Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

A Machine Learning-Based System for Detecting and Alerting DDoS Attacks in Cloud-Based Applications

VijayRaj[1], Dr. Subramanya Bhat[2]

[1]Research Scholar, College of Computer Science and Information Science, Srinivas University, Mangalore, itsvraj87@gmail.com
[2]Research Professor, Computer Science and Information Science, Srinivas University, Mangalore

One of the most frequent network attacks is the distributed denial of service (DDoS) attack. A denial-of-service (DDoS) attack happens when a malevolent user targets a website or server with the intention of flooding it with unsolicited information. This results in a delay in services for authorized users. A Denial of Service (DoS) attack originates from a single source, while a Distributed Denial of Service (DDoS) attack originates from multiple sources, such as a botnet that remotely controls devices for malevolent purposes. A standard benchmark dataset, UNSW-NB15, is used for training and testing purposes. We tested the trained models with the parameters Accuracy, Precision, Recall, and F1-Score. Three machine learning algorithms are selected to detect DDoS attacks, and the best model is found in terms of accuracy, precision, recall, and F1-Score. Among three models we found that Random Forest is the best model by considering all parameters into account. It has produced 98.6% accuracy, 99.6% Precision, 98.3% Recall and 98.5 % F1-Score to detect DDoS attack at the earliest.
Keywords: DDoS attacks, cloud computing, machine learning, Random Forest, K-Nearest Neighbor, Navies Bayer's traffic analysis, cybersecurity.

1. Introduction

The introduction of cloud computing has completely changed the way businesses operate. which provides unparalleled scalability, flexibility, and cost-effectiveness. Both businesses and individuals rely on cloud-based applications for a range of services, including social media, productivity software, email, and storage. The heightened reliance on cloud-centric applications has engendered novel security vulnerabilities, including Distributed Denial-of-Service (DDoS) assaults. DDoS attacks involve overwhelming a cloud-based application with harmful traffic, which stops legitimate users from being able to use it. These attacks could lead

to various negative consequences including financial losses, damage to one's reputation, and exposure of sensitive data. According to a recent study, there was a 50% increase in DDoS attacks in 2022, with an average

attack size of 1.3 Gbps. The advent of cloud computing, which provides unparalleled scalability, flexibility, and cost-efficiency, has fundamentally transformed business operations. Both businesses and individuals now rely on cloud-based applications for a range of services, including social media, productivity software, email, and storage. The heightened reliance on cloud-based applications has resulted in the emergence of novel security vulnerabilities, including Distributed Denial-of-Service (DDoS) attacks. DDoS attacks are characterized by the inundation of a cloud-based application with detrimental traffic, thereby blocking authorized users from using it. The attacks can lead to various negative consequences, such as financial losses, damage to one's reputation, and exposure of sensitive data. A recent study alleges that the number of DDoS attacks in 2022 increased by 50%, with the average size of the attacks being 1.3 Gbps. Conventional methods for detecting DDoS attacks, such as threshold-based detection and rule-based systems, often do not effectively prevent sophisticated and large-scale attacks. The high number of false alarms generated by these methods can lead to unnecessary use of resources and potential disruptions in service. Furthermore, there is a possibility that they may overlook complex attack patterns, allowing harmful traffic to go undetected. Recent advancements in deep learning (DL) and machine learning (ML) have shown potential in recognizing DDoS attacks. ML/DL-based techniques can be used to analyze intricate network traffic patterns and detect small anomalies that could indicate a potential DDoS attack. Nevertheless, existing ML/DL-based DDoS detection systems have several drawbacks including significant computational burden, limited scalability, and inadequate resistance to various types of DDoS attacks. This work goals to report these confines by familiarizing a new machine learning (ML) system for noticing and informing about DDoS attacks in cloud-based applications. We assess the efficiency of our structure in noticing DDoS attacks by using a universal dataset of network traffic outlines. The paper is structured as follows. Section I provides an overview of current DDoS detection methods, highlighting their inadequacies. In Section III, our proposed ML-based system, including its architecture and implementation details, is introduced. Section IV will compare the performance of our system with current state-of-the-art DDoS detection systems using a large dataset. In Section V, this paper is ultimately concluded with a summary.

