CLASSIFICATION OF SOFTWARE DEFINED NETWORK TRAFFIC TO PROVIDE QUALITY OF SERVICE

TEAM - 09

AGENDA

1. Abstract
2. Literature Review
3. Proposed System
4. Objectives
5. System Requirements
6. System Architecture
7. UML Diagrams
8. Modules
9. Status of work

Abstract

Numerous network operations, from security monitoring to accounting, from Quality of Service to giving operators relevant estimates for long-term provisioning, depend on accurate traffic classification. The initial stage in analyzing and classifying the various types of applications running via a network is network traffic classification. This method allows network operators or internet service providers to control the overall performance of a network. We apply machine learning models to categorize traffic by application. This can be done by extracting the features of the traffic. This classified data can be used to stop unnecessary traffic and allow only user required traffic. Basically we prioritize the network traffic based on the features extracted during the classification. Features related to OTT are identified and we try to restrict them in the network for reducing the traffic in the network for providing better quality of service. We intend to stop traffic from Over-the-top(OTT) platforms like Netflix, Prime Videos, etc. Hence, by this the quality of service can be improved for user required applications.

Literature Review

- Gianni D'Angelo et al. suggested a model that begins with statistical characteristics (basic features), taken from traffic flows over a predetermined time period, and creates additional features that explain the correlations between the features (spatial features), as well as changes in those features over time (temporal features). They suggested a deep architecture made up of neural networks based on autoencoders (AEs). The autoencoders encode-decode function contain various combinations of recurrent and convolutional network layers in order to extract such information. The following combinations were looked into: CNN, LSTM, ConvLSTM, CNN-LSTM, and Stacked-CNN-LSTM. The LSTM recurrent network was used to extract temporal features, while the convolutional network was utilised to extract spatial features.

Literature Review

- Kourosh Ahmadi et al. have conducted a fuzzy logic control system (FLCS) that may be used into an SDN controller to improve QoS for different service flows.The path weight for a specific communication line is determined by the FLE-SDN controller using a fuzzy logic control mechanism after it continually collects the QoS measurements of all communication channels between different networks.The process result is the identification of the best route for a particular service flow, which is followed by instructions to forwarding devices to adjust their flow forwarding techniques. Particularly for real-time applications like audio and video, the FLE-SDN approach has a proven track record of assisting SDN controllers in trying to improve the architecture of various service flows and providing greater QoS.The limitation of this study was designing service flows with the goal of providing improved QoS, this technique has not been implemented on SDN controllers.

Literature Review

- Yuyang Zhou at el. proposed a new intrusion detection framework and this framework is based on the feature selection and ensemble learning techniques. CFS-BA algorithm is used for dimensionality reduction, which selects the optimal subset based on the correlation between the features. C4.5,Random Forest and Forest by Penalizing Attributes algorithms are used in ensemble approach along with AOP rule to construct the classification model. Later voting technique is used to combine the probability distributions of the base learners for attack recognition. The data sets used during the experiment were NSL-KDD,AWID and CIC-IDS 2017. The experimental results are promising with an accuracy of classification equal to 99.81%, 99.8% DR and 0.08% FAR with a subset of 10 features for the NSL-KDD dataset, and the obtained results for the AWID provide accuracy of 99.52% and 0.15% FAR with a subset composed of only 8 features. Remarkably, our model achieves the highest accuracy of 99.89% and DR of 99.9% on the subset of 13 features for the CIC-IDS2017 dataset.

Proposed System

- A helpful tool for file analysis and network traffic monitoring is packet capture (PCAP). In order to study the network properties, PCAP files are utilized. By using these qualities, features that will be utilized for classification will be extracted.
- By utilizing their unique qualities, we hope to identify network traffic coming from OTT services. In comparison to earlier articles, we intend to employ more features for classifications. This results in higher accuracy rates.
- We suggest an ensemble model for classification. We intend to use Random Forest, C4.5 and SVM(Support Vector Machine) to form an ensemble model.
- After classifying network traffic to its application, in the deployment state, traffic from OTT platforms will be restricted and we prioritize the network traffic specific to the application.

