DDoS Attack Classification Using Machine Learning

1st Taimoor Fahim
dept. of CS
Brac University
ID-23240193
taimoor.fahim@g.bracu.ac.bd

2nd Mahdi Hossain
dept. of CS
Brac University)
ID-20301194
mahdi.hossain@g.bracu.ac.bd

3rd Kazi Shahed Mamun
dept. of CS
Brac University
ID- 20301471
kazi.shahed.mamun@g.bracu.ac.bd

I. Introduction

A. Overview of DDoS Attacks

Conveyed Refusal of Administration (DDoS) assaults is a conspicuous danger in the scene of organization security, where numerous compromised frameworks flood the transmission capacity or assets of a designated framework, frequently at least one web server. Such goes after upset the administrations of a host associated with the web, making the site or administration inaccessible to clients, which can prompt huge monetary and reputational harm.

B. Objective of the Study

The principal objective of this review is to foster a powerful AI model that can precisely group network traffic as one or the other typical or a piece of a DDoS assault. By utilizing the CIC IDS 2017 dataset for preparation and testing, this task intends to upgrade the capacities of network safety frameworks to really identify and alleviate DDoS dangers.

II. Literature Review

A. Existing Methods in DDoS Detection

The making review uncovers various strategies really embraced to deal with the issue of DDoS attacks. Traditional techniques regularly coordinate close-based structures that require predictable updates as attack plans advance. Before long, these structures regularly come up short in changing in accordance with new and emerging bets, highlighting the need for extra solid and flexible methodologies.

B. Machine Learning Techniques in Cybersecurity

Machine Learning offers promising strategies because of its capacity to procure information and make savvy choices. Methods like Unusual Timberlands, Key Fall away from the Faith, and Mind Affiliations have been used in late appraisals to pack and expect online security gambles with high precision. These models can manage huge datasets and perceive unpretentious models that could show a poisonous turn of events, hence offering an essential benefit over customary methodology.

III. Methodology

A. Data Collection

The dataset utilized in this preliminary is the energetically open CIC IDS 2017 dataset, which organizes a colossal variety of preventions reflected in a controlled setting. The dataset consolidates parts that are fundamental for building the evident confirmation models, for instance, stream terms, full forward social events, firm backward packages, and others.

B. Data Preprocessing

Information Preprocessing The preprocessing steps included overseeing missing qualities, normalizing the information, and encoding outright parts to mathematical attributes to set up the dataset for persuading model availability. The dataset was also isolated into arranging and testing subsets, guaranteeing genuine dispersal of standard and assault models in the two sets.

C. Model Selection

Three man-made knowledge models were picked thinking about their inescapability and adequacy in social occasion errands: 1. Random Forest Classifier - Known for its high precision and power against overfitting. 2. Logistic Regression - Significant for matched demand undertakings and gives probabilities to results, which helps in limit tuning. 3. Neural Network (Multi-Layer Perceptron) - Offers flexibility and learning limits that are especially huge for complex models in information.

D. Training and Testing

The models were trained using the training subset of the dataset, with each model being evaluated on its accuracy, precision, recall, and F1 score using the test data. The performance of these models provides insights into their effectiveness in classifying and predicting DDoS attacks under different network conditions.

IV. Limitations

Dataset Reliance: The model's show is exceptionally subject to the quality and combination of the plan information.

Pushing Assault Frameworks: Aggressors persistently energize new techniques, requiring the model to be animated with information mirroring these gamble

V. System Architecture

A. Overview of the System Setup

The system is designed to efficiently process and analyze network traffic data to detect DDoS attacks. It incorporates various machine learning models that operate on a robust computational framework capable of handling large volumes of data.

B. Workflow and Data Flow Diagrams

The workflow begins with data collection, followed by preprocessing, model training, and finally, testing and validation. The data flow diagram illustrates the process from data ingestion to output generation, highlighting the interaction between different system components.

VI. Synthesis

This task gives a beginning stage to involving recreated knowledge for DDoS assault demand. While obstructions exist, the undertaking remembers the capacity of this procedure for updating network security. Future work could look at cutting-edge assessments, and consistent executions, and perpetually change the model to counter-attack techniques.