Feature-Driven Ensemble Learning for Effective DDoS Detection in Software-Defined Networks

C Sivakumar
Dept of CSSE
Mohan Babu University
(ErstwhileSree Vidyanikethan
Engineering College)
Tirupati, India
sivakumar.c@vidyanikethan.edu

Nambeti Mahathi
Student Department of Cyber Security
Mohan Babu University
(ErstwhileSree Vidyanikethan
Engineering College)
Tirupati, India
mahathidlms@gmail.com

K E Nithish Kumar
Student Department of Cyber Security
Mohan Babu University
(ErstwhileSree Vidyanikethan
Engineering College)
Tirupati, India
kenithishkumar1909@gmail.com

Kogara Malani
Student Department of Cyber Security
Mohan Babu University
(ErstwhileSree Vidyanikethan
Engineering College)
Tirupati, India
malanimalani606@gmail.com

Yalamanchili Venkata Subhash
Student Department of Cyber Security
Mohan Babu University
(ErstwhileSree Vidyanikethan
Engineering College)
Tirupati, India
yalamanchilivenkatasubhash@gmail.com

Abstract: DDoS attacks are serious threats to network security and pose specific challenges in Software-Defined Networks (SDN) due to their centralized control, which comes with its own set of vulnerabilities. State-of-the-art methods have limitations in successfully detecting low-rate and agile multi vector DDoS attacks. In this work, we will introduce a machine learning based system to recognize the DDoS attacks for SDN environments. We use an ensemble approach, combining multiple classifiers to improve detection performance, and we train on a common dataset for DDoS detection in SDNs. The most informative features in the data are highlighted using the feature transformation technique to be used in the model to enhance the model efficiency in detection of the outliers or the electrocardiograms that are deemed risky or dis-functional. Moreover, this research employs Principal Component Analysis (PCA) to achieve dimensionality reduction, thereby enhancing algorithm efficiency by decreasing the number of features while retaining the most important information. This ensemble method takes the strength of the various classifiers yielding a more robust and precise detection system The simulation is conducted with the popular tools in the SDN community, Mininet and Ryu, so the model fully utilizes the SDNs' real-life application. The model yields higher precision and recall scores than conventional approaches. Although the attention mechanism is primarily used in this work for effective detection of DDoS attacks, this system provides a general-purpose scalable solution for detecting DDoS attacks in SDN, with the potential for real-time monitoring, paving the way for future work on automated detection and response against different threats in SDN environments.

Keywords: Mininet, Ryu, Software-Defined Network(SDN) , DDoS Detection, Dynamic feature Selection,Ensemble Learning, Network Security, Machine Learning, Dimensionality Reduction .

I. Introduction

With the ever-changing networking environment, Software-Defined Networks have emerged as a paradigm shift that reshapes how we design, manage, and operate networks.The Control and Data planes manage data packet transport through the network by using routers and switches. This decoupling allows for centralized management of network resources through a logically centralized SDN controller, facilitating scalability, flexibility, and agility in network management that has never been achieved before. With a growing number of organizations migrating towards software-defined networking to address dynamic demands and architectural challenges, securing those networks is more important than ever.

One of the most significant challenges confronting SDN environments is the ubiquitous threat of Distributed Denial of Service attacks. Attacks that target the availability of a network by overwhelming it with a flood of malicious traffic are commonly known as denial-of-service attacks. Furthermore, low-rate DDoS attacks that use stealth approaches make it even harder to detect attacks, as these attacks escape most threshold-based detectors. Such attacks are exemplified by techniques such as Slowloris and RUDY, which target failures in web servers and instead burn resources without using a large amount of bandwidth. Motivated as such, S-DDoS is emerging to solve these limitations of a static and conventional DDoS detection system on recent advancements at incorporating Machine learning into an SDN environment.

