Table of contents

Table of figures

Introduction

Chapter I
Problematic

The rise of Distributed Denial of Service (DDoS) attacks poses a significant threat to network security, making it crucial to develop effective detection mechanisms. In the context of machine learning, the challenge lies in accurately identifying DDoS attacks from normal network traffic data.

Detecting DDoS attacks presents several challenges due to the dynamic and evolving nature of these attacks. Traditional detection methods often fall short in accurately identifying DDoS attacks, leading to increased risk and potential downtime for organizations. There is a critical need for an effective and efficient DDoS detection solution that can adapt to the changing nature of attacks. Machine learning offers a promising approach to DDoS detection, thanks to its ability to analyze large volumes of data and identify complex patterns that may indicate an ongoing attack.

In this project, our objective is to develop a web application that provides users with an interface to upload their Packet Capture (PCAP) files. The application will then process these files and present the results in a tabular format, indicating whether the traffic is classified as NORMAL or as a DDoS attack.

1. What is a DDoS attack ?

A distributed denial-of-service (DDoS) attack is analogous to a group of people crowding the entry door of a shop, making it hard for legitimate customers to enter, thus disrupting trade and losing the business money.

It is a malicious attempt to disrupt the normal traffic of a targeted server, service or network by overwhelming the target or its surrounding infrastructure with a flood of Internet traffic.

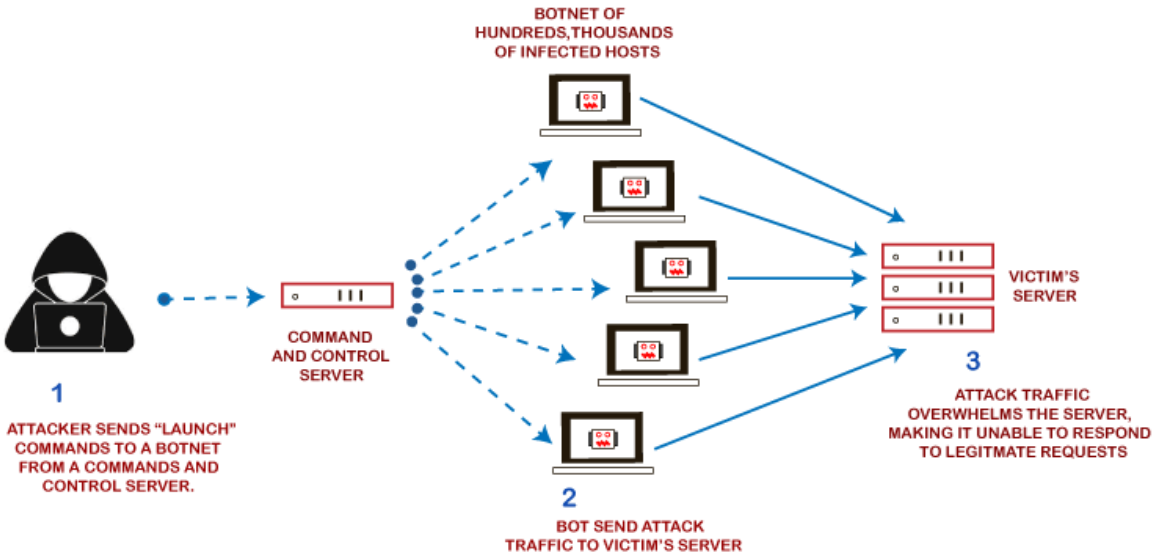

Figure I.1: DDoS attack

2. Structure of a PCAP file

A PCAP file is a binary file format that stores network traffic data. It captures packets in a structured manner, preserving the details of each communication unit traversing a network. These files are instrumental for network administrators, analysts, and cybersecurity professionals in diagnosing network issues, monitoring activities, and investigating security incidents.

The PCAP's file structure is defined by three fundamental components : the Global Header (PCAP Header), the Packet Headers, and the Packet Data.

