Course Code: CS64123

Task 1: CLASSIFICATION OF DDOS DETECTION AND MITIGATION DATASET USING DEEP LEARNING

Prepared by:

Kantipudi Sruthi	2206222
P Sravya	2206083
Sanapala Sai Siddardha	2206230

1. Introduction

DDoS attacks pose a critical cybersecurity threat, disrupting networks by overwhelming target servers with malicious traffic. A distributed denial of service attack is a DoS attack, in which multiple hosts perform DoS attacks in a coordinated fashion to one or more targets. Software-Defined Networking (SDN) offers a dynamic, centralized approach to traffic management, making it a powerful solution for detecting and mitigating such attacks.

This project leverages deep learning models—Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM)—to accurately classify normal and attack traffic in an SDN environment.

2. Dataset Collection

The dataset was generated using Mininet and an SDN controller in a simulated SDN environment, replicating realistic DDoS attack scenarios. It comprises:

- **Benign Traffic**: Normal network packets.
- **DDoS Attack Traffic**: Simulated attack packets created with custom scripts.

The dataset contains 15,462,150 entries with 26 features, including packet sizes, inter-arrival times, protocol types, and flow statistics (e.g., packet_count, byte_count, flow_duration_sec). After preprocessing (removing NaN and infinity values), the dataset was reduced to 2,667,523 entries, ensuring data quality for robust model training and evaluation.

3. Types of DDoS Attacks

The dataset includes various DDoS attack types:

- **UDP Flood**: Overwhelms targets with excessive UDP packets.
- **TCP SYN Flood**: Exploits the TCP handshake process to exhaust resources.
- **ICMP (Ping) Flood:** Overloads networks with ICMP echo requests.
- **HTTP Flood:** Bombards web servers with high HTTP requests.
- **Slowloris Attack:** Keeps multiple connections open without completing them, leading to resource depletion.
- **DNS Amplification Attack**: Uses publicly accessible DNS servers to flood a target with traffic.
- **NTP Reflection Attack:** Exploits Network Time Protocol (NTP) servers to amplify attack traffic.

4. Software-Defined Networking (SDN)

SDN is a revolutionary networking paradigm that decouples the control plane from the data plane, offering:

- **Centralized Traffic Management:** SDN controllers efficiently monitor and route network traffic.
- **Adaptive Security Measures:** Enables dynamic, real-time response to cyber threats.
- **Scalability & Efficiency:** Optimizes resource allocation and traffic flow.
- **Improved Network Visibility:** Provides real-time monitoring of network traffic patterns.

By leveraging SDN controllers, this project enables real-time **DDoS attack detection and mitigation.**

5. Feature Engineering and Selection

To improve model performance and reduce complexity, feature selection techniques were applied:
Correlation Analysis: Features highly correlated with others were identified and removed to prevent redundancy.
Statistical Methods: Features with low variance and minimal impact on classification were eliminated.
Feature Scaling: Normalization techniques were used to standardize data for better model convergence.

These preprocessing techniques enhanced model interpretability and efficiency while ensuring robust performance

6. Deep Learning Models for DDoS Detection

6.1 Long Short-Term Memory (LSTM)

- Designed for sequential data processing.
- Captures time-based anomalies in network traffic for enhanced detection.
- Effective in detecting slow and evolving DDoS attacks.

6.2 Artificial Neural Network (ANN)

- Fully connected layers learn complex patterns in network traffic.
- Differentiates between normal and attack packets efficiently.
- Provides a lightweight model for real-time traffic classification.

6.3 Convolutional Neural Network (CNN)

- Extracts spatial features from network traffic patterns.
- Enhances feature representation for effective DDoS detection.

- Detects anomalies in network flows with high precision.

7. Implementation and Results

7.1 Model Training and Hyperparameters

- Epochs: 5
- Batch Size: 128
- Optimizer: Adam
- Loss Function: Binary Crossentropy
- Activation Functions: ReLU

7.2 Performance Evaluation Metrics

To assess model effectiveness, the following metrics were used:

- **Accuracy:** Measures the percentage of correctly classified instances.
- **Precision:** Indicates how many detected attacks were actual attacks.
- **Recall:** Measures the model's ability to detect all actual attacks.
- **ROC-AUC**: Trade-off between true and false positive rates.

7.3 Model Performance and Comparison

- **LSTM:** Outperformed other models in identifying time-dependent attack behaviours, making it the most effective for sequential network data.

- **CNN:** Performed well in detecting attacks with structured patterns but struggled with temporal dependencies.
- **ANN:** Provided a lightweight alternative but lacked the robustness of CNN and LSTM.