However, as the DDoS landscape continues to evolve, the number of features capable of being exploited to train a model continues to be less available, cutting out some of the adaptability and efficacy of the models against more advanced attack types. In this paper, we propose an ensemble machine learning-based DDoS attack detection approach in an SDN environment. Using a classic detection method, Bernoulli Naive Bayes, Passive-Aggressive, and Multi-Layer Perceptron classifiers as input to a Stacking Classifier, the proposed system overcomes traditional detection limitations. The model is trained on the InSDN dataset and uses feature transformation techniques such as target encoding to represent the data effectively. When it comes to improving the results of machine learning models, the task of feature selection is of great importance. Mutual information has been one of the potential techniques for performing feature selection. the goal of mutual information is to prioritize and concentrate on the most relevant features, such that the information lost in the model is minimized while the accuracy of the model in predicting the target variable is enlarged. Apart from feature extraction, maximally reducing the dimensions of the data is also very relevant for increasing the efficiency of the computations, especially if one is working with multi-dimensional data sets. PCA is one of the most popular dimensionality reduction techniques.

PCA first identifies a set of linear combinations of the original variables, called principal components. Then it projects the data onto a lower dimensional space spanned by these principal components. These components contain most of the variability in the data and therefore will yield the most important aspects of the data while minimizing excess noise. The mutual information for feature selection together with PCA for dimensionality reduction is integrated into the proposed system which utilizes a reduced feature space that is efficient and relevant. Not only does this strategy improve computational resources needed by cutting down on the volume of input data but it also helps to reduce overfitting by keeping only the necessary features. Such improvements and optimizations are important in building a strong and extensible machine learning framework for DDoS attack detection in Software-Defined Networks. We conduct an evaluation of the proposed system in a simulated software defined network environment making use of Mininet for the network construction and Ryu to implement the SDN controller thus enabling us to provide the approach with a real-world testing platform. Results of this work provide a stepping-stone towards a scalable and robust framework for real-time DDoS detection in SDN networks.

II. Related Work

This chapter discusses various important methodologies for DDoS attack detection and prevention in SDN environments and highlights the role of ML in improving network security. .

Ribeiro et al. discussed a flexible architecture that combines MTD with SDN, which can be used against DDoS attacks, where the system redirects attack traffic to a controlled server, thereby ensuring that legitimate users maintain uninterrupted service while isolating the attack paths. Ensemble modeling is used in the process of classifying traffic, and different algorithms such as Gaussian Naive Bayes (GNB), Support Vector Machine (SVM), Random Forest (RF), and Multi-layer Perceptron (MLP) are used in it. SDN with MTD can also enhance adaptability and robustness. The use of ML-driven DDoS will increase the accuracy. Advanced proactive defense mechanisms also have the ability to anticipate prospective threats. However, the architecture is highly dependent on specific network conditions and the availability of secondary servers, which can limit its practical deployment in dynamic or resource-constrained environments.

Tonka et al. present a machine learning framework using most Neighborhood Component Analysis (NCA) for feature selection to optimize DDoS detection in SDN. Their ensemble learning approach achieves exceptional accuracy, with Decision Trees (DT) achieving a 100% classification success rate. Despite this success, the reliance on predefined feature sets poses a limitation, reducing the model's effectiveness against novel or evolving attack patterns that deviate from known characteristics. This static approach makes it less adaptable to real-world scenarios where attack strategies evolve rapidly.

Deepa et al propose an ensemble learning methodology that uses combination of multiple algorithms. Different algorithms used in the framework are K-Nearest Neighbor (KNN), Naive Bayes, SVM, and SOM, each of which identifies anomalous traffic behavior in SDN controllers. The ensemble method reduces false alarms, increases the number of detection rates, and improves accuracy as compared with single-algorithm-based solution. This approach significantly enhances the detection and mitigation of DDoS attacks, especially in dynamic SDN environments by addressing the limitations of standalone models.

Jess et al. Introduce a modular architecture to train an intrusion detection system (IDS) using six varied machine learning models. They achieved a 95% detection rate, proving this system is effective in spotting threats. However, on large-scale networks, several drawbacks have been seen with implementation, such as increased controller overhead and decreased response efficiency, which poses a need to optimize the scalability and resources.

Khamkar et al. focus on mitigating Low-Rate Distributed Denial-of-Service (LDDoS) attacks through an SVM-based framework, achieving a high accuracy rate of 99%. While the framework is effective in identifying and mitigating LDDoS threats, it lacks robust mechanisms for feature selection and rule creation, which can impede its efficiency in broader deployment scenarios. Addressing these gaps is essential for enhancing its scalability and adaptability.