2. Related Work

We will now provide an overview of current literature discussing the detection methods for DDoS attacks in this section. We will provide a short overview of the recent trends in DDoS attack detection mechanisms.

Pande et.al [5] proposed an approach using machine learning (ML) to identify Distributed Denial of Service (DDoS) attacks. The authors utilized a blend of feature extraction and machine learning algorithms in order to identify DDoS attacks. The CICIDS2017 dataset, utilized by the authors, comprises a combination of both harmless and harmful network traffic, and implemented using three ML algorithms: Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). In their study, the researchers found that Random

Forest (RF) outperformed SVM and KNN in accurately detecting DDoS attacks, achieving an accuracy of 98.45%. dhammad, M., et.al [6] introduce a semi-supervised method for identifying DDoS attacks. The researchers used a combination of labeled and unlabeled data to train a machine learning model that can effectively detect DDoS attacks. They found that semi-supervised machine learning methods can successfully identify these attacks. Kumar et.al [7] proposed using an online machine learning approach to adapt to changing network traffic patterns and detect attacks in real-time. Typically, this approach is evaluated using standard sets of network traffic data that include both normal and malicious traffic. Priya, S. S., et.al.[8] examines and contrasts the efficiency of different machine learning (ML) and deep learning (DL) methods in identifying Distributed Denial-of-Service (DDoS) attacks. The authors' conclusion was that deep learning models frequently surpass traditional machine learning models in terms of accuracy and detection rate, particularly for complex and large datasets. In their article, Detecting and Classifying Network Traffic Flows Using Machine Learning and WEKA Doshi, R., [9] suggest using a machine learning approach with WEKA to identify and categorize various network traffic flows, including contemporary attack types such as HTTP flood and SID DoS, as well as normal traffic. The J48 classifier outperformed the other 2 classifiers based on the results provided by three classifiers. J48 achieved 98.64% accuracy. The study conducted by Dasari, K. B., & Devarakonda, N. [10] examined the effectiveness of different machine learning algorithms in identifying Botnet-based DDoS attacks. The authors' conclusion was that SVM and ANN, other algorithms such as Naïve Bayes and Decision Tree. Lima Filho, F. S. D., et.al [11], aim to systematically review the current state-of-the-art in deep learning-based DDoS attack detection in their article It examines and sorts different deep learning methods, data sets, evaluation measures, and obstacles in this area. of different approaches by methodically analyzing various deep learning architectures, datasets, and evaluation metrics. Al-Shareeda, M. A., et.al [12] have introduced a predictive integrated approach for DDoS detection and prevention in utilizes machine learning algorithms such as Jrip, J48, and k-NN on the CICIDS2017 dataset to successfully identify different types of DDoS attacks, including brute force, Heartbleed, infiltration, botnet, and port scan attacks. Their focus is specifically on preventing attacks by blocking malicious nodes that are identified as participating in them. The research findings suggest that combining detection and prevention methods can significantly help in reducing the risk of DDoS attacks. Suresh, M., & Anitha, R [13] found that by conducting effective feature engineering, particularly feature selection, DDoS attack detection accuracy can be greatly improved through the reduction of data dimensionality and the mitigation of problems such as overfitting. The suggested framework highlights the significance of using a organized method for creating features in order to develop strong and effective detection systems. This results in improved performance of supervised machine learning models in detecting DDoS attacks. Aytaç, T., et al. [14] determined in their study that the accuracy of DDoS detection with ML techniques is greatly affected by feature selection. As per their examine, the performance of classifiers can be improved by choosing relevant features, which reduces noise and complexity. In their study, Khempetch and Wuttidittachotti [15] discovered that deep learning models are effective in detecting DDoS attacks, yielding promising results. The study demonstrates the ability of deep learning to automatically learn intricate patterns from network traffic data, leading to the successful identification of various attack types. This suggests that deep learning offers a viable and potentially superior approach compared to other methods. Tuan, T. A. [16]

