International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 09 | Sept – 2025 DOI: 10.55041/ISJEM05043
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

A Deep Learning- Based Intrusion Detection and PreventionSystem For Detecting and Preventing Denial-of-Service Attacks

Chebolu pravallika[1],G. Vijaya lakshmi[2]
[1]Mtech pursuing,Sanketika Vidya Parishad Engineering College, Vishakhapatnam, Andhra Pradesh, India
[2]Assistant Professor,Head of the Department,Sanketika Vidya Parishad Engineering College, Vishakhapatnam, Andhra Pradesh, India

Abstract
Denial-of-Service (DoS) attacks pose a significant threat by overwhelming network resources and rendering services inaccessible to legitimate users, while traditional detection systems often fail to keep pace with evolving attack patterns. To overcome this limitation, the proposed project applies machine learning techniques using the dataset *dataset_sdn.csv* to train and evaluate supervised models such as Logistic Regression, K-Nearest Neighbors, Support Vector Machine (SVM), Decision Tree, Random Forest, and Gradient Boosting, comparing their performance in terms of accuracy, precision, recall, and computational efficiency. To enhance detection further, a deep learning-based Long Short-Term Memory (LSTM) neural network is integrated, leveraging its ability to capture sequential dependencies and time-series anomalies that are critical in identifying sophisticated DoS attack behaviors. This combined approach not only improves detection accuracy but also supports real-time adaptation, with the final aim of embedding the framework into an Intrusion Detection and Prevention System (IDPS) for proactive threat mitigation. By uniting machine learning, deep learning, and real-time deployment, the project offers a robust and resilient defense mechanism against DoS attacks.
Keywords: Denial-of-Service, Machine Learning, Deep Learning, LSTM, Intrusion Detection.

Introduction

Denial-of-Service (DoS) attacks have become one of the most severe threats in cybersecurity, capable of overwhelming network resources and disrupting essential services for legitimate users. Traditional detection techniques often struggle to adapt to the evolving strategies employed by attackers, resulting in delayed response times and reduced accuracy. To address these challenges, the proposed system employs a hybrid detection framework that integrates machine learning and deep learning models. Supervised algorithms such as Logistic Regression, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree, Random Forest, and Gradient Boosting are trained and benchmarked on the *dataset_sdn.csv* to establish a robust baseline for attack detection by evaluating accuracy, precision, recall, and computational efficiency.

While traditional classifiers provide strong predictive performance, they are limited in detecting time-dependent anomalies and sequential attack behaviors. To overcome this, the framework incorporates a Long Short-Term Memory (LSTM) neural network, which excels in processing sequential and time-series data. LSTMs capture long-range dependencies within traffic flows, enabling the detection of anomalies that gradually emerge during the course of an attack. By combining conventional machine learning models with LSTM, the system enhances adaptability, accuracy, and resilience against both immediate anomalies and evolving attack patterns, ensuring a more reliable defense mechanism against sophisticated DoS threats.

The ultimate objective of this research is to integrate the hybrid detection system into a real-time Intrusion Detection and Prevention System (IDPS) tailored for Software-Defined Networking (SDN) environments. SDN provides centralized control, scalability, and programmability, making it a suitable platform to deploy intelligent defense solutions. Embedding the proposed framework into SDN not only enables timely detection but also allows proactive mitigation of attacks before they impact critical services. This integration contributes toward building intelligent, scalable, and adaptive cybersecurity infrastructures, representing a significant step forward in protecting digital networks from increasingly complex cyber threats.