Sudar et al. developed a flow-based detection and mitigation framework, achieving an accuracy of 93% in the classification of network traffic while optimally reducing the consumption of resources. Nevertheless, the system has a very high false positive rate for particular traffic, such as ICMP, thereby limiting its usability in varied traffic environments. Despite achieving resource efficiency, more enhancement is necessary in terms of reducing false positives to achieve precision in detection.

These methods highlight the challenges of dealing with emerging and zero-day attacks. Traditional machine learning models often rely on static datasets, historical data, and predefined features, which makes them not very effective in detecting novel attack patterns or adjusting to changing tactics from malicious actors. The reliance on static data weakens the models' capability to detect zero-day threats and respond accordingly to evolving attack scenarios.

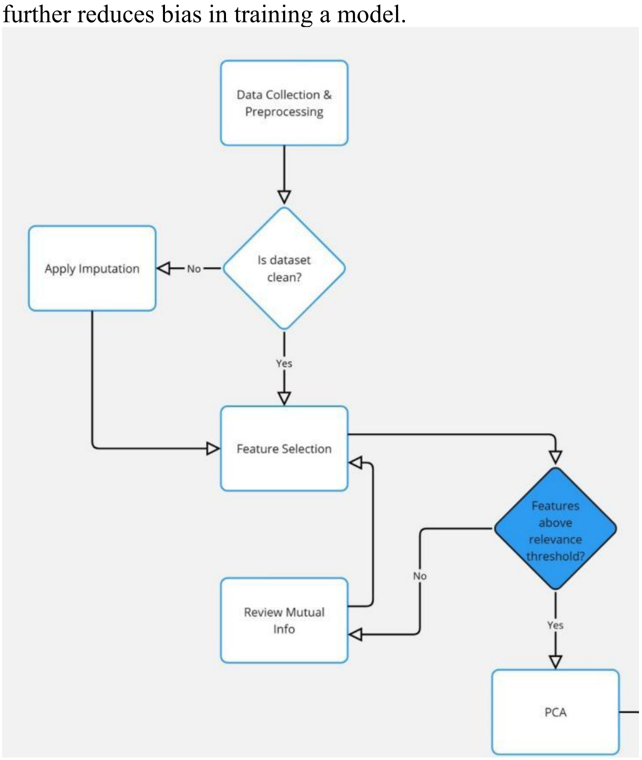
To overcome these limitations, a dynamic framework combining several innovations is proposed. The framework leverages online ensemble learning to dynamically adapt to evolving network conditions and real-time concept drift. Continuous feature selection allows the model to update the most relevant features based on real-time traffic characteristics, enhancing detection accuracy with minimal computational overhead. Designed to handle voluminous data efficiently, it ensures scalability across diverse and complex network environments. Unlike conventional methods that depend on static datasets and offline training, this approach learns from real-time streaming data, allowing for immediate responses to emerging threats. This dynamic framework of optimized resource usage and high efficacy in detection is a step forward in securing SDN environments.

III. Methodology

The methodology for designing an ensemble-based DDoS detection system for Software-Defined Networks (SDN) would be a systematic approach combining data preprocessing, point selection, model development, and evaluation in a simulated SDN terrain. The crucial stages of the methodology are outlined below

1. Data Collection and Preprocessing :

The first phase is carrying a high-quality dataset containing business data reflective of normal and vicious conditioning. For this experiment, the InSDN dataset is used because it can be applied to SDN-based environments. The gathered dataset is preprocessed to get ready for training and testing. originally, the dataset is divided into lower packets that are distributed into benign and vicious business. Missing or noisy data values are gutted by insinuation styles to ensure data thickness. To enhance model comity, categorical features are converted into numerical forms using target garbling. besides that, numerical features are formalized for invariant scaling that further reduces bias in training a model.