concludes that AI approaches show excessive potential in effectively detecting and preventing DDoS attacks. Their research demonstrates network security by improving detection accuracy and enabling proactive risk mitigation strategies. According to Bindra and Sood [17] , Multiple Linear Regression (MLR) is a useful method for detecting DDoS attacks. The inquire about conducted by them appears that MLR is competent of recognizing DDoS assaults by analyzing organize activity characteristics. Li, Q., Meng, L., Zhang, Y., & Yan, J [18] suggest that machine learning models show potential for effectively detecting DDoS attacks and identifying different types of attacks. The comparison of their performance emphasizes to progress existing detection systems.

3. Intrusion Detection Methods

There are 3 prime categories of intrusion detection methodologies: Signature-based Detection (SD), Anomaly-based Detection (AD), and Stateful Protocol Analysis (SPA). Next, we will outline the advantages and disadvantages of the three detection methods.

Signature-based Detection: Signature-based intrusion detection systems can categorize intrusions by monitoring events and recognizing patterns that match the known attack signatures [19]. An attack signature provides a detailed plan for executing an attack, including the required steps and their specific sequence. Moreover, it exclusively detects assaults for which their signatures have previously been stored in a database. The signatures need to be regularly updated to ensure efficiency [20]. Your hosts are consistently facing new threats, and as a result, there is a continuous need to update signatures in response to the frequent release of new threats. This method is only successful in combating a specific set of behaviors. They cannot effectively manage attacks from humans or worms that have self-modifying behavioral characteristics.

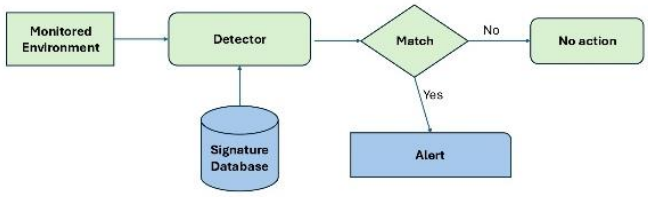

Figure 1: Signature Based Detection [21]

Anomaly-based Detection: Many researchers have been intrigued by the AD's skill in detecting new attacks. Detecting network behavior is built upon the foundation of determining it. The network's behavior aligns with the predetermined behavior. Following that, it will be either approved or else it will initiate the anomaly detection process [21]. Network administrators can prepare or learn the accepted behavior of the network according to its specifications. One of the main benefits of AD over signature-based engines is its ability to detect a new spell that does not have a predefined signature, as long as it deviates the traffic patterns.

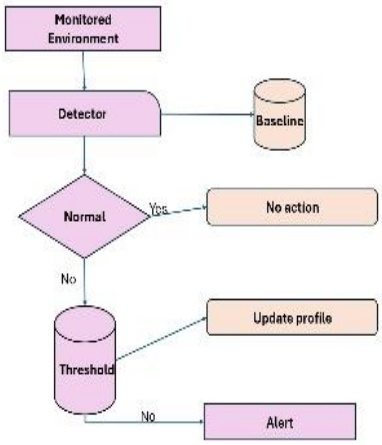

Figure 2: Anomaly Based Detection

Stateful protocol analysis: The Stateful protocol analysis methodology involves comparing the expected behavior of protocols with their actual behavior, using established profiles as a reference. The vendor designs and establishes the established protocol profiles. In contrast to the signature-based approach that only checks observed behavior against a list, Stateful protocol analysis comprehensively understands how protocols and applications are supposed to function. This comprehensive understanding and analysis require a high level of resources for the systems [22]. The integration of stateful protocol analysis with other IDPS methodologies has proven to be effective, leading to the development of hybrid methodologies. Figure 4 illustrates the overall structure of Stateful protocol analysis. The construction of this manner is the same as the signature-based methodology, except for one difference: instead of using a signature database, it utilizes a database of acceptable protocol behavior for Stateful protocol analysis [23].