Objectives

- The sole objective of the project is to improve the quality of service.
- To classify the network traffic to its application base.
- To stop unnecessary traffic and allow only user required traffic.
- To prioritize necessary applications

System Requirements

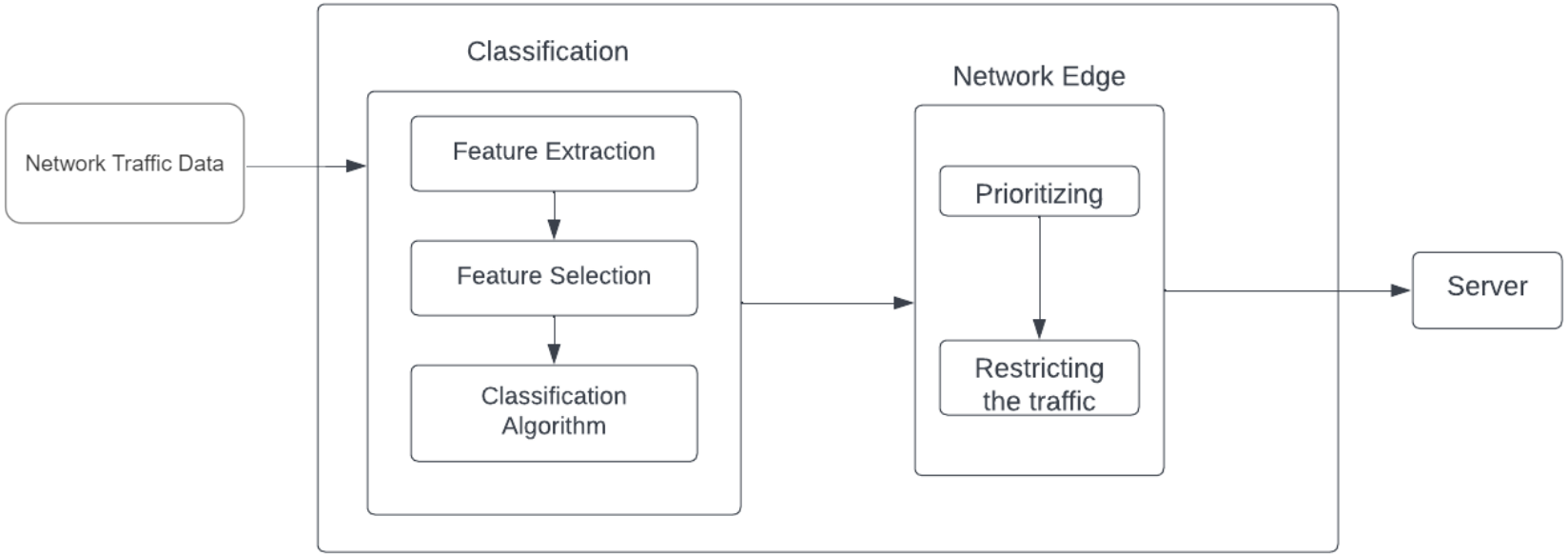
Functional Requirements

- Replaying of Pcap files
- Extracting of required features
- Apply the classification model on data
- Restrict the unwanted incoming traffic