2. point Selection and Dimensionality Reduction :

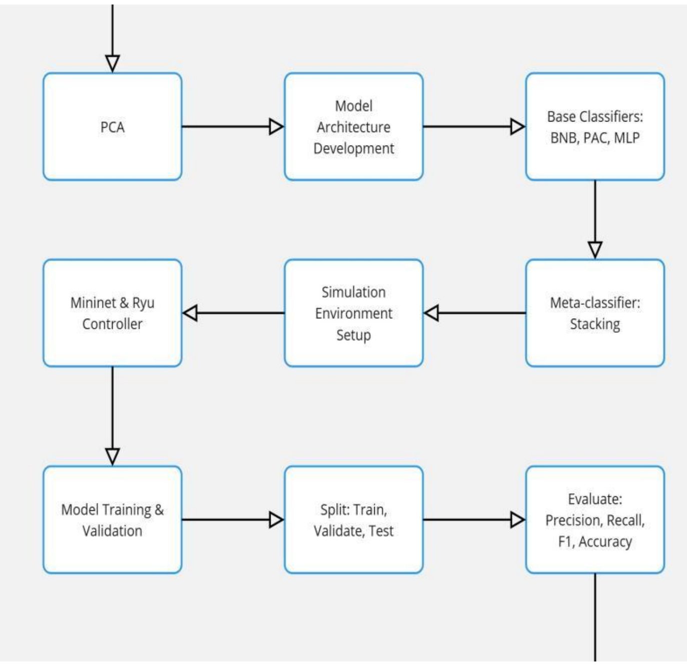
point selection plays a key role in reducing computation complexity but also conserving meaningful information. In this step, features are ranked based on their collective information scores to determine their applicability to the target variable. Only the most instructive features are retained. similarly, star element Analysis(PCA) is implemented to reduce the dimensionality of the dataset by confirmation set, and hyperparameters are optimized for the reduces model redundancy, minimizes spare information, and refines computation power.

3. Discovery of Model Architecture:

This discovery system combines different kinds of classifiers using an ensemble literacy method. The armature of the proposed system has been divided into the base classifiers as well as the mounding classifier:

These are base classifiers as follows: Bernoulli Naive Bayes(BNB) for capturing the probabilistic connection in the double data, a Passive-Aggressive Classifier(PAC) which has been highly efficient for the purposes of streaming data learning thereby suitable in dynamic SDN surroundings. Multi-Layer Perceptron(MLP), a neural network-grounded classifier for landing the nonlinear pattern in the business data.

Here, the combination of mounding takes place using the base classifiers that combine their individual prognostications by use of meta-classifier with enhanced performance in general discovery. This layered framework provides robustness against various kinds of attacks.


PCA
Model Architecture Development
Base Classifiers: BNB, PAC, MLP
Meta-classifier: Stacking
Simulation Environment Setup
Mininet & Ryu Controller
Model Training & Validation
Split: Train, Validate, Test
Evaluate: Precision, Recall, F1, Accuracy

4. Simulation Environment Setup:

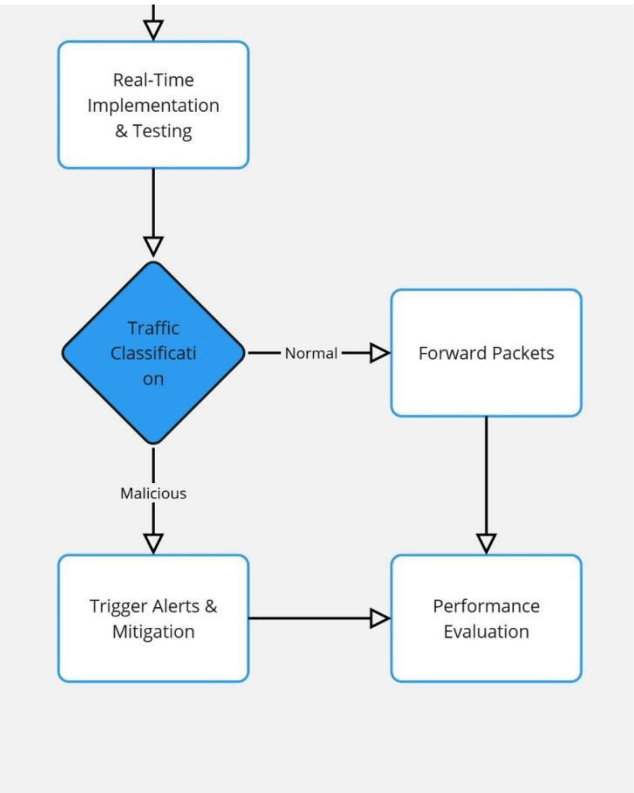
For estimating the proposed system, a simulated SDN terrain is created using highly espoused tools. Mininet is applied in the creation of virtual SDN topology, bluffs concerning network switches, hosts, and regulators, while the Ryu Controller provides the implementation of the SDN control aeroplane with the centralized business operation. Simulation terrain creates business flows which can be used to represent the real conditions such as normal and DDoS attack scripts. Thus, with this, it can reasonably test the real operation of the discovery system.