Literature review

The growing frequency and sophistication of cyber threats have made Network Intrusion Detection Systems (NIDS) a crucial component of modern cybersecurity. Machine learning techniques have been widely explored for their ability to classify normal and malicious traffic patterns with high accuracy. For instance, Bertoli (2021) proposed the AB-TRAP framework, which integrated machine learning algorithms with updated datasets to improve intrusion detection and achieved an F1-score of 0.96 in LAN environments [1]. Similarly, Santos et al. (2020) investigated Distributed Denial-

of-Service (DDoS) detection in Software-Defined Networks (SDN) using machine learning classifiers, highlighting the method's effectiveness but also its challenges in handling concurrent data flows and computational overhead [2]. Earlier studies, such as Dong and Sarem (2019), applied the K-Nearest Neighbor (KNN) algorithm for DDoS detection but observed limitations in scalability and robustness when working with large, noisy datasets [3]. These works collectively demonstrate that while machine learning provides powerful tools for anomaly detection, challenges related to computational efficiency, dataset quality, and adaptability persist in real-world deployments.

Recent research emphasizes the importance of hybrid and advanced detection techniques to overcome the shortcomings of traditional classifiers. Gu, Li, and Guo (2019) proposed a semi-supervised hybrid feature selection method that combined clustering with supervised learning to enhance detection accuracy, though limitations in adaptability remained [4]. Other approaches, such as Midzic, Avdagic, and Omanovic (2019), leveraged hybrid AI methods by combining self-organizing maps with fuzzy inference systems, achieving promising detection rates but struggling with false alarm rates in ambiguous traffic scenarios [5]. Furthermore, studies like Wang, Zhang, and Singhal (2019) introduced metrics such as k-Zero Day Safety to quantify risks associated with unknown vulnerabilities, reinforcing the need for dynamic models that can respond to evolving threats [6]. Collectively, these findings underline the critical need for scalable, adaptive, and computationally efficient NIDS frameworks that integrate machine learning, deep learning, and real-time deployment strategies to safeguard modern high-speed networks against evolving attack vectors.

Scope of the Present Investigation

Software-Defined Networking (SDN) provides applications with a unique capability to access complete network information, offering centralized visibility and fine-grained control over traffic flows. This centralized perspective enables applications to make informed decisions for optimizing resource allocation, managing network congestion, and improving performance during high-traffic periods. By integrating multiple applications, SDN further supports advanced functionalities such as load balancing and intrusion detection, ensuring that resources are efficiently utilized while threats are promptly addressed. One of the major strengths of SDN lies in its adaptability; when anomalies or malicious activities are detected, the SDN controller can dynamically instruct the data plane to reconfigure itself, thereby mitigating potential issues in real time. Since both the control plane and data plane operate across distributed routers with open interfaces, the programmability of SDN ensures flexibility, scalability, and enhanced security, all while maintaining smooth and reliable network operations.

Existing System

In recent years, extensive research has been carried out on the use of Software-Defined Networking (SDN) for Distributed Denial-of-Service (DDoS) detection, with many works employing machine learning techniques such as Naive Bayes, K-Nearest Neighbors (KNN), Neural Networks, Support Vector Machines (SVM), and Self-Organizing Maps (SOM). Each of these algorithms offers certain benefits but also faces challenges in terms of accuracy, computational efficiency, scalability, and adaptability to evolving attack strategies. Among the surveyed approaches, Support Vector Machine (SVM) and Random Forest have been selected for this project due to their complementary strengths. SVM is known for constructing optimal decision boundaries to distinguish between normal and malicious traffic, while Random Forest employs an ensemble of decision trees to achieve high predictive accuracy and minimize overfitting. Together, these algorithms enhance detection capabilities by combining SVM's precision with Random Forest's robustness, allowing the system to effectively analyze essential traffic features such as flow duration, packet size, and connection frequency. This combination provides proactive defense against DDoS attacks, thereby ensuring service availability and securing SDN infrastructures. However, the limitations of the existing system lie in the experimental accuracy levels and the increased complexity of implementation, which hinder real-time efficiency.