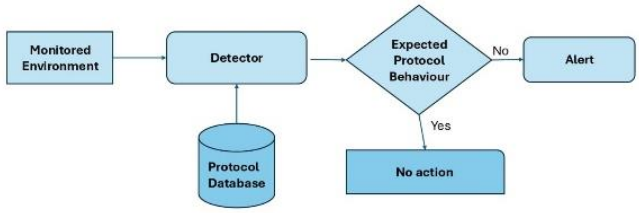

Figure 3: Stateful protocol analysis detection

4. Methodology

Employing ML methods for categorizing the requested data and implementing this model in an intelligent detection system to alert network administrators about potentially harmful

requests. The value of existing data is maximized through the utilization of machine learning techniques. As the availability of data increases, the accuracy of classification also improves. ML approaches are employed to tackle the limitations of the current systems. [24]. The objective is to create a classification model using the provided data. The model must have the capability to categorize any unfamiliar case as either a harmful (DoS) assault or a harmless inquiry. Methods such as Random Forest and K-Means Clustering can be utilized for categorization issues of this nature. The goal is to solve classification problems [25].

Dataset: This research utilizes the UNSW-NB15 dataset, a benchmark for evaluating DDoS detection systems. Collected at the University of New South Wales' Cyber Range Lab, this dataset comprises 257,673 network traffic records, meticulously labeled as either "normal" (164,673 records) or "abnormal" (93,000 records, encompassing various DDoS attacks)

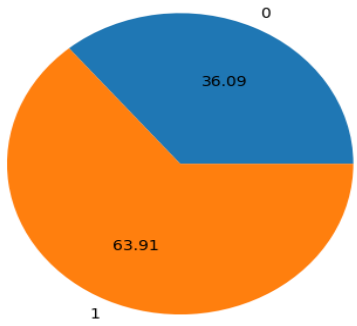

Figure 4: Percentage Report of Normal and DDoS attacks

A. Preprocessing

When machine learning begins experiments, data pre-processing is conducted to improve the model's capability to be effective in various situations. The initial stage of Pre-processing involves removing socket features, and then eliminating records that contain missing values. Employ a data split of 80:20, allocating 80% of the labeled dataset for prototypical training and reserving the remaining 20% for independent testing. Prior to training, regularize the training data to enhance the classification routine of the models."

B. Classification Algorithms

Machine learning procedures are applied to address a variety of network security issues due to their significant benefits and advanced capabilities, which include high accuracy, continuous enhancement, and the capacity to handle large volumes of data. The research employed various classification algorithms, including Random Forest(RF), K Nearest Neighbors(K-NN), and Naive Bayes(NB), to identify DDoS attacks within the dataset. classifier is user-friendly, has quick computation, and is well-suited for analyzing large datasets with many variables.

5. Findings & Analysis

Dissimilar performance measures are used to evaluate how well classification algorithms are able to detect DDoS attacks and accurately differentiate between attacks and normal activities.

Accuracy

The accuracy of a classifier indicates how accurately it assigns class labels. The logistic regression (LR), Gradient Boost (GB), and Naive Bayes (NB) models all attained a classification accuracy of 99.58%, which was the highest among all the models. The accuracy of the K-Nearest Neighbor(K-NN) model is not good. Figure 1 displays a figure depicting the accuracy standards of the arrangement algorithms for detection DDoS attacks.

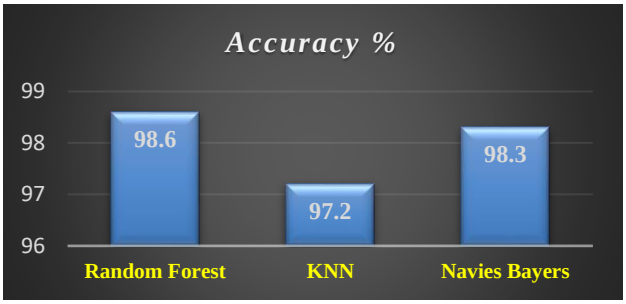

Figure 5: Accuracy report of RF, K-NN & NB

Table 1. Classifier precision, recall, and F-Score values for detecting DDoS attacks