5. Model Training and Verification :

The preprocessed dataset is split into training, verification, and testing subsets to measure the performance of the model. In the training phase, the ensemble model is trained on labeled business data. ways similar to early stopping and powerhouse are used to prevent overfitting. The performance of the model is tested on the best performance. Metrics such as perfection, recall, F1-score, and delicacy are computed to test the performance of the system in detecting DDoS attacks.

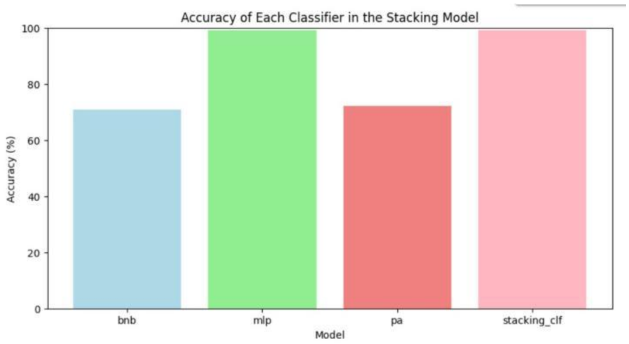
6. Real- Time execution and Testing:

The trained model is implemented in the dissembled SDN environment for real- time DDoS detection. This is an automatic system that watches business incoming flows, correlating implicit anomalies in real time. After detecting a DDoS attack, warnings are sent and mitigation measures implemented by the Ryu controller. This ensures the system can work at efficiency in dynamic networks.



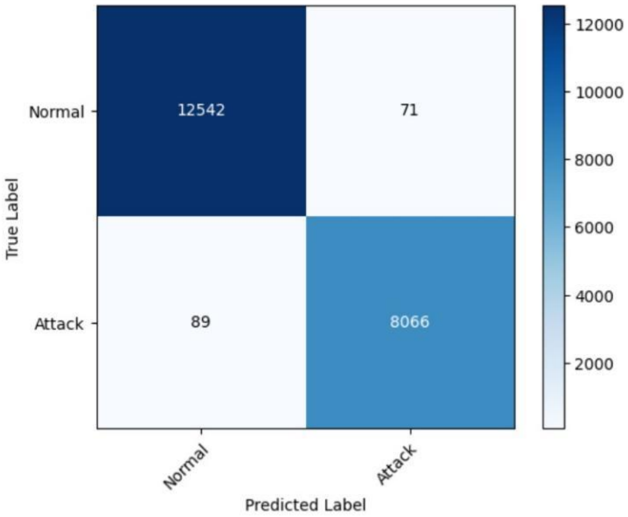
7. Performance Evaluation:

The performance of this system is measured in the various network conditions under which the system can work. Scalability tests will be performed to determine the system's ability to handle increasing business volumes. The model will be tested against various attack types, including low-rate and multi-vector DDoS attacks. Response times for detecting and mollifying attacks would measure real-time capabilities, pressing the system's practical connection.



IV. Result and Discussion

The results demonstrate that the proposed ensemble-based approach achieves high accuracy (99.36%) in detecting DDoS attacks, with minimal false positives and false negatives. The high precision (99.13%) and recall (98.90%) values highlight the model's robustness in identifying malicious traffic without compromising the detection of benign traffic. The F1-score (99.01%) further indicates a balanced performance across precision and recall.
These metrics confirm the efficacy of the ensemble-based detection system in real-world scenarios, ensuring reliable detection of DDoS attacks with minimal overhead.