Proposed System

The proposed system adopts a hybrid detection framework that integrates both machine learning and deep learning techniques to ensure highly accurate and efficient detection of Denial-of-Service (DoS) and Distributed DoS (DDoS) attacks. Initially, several supervised machine learning models—such as Logistic Regression, K-Nearest Neighbors (KNN), Decision Trees, Random Forest, and Support Vector Machines (SVM)—are trained and benchmarked to evaluate their precision, recall, and overall performance in distinguishing malicious traffic from normal traffic. This evaluation establishes a reliable baseline and identifies the most effective algorithms for real-time deployment within an SDN environment. To further strengthen the framework, a Long Short-Term Memory (LSTM) neural network is

integrated, as it excels at processing sequential data and capturing time-dependent patterns in network traffic. By analyzing temporal dependencies and subtle behavioral changes, the LSTM enhances detection accuracy beyond traditional classifiers, enabling the system to recognize sophisticated attack strategies. The hybrid approach therefore delivers a highly robust detection mechanism that combines the computational efficiency of machine learning with the adaptive learning power of deep learning.

Advantages of the Proposed System

The proposed hybrid system offers several advantages: it achieves high accuracy in detecting network intrusions, provides effective identification of DoS and DDoS attacks, and ensures low computational complexity for faster processing. Additionally, the framework supports real-time detection and mitigation of threats, which is essential for maintaining uninterrupted services. Its scalable design also makes it adaptable for deployment across different network sizes and varying traffic loads, thereby ensuring reliability and resilience in dynamic environments.

Methodology

Dataset and Pre-processing:

For this project, a comprehensive dataset containing both fake and genuine social media profiles was utilized to develop and evaluate classification models. The dataset includes attributes such as the number of friends, followers, status counts, profile activity metrics, and engagement patterns, which serve as key indicators for distinguishing authentic accounts from fraudulent ones. These features capture subtle behavioral differences that may not be evident through manual inspection but can be leveraged by machine learning algorithms to identify fake profiles. To ensure data quality, preprocessing techniques such as handling missing values through imputation, encoding categorical variables, and scaling numerical features were applied. This step reduces noise, minimizes bias, and ensures consistency across the dataset. The processed data was then split into training and testing subsets, with 80% used for training the models and 20% reserved for testing, enabling unbiased evaluation of classification accuracy, precision, and recall.

Feature Selection:

Feature selection plays a vital role in enhancing classification model performance by identifying the most relevant and independent attributes while minimizing redundancy. For this project, features were carefully selected from a publicly available dataset containing 1,337 fake users and 1,481 genuine users. Attributes such as user name, status count, number of friends, follower count, favorites, and languages known were considered. These features provide insights into user engagement, with anomalies such as disproportionate follower-to-friend ratios or unusually low status counts signaling suspicious activity. By focusing on meaningful attributes, the model is able to detect complex behavioral patterns and interactions that help differentiate fake accounts from real ones. Effective feature selection reduces false positives and false negatives, ultimately leading to higher model precision and reliability.

Classification:

Classification forms the core of the methodology, where the dataset is analyzed and each profile is assigned to either the fake or genuine category. In this project, the Support Vector Machine (SVM) algorithm was chosen as the primary classifier due to its effectiveness in handling high-dimensional and complex datasets. SVM works by transforming input data into a higher-dimensional feature space, where it identifies an optimal hyperplane that maximally separates the two classes. This ensures accurate classification even in non-linear scenarios. The use of kernel functions such as radial basis function (RBF) allows the SVM to capture intricate relationships between features that simpler linear models might miss. Once trained, the SVM classifier generalizes well to unseen data, effectively minimizing false positives and false negatives. Its robustness makes it a reliable tool for detecting fraudulent accounts in real-world scenarios and improving the security of social media platforms

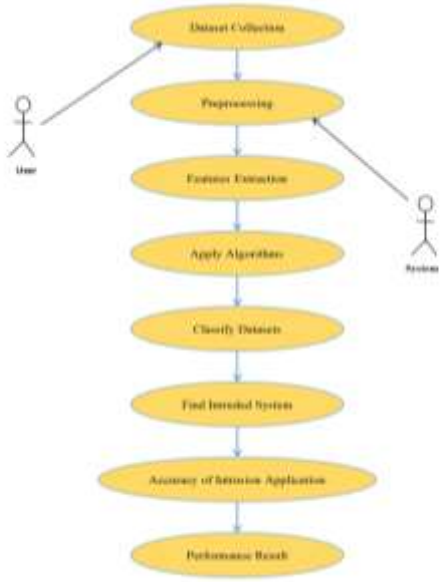


Fig 1: Case Diagram.