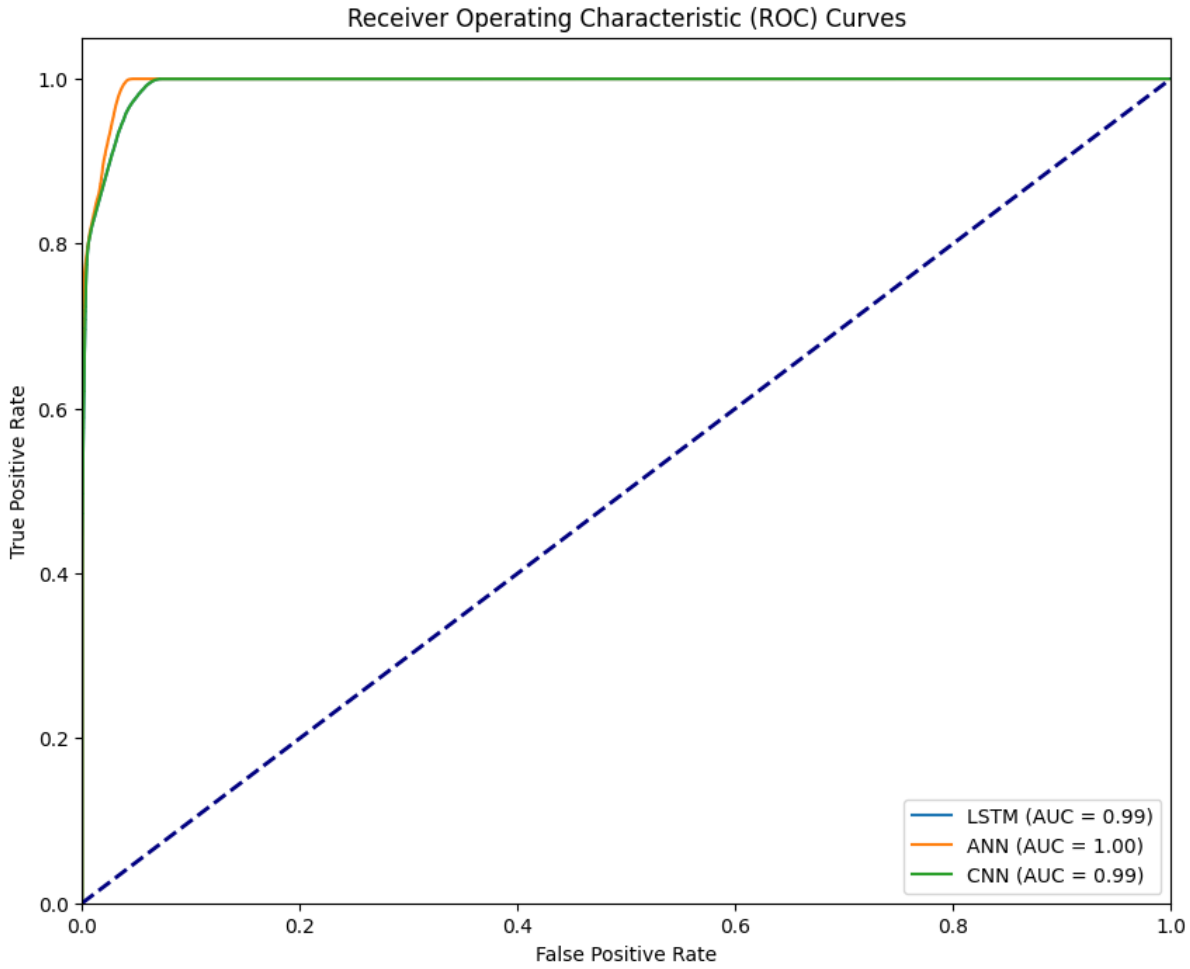
	Model	Accuracy	Precision	Recall
0	LSTM	0.9831	0.9752	0.9999
1	CNN	0.9742	0.9623	0.9993
2	ANN	0.9810	0.9720	0.9997

7.4 ROC-AUC Curves & Loss Curves

ROC-AUC Curve: Evaluates the trade-off between true positive and false positive rates. A higher AUC indicates better model discrimination ability.

Loss Curve: Represents how well the model is learning. A smoothly decreasing loss curve suggests effective training, while fluctuating loss may indicate overfitting or suboptimal hyperparameters.

Our models achieved high accuracy in detecting DDoS attacks, proving the effectiveness of deep learning in SDN-based cybersecurity. The LSTM model outperformed others in identifying time-based attack patterns, while CNN provided superior feature extraction capabilities.



8. Conclusion

This project highlights how deep learning can intelligently detect and mitigate DDoS attacks in SDN environments.

Integrating network anomaly detection with proactive defense mechanisms can further enhance SDN's ability to handle sophisticated cyber threats. Additionally, expanding datasets with real-world attack traffic can improve model generalization and resilience.

References: https://ijcionline.com/paper/12/12423ijci08.pdf