	Random Forest	K-NN	Navies Bayers
Precision	99.6	98.7	98.8
Recall	98.3	97.8	98.1
F1-Score	98.5	97.7	97.3

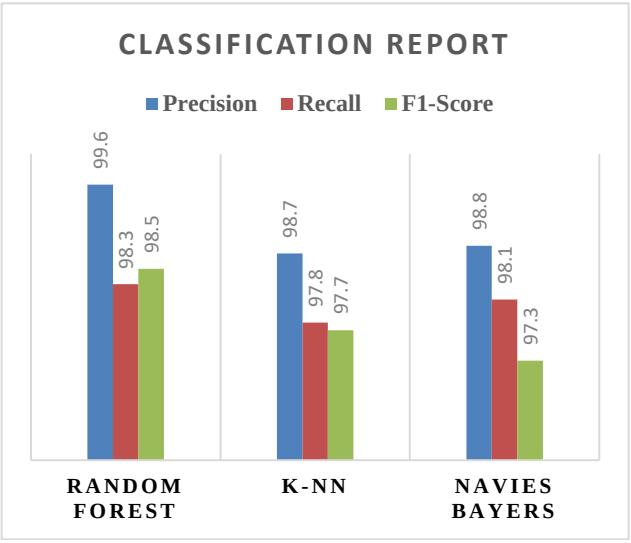

Figure 2: Classification Report

6. CONCLUSION

DDoS attacks pose a significant threat to the security of networks. They disrupt the operation of online applications and network services and resources, causing them to come to a stop. In the midst of a DDoS attack, valid users were left waiting for services as the system was preoccupied with handling the fraudulent requests from bots. The dataset UNSW-NB15 was utilized in this research. Comprising a total of 257,673 records, this dataset consists of 164,673 normal entries (0) and 93,000 abnormal entries (attacks) (1). Three various classifiers, Random Forest, Knn & Random Forest, and Naïve Bayes, were employed to categorize the attacks in the dataset. After analyzing the results from three classifiers and detected that the Random Forest classifier performed better than the other two classifiers. Random forest achieved an accuracy of 98.64%, while the K-nn algorithm and Navies Bayer's achieved 97.2% and 98.3% accuracy, respectively.

References

[1] Hoque, N., Bhattacharyya, D. K., & Kalita, J. K. (2015). Botnet in DDoS attacks: trends and challenges. IEEE Communications Surveys & Tutorials, 17(4), 2242-2270.

[2] Nazario, J. (2008). DDoS attack evolution. Network Security, 2008(7), 7-10.

[3] Douligeris, C., & Mitrokotsa, A. (2003, December). DDoS attacks and defense mechanisms: a classification. In Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology (IEEE Cat. No. 03EX795) (pp. 190-193). IEEE.

[4] Srivastava, A., Gupta, B. B., Tyagi, A., Sharma, A., & Mishra, A. (2011, September). A recent survey on DDoS attacks and defense mechanisms. In International Conference on Parallel Distributed Computing Technologies and Applications (pp. 570-580). Berlin, Heidelberg: Springer Berlin Heidelberg.

[5] Pande, S., Khamparia, A., Gupta, D., & Thanh, D. N. (2021). DDOS detection using machine learning technique. In Recent Studies on Computational Intelligence: Doctoral Symposium on Computational Intelligence (DoSCI 2020) (pp. 59-68). Springer Singapore.

[6] Idhammad, M., Afdel, K., & Belouch, M. (2018). Semi-supervised machine learning approach for DDoS detection. Applied Intelligence, 48, 3193-3208

[7] Kumar, D., Pateriya, R. K., Gupta, R. K., Dehalwar, V., & Sharma, A. (2023). DDoS detection using deep learning. Procedia Computer Science, 218, 2420-2429.

[8] Priya, S. S., Sivaram, M., Yuvaraj, D., & Jayanthiladevi, A. (2020, March). Machine learning based DDoS detection. In 2020 International Conference on Emerging Smart Computing and Informatics (ESCI) (pp. 234-237). IEEE.