The working procedure of the proposed system is implemented in two main stages: **traditional machine learning classifiers** and **deep learning with LSTM**. First, the dataset (*dataset_sdn.csv*) is loaded and cleaned by removing missing values, after which categorical variables like protocol type, source, and destination addresses are encoded using LabelEncoder to convert them into numeric form. Features are separated from the target label, and all numerical attributes are standardized using StandardScaler to ensure uniformity across the dataset. The data is then split into training and testing subsets (80–20 ratio), maintaining stratification to preserve class balance. Several machine learning models, including Logistic Regression, K-Nearest Neighbors, Support Vector Machine, Decision Tree, Random Forest, and Gradient Boosting, are initialized and trained on the training set. Each model is evaluated on the testing set using metrics like accuracy, precision, recall, and F1-score, and the results are compared both numerically and visually through a bar chart for accuracy.

In the second stage, a **Long Short-Term Memory (LSTM) neural network** is introduced to enhance sequential pattern recognition. After preprocessing, the feature matrix is reshaped into a 3D format suitable for LSTM input, where each sample is treated as a time step with its features. The target labels are one-hot encoded using to_categorical to facilitate multi-class classification. The LSTM model is then built with an architecture consisting of an LSTM layer with 64 units, a dropout layer for regularization, a dense hidden layer of 32 neurons with ReLU activation, and an output layer with softmax activation for binary classification (normal vs. attack). The model is compiled using the Adam optimizer and trained for 10 epochs with a batch size of 64, validating performance on the test set. Finally, the model's accuracy is evaluated, providing a benchmark to compare its performance against traditional machine learning classifiers. This hybrid methodology ensures robust detection of DoS/DDoS attacks, leveraging both classical statistical learning and deep sequential modeling.

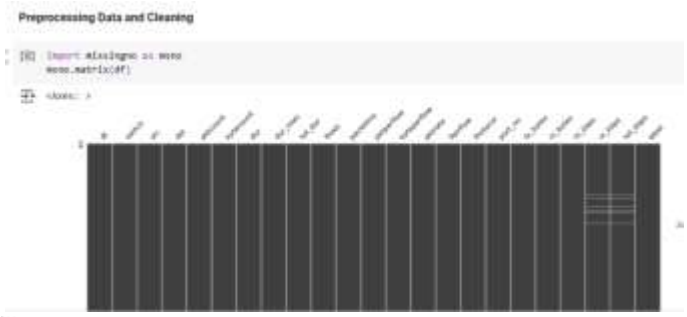
Results

Reading Dataset

```
path = "dataset_sdn.csv"
df = pd.read_csv (path)
print(df)
```

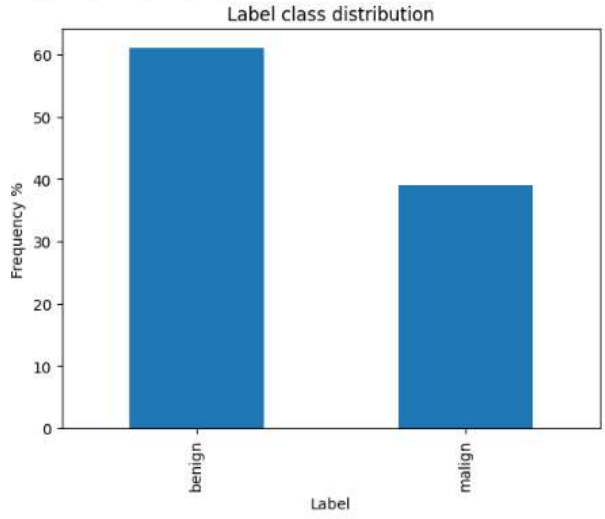
```
            dt  switch       src       dst  pktcount  bytecount  dur  \
0        11425       1  10.0.0.1  10.0.0.8     45304   48294064  100
1        11605       1  10.0.0.1  10.0.0.8    126395  134737070  280
2        11425       1  10.0.0.2  10.0.0.8     90333   96294978  200
3        11425       1  10.0.0.2  10.0.0.8     90333   96294978  200
4        11425       1  10.0.0.2  10.0.0.8     90333   96294978  200
...        ...     ...       ...       ...       ...        ...  ...
104340    5262       3  10.0.0.5  10.0.0.7        79       7742   81
104341    5262       3  10.0.0.5  10.0.0.7        79       7742   81
104342    5262       3  10.0.0.11  10.0.0.5       31       3038   31
104343    5262       3  10.0.0.11  10.0.0.5       31       3038   31
104344    5262       3  10.0.0.11  10.0.0.5       31       3038   31
```