1.Behaviour Analysis

The system achieves high accuracy, Effectively handles binary categorical data by capturing probabilistic relationships.Ensures rapid adaptability to dynamic SDN traffic.Captures non-linear patterns, enhancing robustness. The ensemble model mitigates individual weaknesses by combining predictions through a stacking classifier.Demonstrates low false positive rates, crucial in SDNs to minimize unnecessary mitigation and maintain network stability/Effectively detects diverse attack types, including low-rate and multi-vector DDoS attacks, with recall rates >90%.Handles increased traffic volumes efficiently using dimensionality reduction (e.g., PCA). Response times for detecting and mitigating attacks are within milliseconds to seconds, suitable for real-time SDN environments.

V.Conclusion

This study proposes a machine learning-based ensemble framework for DDoS attack detection and mitigation in Software-Defined Networks (SDNs). The proposed system improves detection accuracy as well as robustness by applying feature selection through Mutual Information, dimensionality reduction via PCA, and an ensemble method of classifiers: Bernoulli Naive Bayes, Passive-Aggressive Classifier, and Multi-Layer Perceptron all fitted together by a Stacking Classifier. The InSDN dataset is used to evaluate the model and tested in a simulated SDN environment with Mininet and Ryu to prove its real-world applicability.The proposed system improves the detection capability while ensuring computational efficiency and reducing overfitting in the case of low-rate and multi-vector DDoS attack detection under SDN securi

-ty challenges. The experimentation proves scalability along with effectiveness in real-time traffic with high precision and recall.This work contributes to the scalable, flexible framework for real-time DDoS detection in SDNs, thereby allowing further developments on automated detection and response mechanisms dealing with new threats in SDN contexts.

V. References

[1] Wang Jin and Liping Wang, "SDN-Defend: A Lightweight Online Attack Detection and Mitigation System for DDoS Attacks in SDN", Sensors, vol. 22.21, pp. 8287, 2022.

[2] Mohammad Adnan Aladaileh et al., "Effectiveness of an Entropy-Based Approach for Detecting Low-and High-Rate DDoS Attacks against the SDN Controller: Experimental Analysis", Applied Sciences, vol. 13.2, pp. 775, 2023.

[3] K. M. Sudar, M. Beulah, P. Deepalakshmi, P. Nagaraj and P. Chinnasamy, "Detection of Distributed Denial of Service Attacks in SDN using Machine learning techniques", 2021 International Conference on Computer Communication and Informatics (ICCCI), pp. 1-5, 2021.

[4] C. Kaushik;D. VarunTeja;M Shiva Krishna;S. Jaavali 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT).

[5] Ozgur Tonkal, Huseyin Polat, Erdal Ba, Zafer Cömert saran and Ramazan Kocao glu, "Machine Learning Approach Equipped with Neighbourhood Component Analysis for DDoS Attack Detection in Software-Defined Networking", Electronics, vol. 20, no. 11, pp. 1227.

[6] Nisha Ahuja, Gaurav Singal, Debajyoti Mukhopadhyay and Neeraj Kumar, "Automated DDOS attack detection in software defined networking", [online] Available: https://doi.org/10.1016/j.jnca.2021.103108.

[7] H. Lin and P. Wang, "Implementation of an SDNbased security defense mechanism against DDoS attacks", Proceedings of the 2016 Joint International Conference on Economics and Management Engineering (ICEME 2016) and International Conference on Economics and Business Management (EBM 2016), 2016.

[8] D. Yin, L. Zhang and K. Yang, "A DDoS attack detection and mitigation with software-defined Internet of Things framework", IEEE Access, vol. 6, pp. 2469424705, 2018.

[9] S. Singh, R. Sulthana, T. Shewale, V. Chamola, A. Benslimane and B. Sikdar, "Machine-Learning-Assisted Security and Privacy Provisioning for Edge Computing: A Survey", IEEE Internet of Things Journal, vol. 9, no. 1, pp. 236-260, 2022.

[9] O. P. Badve, B. B. Gupta, S. Yamaguchi and Z. Gou, "DDoS detection and filtering technique in cloud environment using GARCH model", IEEE Global Conference on Consumer Electronics (GCCE), pp. 584-586, 2015.

[10] S. Haider, A. Akhunzada, I. Mustafa, T. B. Patel, A. Fernandez, K.-K. R. Choo, et al., "A Deep CNN Ensemble Framework for Efficient DDoS Attack Detection in Software Defined Networks", IEEE Access, vol. 8, pp. 53 972-53 983, 2020.