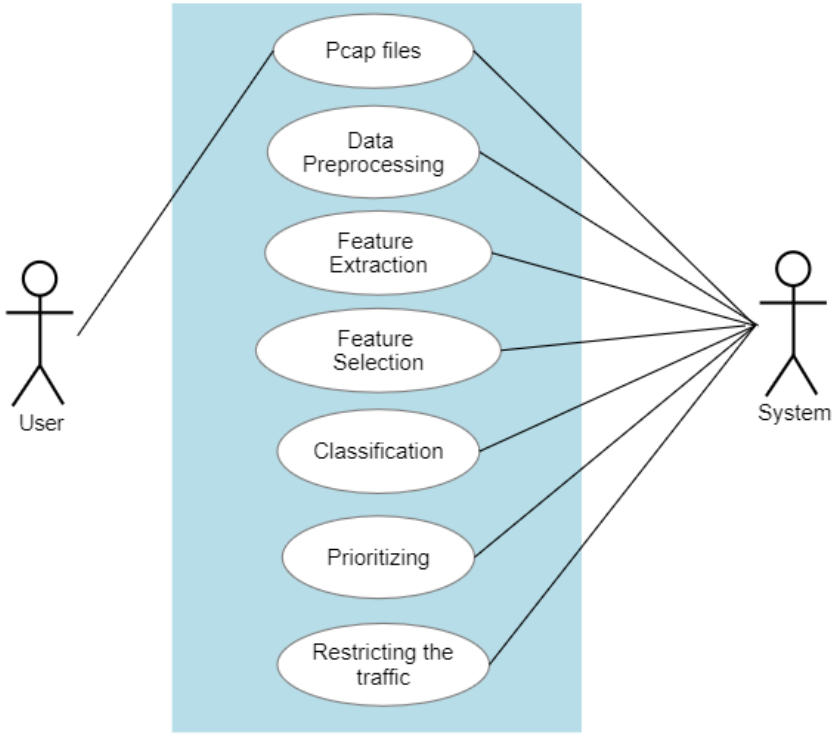
Non functional Requirements

- Scalability
- Performance
- Maintainability

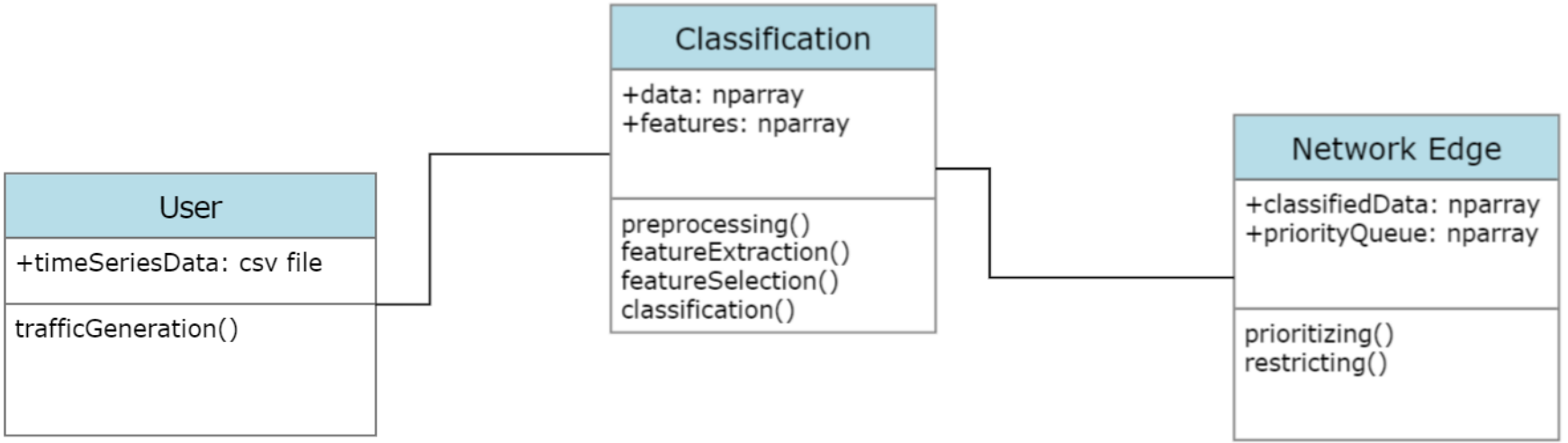
System Architecture



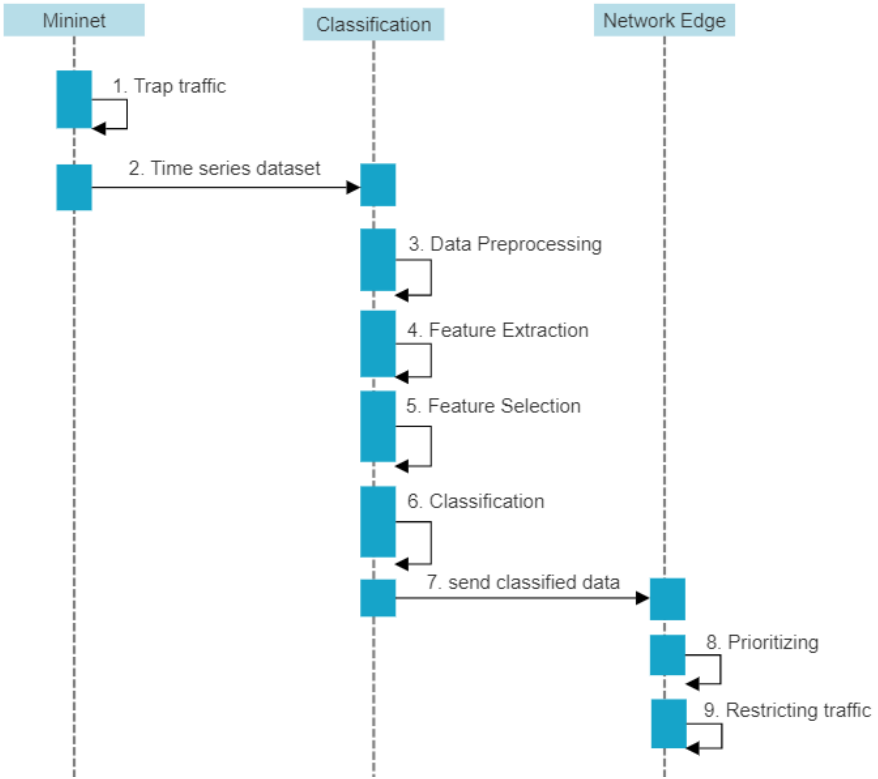
UML Diagram - Usecase Diagram



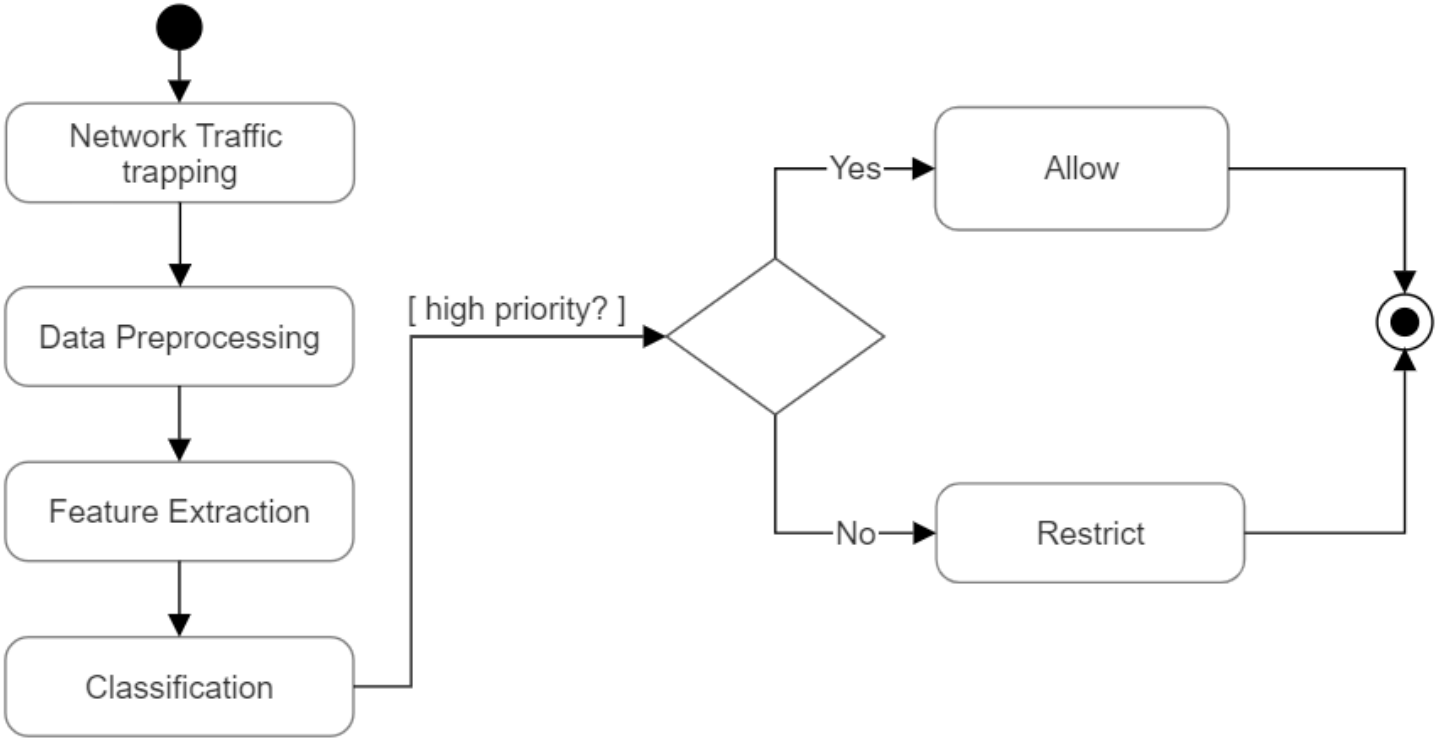
UML Diagram - Class Diagram



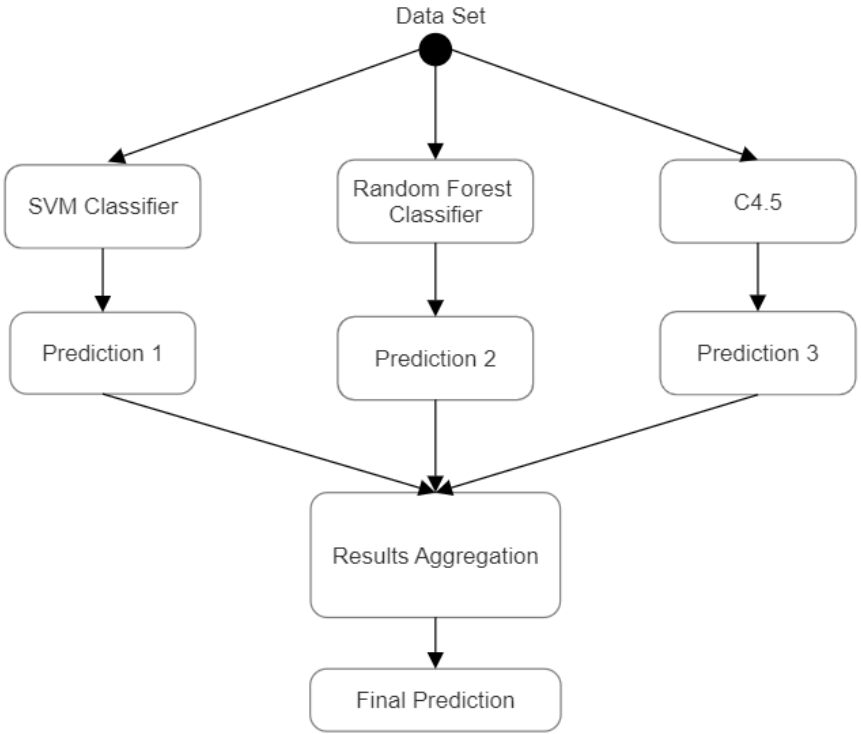
UML Diagram - Sequence Diagram



UML Diagram - Activity Diagram



Ensemble Model



Modules

- **Preprocessing Module**
  - Traffic trapping and data preprocessing is done in this module
- **Classification Module**
  - Feature extraction, feature selection and classification is done in this module
- **Prioritizing Module**
  - Prioritizing the applications using their features

Status of Work

- Required data set is collected.
- Identified algorithms used for classification - SVM, Random Forest, C4.5

Prioritizing Module

Applications

Devices -

- Principal
- HOD
- Students

High Priority -

Microsoft Teams
Google Meet
Zoom
Cisco

Low Priority -

Netflix
Amazon Prime
Aha
Hotstar
Youtube

Features

- Time
- Source IP address
- Destination IP address
- Source Port Number
- Destination Port Number
- Frame Length
- Sequence Number
- Next Sequence Number
- Acknowledgement Number
- Bytes in flight
- Check Sum
- Time Stamp
- Protocol
- Root_Shell
- Flag
- Down/Up Ratio

THANK YOU

TEAM - 09