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 09 | Sept – 2025 DOI: 10.55041/ISJEM05043
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

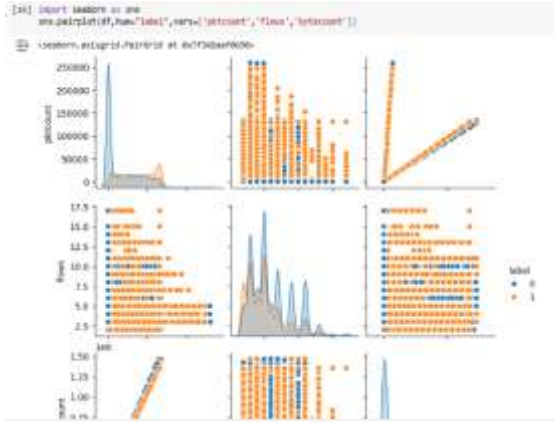



```
[14] # Let's plot the Label class against the Frequency
     labels = ['benign','malign']
     classes = pd.value_counts(df['label'], sort = True) / df['label'].count() *100
     classes.plot(kind = 'bar')
     plt.title("Label class distribution")
     plt.xticks(range(2), labels)
     plt.xlabel("Label")
     plt.ylabel("Frequency %")
```

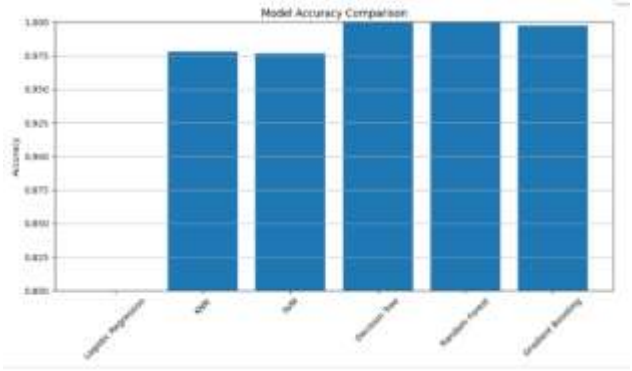
The graph illustrates the label class distribution of the dataset, highlighting its binary classification nature with two categories: benign (normal) and malign (abnormal/attack). Benign cases form the majority class at about 61% of the data, while malign cases represent the minority with around 39%. This moderate imbalance indicates that benign instances occur more frequently, which may cause machine learning models to lean toward predicting the benign class, thereby reducing sensitivity in detecting malign cases. To address this, techniques such as oversampling the minority class, undersampling the majority class, or applying SMOTE can be used to balance the dataset. The ultimate goal is to develop a model that performs fairly and accurately for both classes, ensuring reliable and unbiased predictions.



Graph1: Label



The graph presents a Model Accuracy Comparison among six machine learning algorithms—Logistic Regression, KNN, SVM, Decision Tree, Random Forest, and Gradient Boosting—showing their performance on the dataset. Accuracy values are plotted on the Y-axis, ranging between 0.80 and 1.00, reflecting the very high overall effectiveness of all models. Both K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) achieve accuracies of around 97.5%, indicating strong reliability in classification tasks. Logistic Regression performs nearly perfectly, reaching close to 100% accuracy, while Decision Tree, Random Forest, and Gradient Boosting models also score exceptionally well, with results in the 99.9–100% range. Although all classifiers demonstrate excellent performance, tree-based ensemble methods slightly outperform the others, establishing themselves as the most accurate approaches for this dataset.



