad, "Sliding Time Analysis in Traffic Segmentation for Botnet Activity Detection," 2022 5th International Conference on Computing and Informatics (ICCI), New Cairo, Cairo, Egypt, 2022, pp. 286-291, doi: 10.1109/ICCI54321.2022.9756077.

[8] G. Qaiser, S. Chandrasekaran, R. Chai and J. Zheng, "Classifying DDoS Attack in Industrial Internet of Services Using Machine Learning," 2023 15th International Conference on Computer and Automation Engineering (ICCAE), Sydney, Australia, 2023, pp. 546-550, doi: 10.1109/ICCAE56788.2023.10111178.