DDoS Protection System for Cloud: Architecture and Tool

Panel No. 06
Supervisor Name
Dr. Balaji Srikaanth P, AP/NWC
Dr. S. Nagendra Prabhu, AP/CINTEL

Batch No. NW000156
Shaurya Singh Srinet – RA2111032010006
Shounak Chandra – RA2111032010026
Charvi Jain – RA2111047010113

Functional Document for User Story 2: Simulate Traffic and Generate Dataset

1. Introduction

This user story focuses on simulating traffic and generating datasets using Apache JMeter for training the Isolation Forest anomaly detection model. The datasets will include both benign and DDoS attack traffic to help develop a robust system for detecting anomalies in real-time cloud environments.

2. Product Goal

The goal is to simulate realistic network traffic patterns using JMeter and generate labelled datasets that will be used for training the anomaly detection system. The dataset must accurately represent real-world scenarios to ensure the model's reliability and effectiveness.

3. Demography (Users Location)

- **Target Users:** Cloud security engineers, data scientists, AI developers.
- **User Characteristics:** Users familiar with traffic simulation, data preprocessing, and model training workflows.
- **Location:** Organizations globally, with a focus on those operating in cloud-hosted environments.

4. Business Processes

- **Traffic Simulation:**
  - Configure JMeter to simulate benign traffic with varying patterns.
  - Introduce DDoS attack scenarios using stress-testing configurations.
- **Data Logging:**
  - Capture traffic data and export logs in CSV format.
  - Record key metrics such as source IP, destination IP, traffic volume, and timestamp.

- **Data Preprocessing**:
    - Label traffic entries as **benign** or **malicious** based on predefined criteria.
    - Normalize and clean the data for consistency.

5. Features

- **Traffic Simulation with JMeter:**
    - Simulate traffic using pre-configured JMeter test plans.
    - Generate scenarios for both normal usage and DDoS attacks.
- **Data Export and Labelling:**
    - Export raw traffic logs to CSV format.
    - Label datasets with accuracy to differentiate between benign and malicious activities.
- **Dataset Review:**
    - Validate and balance the dataset for model training, ensuring an equal representation of benign and attack traffic.

6. Authorization Matrix

ROLE	**Access Level**
Developer	Full access to configure and manage AWS resources.
Data Scientist	Access to processed and labelled datasets.
Security Analyst	Access to simulation logs for analysis.

7. Assumptions

- The JMeter test plan runs without errors during simulations.
- Dataset labelling criteria are well-defined and consistently applied.
- Necessary tools and dependencies for data preprocessing are available and functional.

8. Target Audience

- **Audience:** Data scientists, cloud security teams, and developers working on anomaly detection systems.
- **Effort Estimation:** Approximately 5 days for simulation, data generation, and preprocessing.

9. Acceptance Criteria

- Traffic logs are generated and exported successfully using JMeter.
- Logs are processed into CSV format without errors.
- Data entries are accurately labelled as benign or malicious.
- The final dataset is balanced and ready for model training.

10. Checklist

- JMeter configured for traffic simulation.
- Traffic logs exported to CSV format.
- Data labelled correctly for benign and malicious traffic.
- Dataset reviewed and balanced for model training.