[9] Doshi, R., Apthorpe, N., & Feamster, N. (2018, May). "Machine learning ddos detection for consumer internet of things devices". In 2018 IEEE Security and Privacy Workshops (SPW) (pp. 29-35). IEEE.

[10] Dasari, K. B., & Devarakonda, N. (2021). Detection of Different DDoS Attacks Using Machine Learning Classification Algorithms. Ingénierie des Systèmes d Inf., 26(5), 461-468.

[11] Lima Filho, F. S. D., Silveira, F. A., de Medeiros Brito Junior, A., Vargas-Solar, G., & Silveira, L. F. (2019). Smart detection: an online approach for DoS/DDoS attack detection using machine learning. Security and Communication Networks, 2019(1), 1574749

[12] Al-Shareeda, M. A., Manickam, S., & Saare, M. A. (2023). DDoS attacks detection using machine learning and deep learning techniques: Analysis and comparison. Bulletin of Electrical Engineering and Informatics, 12(2), 930-939.

[13] Suresh, M., & Anitha, R. (2011). Evaluating machine learning algorithms for detecting DDoS attacks. In Advances in Network Security and Applications: 4th International Conference, CNSA

2011, Chennai, India, July 15-17, 2011 4 (pp. 441-452). Springer Berlin Heidelberg.
[14] Aytaç, T., AYDIN, M., & ZAİM, A. (2020). Detection DDOS attacks using machine learning methods. Electrica, 20(2).
[15] Khempetch, T., & Wuttidittachotti, P. (2021). DDoS attack detection using deep learning. IAES International Journal of Artificial Intelligence, 10(2), 382.
[16] Tuan, T. A., Long, H. V., Son, L. H., Kumar, R., Priyadarshini, I., & Son, N. T. K. (2020). Performance evaluation of Botnet DDoS attack detection using machine learning. Evolutionary Intelligence, 13(2), 283-294.
[17] Bindra, N., & Sood, M. (2019). Detecting DDoS attacks using machine learning techniques and contemporary intrusion detection dataset. Automatic Control and Computer Sciences, 53(5), 419-428.
[18] Li, Q., Meng, L., Zhang, Y., & Yan, J. (2019). DDoS attacks detection using machine learning algorithms. In Digital TV and Multimedia Communication: 15th International Forum, IFTC 2018, Shanghai, China, September 20–21, 2018, Revised Selected Papers 15 (pp. 205-216). Springer Singapore.
[20] Srivastava, A., Gupta, B. B., Tyagi, A., Sharma, A., & Mishra, A. (2011, September). A recent survey on DDoS attacks and defense mechanisms. In International Conference on Parallel Distributed Computing Technologies and Applications (pp. 570-580). Berlin, Heidelberg: Springer Berlin Heidelberg.
[21] Santanna, J. J., van Rijswijk-Deij, R., Hofstede, R., Sperotto, A., Wierbosch, M., Granville, L. Z., & Pras, A. (2015, May). Booters—An analysis of DDoS-as-a-service attacks. In 2015 IFIP/IEEE International Symposium on Integrated Network Management (IM) (pp. 243-251). IEEE.
[22] Somani, G., Gaur, M. S., Sanghi, D., Conti, M., & Buyya, R. (2017). DDoS attacks in cloud computing: Issues, taxonomy, and future directions. Computer communications, 107, 30-48.
[23] Rajapraveen, K. N., & Pasumarty, R. (2021, December). A machine learning approach for DDoS prevention system in cloud computing environment. In 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS) (pp. 1-6). IEEE.
[24] Sanjalawe, Y., & Althobaiti, T. (2023). DDoS Attack Detection in Cloud Computing Based on Ensemble Feature Selection and Deep Learning. Computers, Materials & Continua, 75(2).
[25] Makkawi, A. M., & Yousif, A. (2021, February). Machine learning for cloud DDoS attack detection: a systematic review. In 2020 International conference on computer, control, electrical, and electronics engineering (ICCCEEE) (pp. 1-9). IEEE.