Graph 2: Model accuracy Comparison

Conclusion

This project demonstrates the effectiveness of integrating machine learning and deep learning techniques for real-time detection and prevention of Denial-of-Service (DoS) attacks. By testing multiple supervised classifiers on real-world datasets, the study revealed the strengths and limitations of each approach, identifying which models are most reliable for malicious traffic detection in Software-Defined Networks (SDNs). Among them, the Long Short-Term Memory (LSTM) model proved highly capable, as it efficiently captured temporal patterns in sequential traffic data, offering superior accuracy in identifying complex and evolving attack behaviors. Incorporating such models into an Intrusion Detection and Prevention System (IDPS) enhances network resilience by ensuring continuous monitoring, instant detection, and automated mitigation of threats, thereby minimizing disruptions and safeguarding stable network performance.

Future Scope

Future research can focus on strengthening the framework by incorporating larger and more diverse datasets, particularly real-time traffic from multiple network environments, to improve model robustness and generalization. Exploring hybrid detection approaches that combine the adaptability of deep learning with the speed of feature-engineered machine learning models could further enhance accuracy. Adaptive learning techniques, such as online and incremental training, can be applied to enable the system to evolve with new attack patterns without the need for full retraining. Finally, optimizing the framework for low-latency performance will allow seamless deployment in large-scale, distributed IDPS environments, ensuring faster and more effective real-time cyber defense.

References

[1]. GustavodeCarvalhoBetroli (2021). Network Intrusion Detection System With the Degree of An End-to-End Framework for Machine Learning. IEEE Access 10.1109/ACCESS.2021.3101188

[2].Dong,S.,&Sarem,M.(2020).DDoS Attack Detection Method Based on Improved KNN With the Degree of DDoS Attack in Software-Defined Networks. IEEEAccess,8, 5039-5048.

[3]. Gu, Y., Li, K., Guo, Z., & Wang, Y. (2019). Semi supervised K-means DDoS detection method using hybrid feature selection algorithm. IEEE Access, 7, 64351- 64365.

[4]. Meti, N., Narayan, D. G., &Baligar, V. P. (2017, September). Detection of distributed denial of service attacks using machine learning algorithms in software defined networks. In 2017 international conference on advances in computing, communications and informatics (ICACCI) (pp. 1366-1371). IEEE

[5]. 15th International Symposium on Pervasive Systems, Algorithms and Networks IEEEDDoSAttackIdentificationandDefenseusingSDNbasedonMachineLearning Method, 2018

[6]. Muthamil Sudar, K., & Deepa lakshmi, P. (2020). A two level security mechanism to detect a DDoS flooding attack in software-defined networks using entropy-based and C4. 5 technique. Journal of High Speed Networks, (Preprint), 1-22.

Acknowledgement

Mrs.G. vijayalakshmi working as an Assistant professor and Head of the Departmentin Department of CSE in sanketika vidya parishad engineering college, Visakhapatnam,AP with 5 yrs teaching experience and member of IAENG, accredited by Naac with herareas of interests in C, Data warehousing and data mining, Design and analysis ofalgorithm, python, formal languages and automata theory,compiler design.

Chebolu Pravallika is pursuing her final semester M.Tech in Sanketika Vidya ParishadEngineering College, accredited with A grade by NAAC, affiliated by Andhra University andapproved by AICTE. With interest in Machine Learning Chebolu Pravallika has taken up herPG project on A Deep Learning- Based Intrusion Detection and PreventionSystemFor Detecting and Preventing Denial-of-Service Attacksand published the paper in connection to the project under the guidance of Mrs.G.vijayalakshmi, Assistant Professor, SVPEC