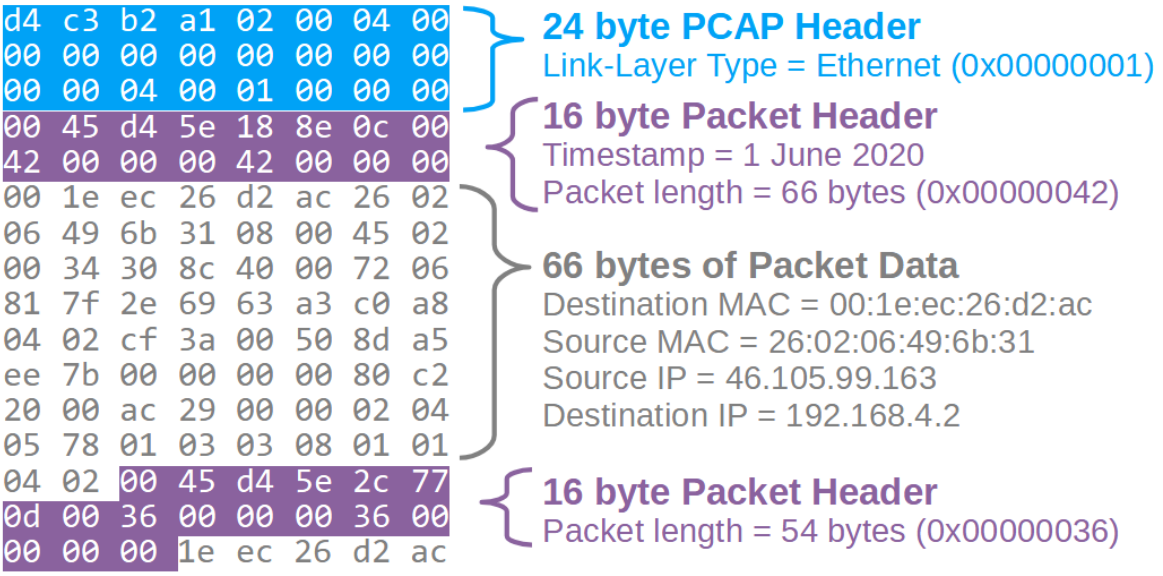

Figure I.2: PCAP file structure

Chapter II
Data preparation and analysis

1. **Description of the dataset**

Source of the dataset : Dataset link

2. **Data Preparation**
3. **Data analysis**

Chapter III
Model training

1. Random Forest algorithm

1.1. Decision Trees

A decision tree is a type of supervised learning algorithms that is used for both classification and regression tasks. Decision trees learn a series of hierarchical 'if/else' questions to classify data or predict outcomes.

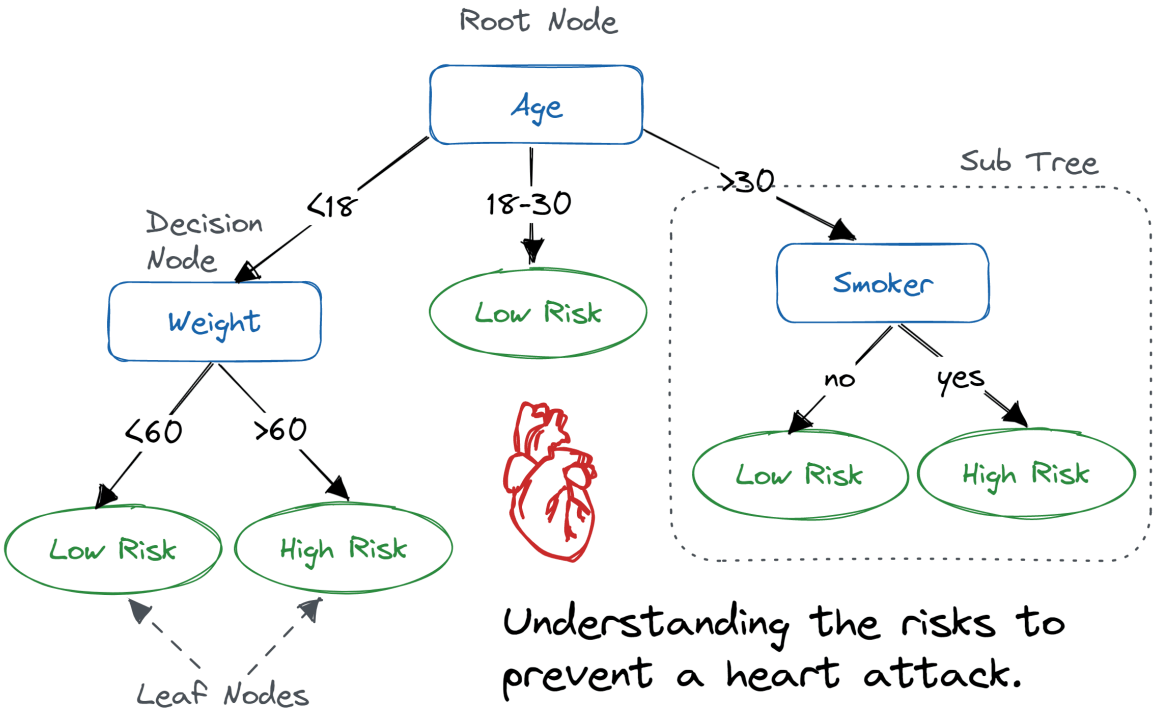

Figure III.1: Decision Tree example

1.2. Random Forests

An ensemble learning method that uses a collection of decision trees to make predictions.

It builds multiple decision trees during training and outputs the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.

2. **Feature extraction and engineering**

3. **Training the model**

Chapter IV
App Development

1. **Tools**
2. **Backend**
3. **Frontend**

Conclusion

This is a genral conclusion