



INTRODUCTION

Analyzing network traffic is one of the ways we can use to detect malware. Network traffic is stored in a pcap extension file. We can obtain the network flows which are the sequences of network packets using flowmeter such as CIC-flowmeter. The flowmeter will output a csv file containing the network flows that contains 84 features. From this, we can analyze the network traffic whether it is benign or contain some form of malware.



PROBLEM STATEMENT

- The network traffic data is huge and takes a lot of time to analyze the network for potential malware. Many of the features does little or not at all contributing to the detection of malware. Plus, many of the network flow also contains a lot of unnecessary noises of network packets that a normal user generate.

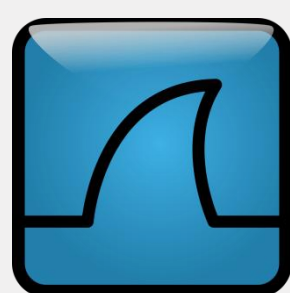


OBJECTIVES

- To classify the network traffic into benign and Malicious
- To reduce the processing time for categorizing and analyzing network packets with and without malware



TOOLS



METHODOLOGY

Obtain Preprocess
Raw Data

Dataset used is CICIDS2017 from University of New Brunswick, Canada

Features Selection

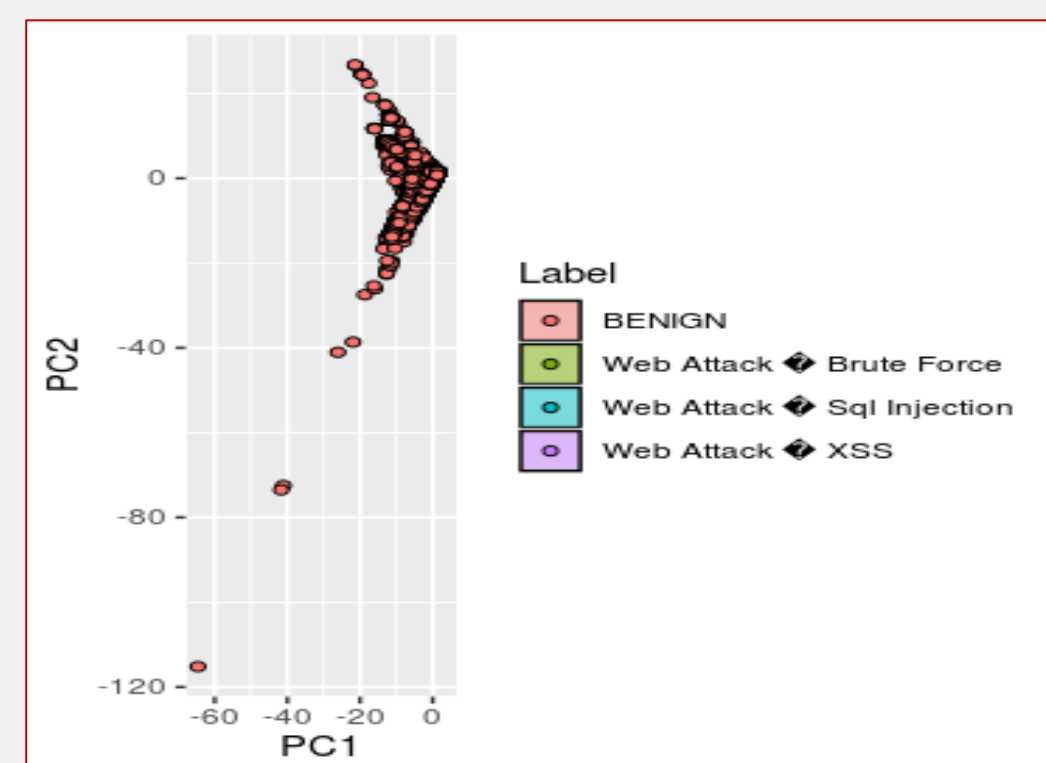
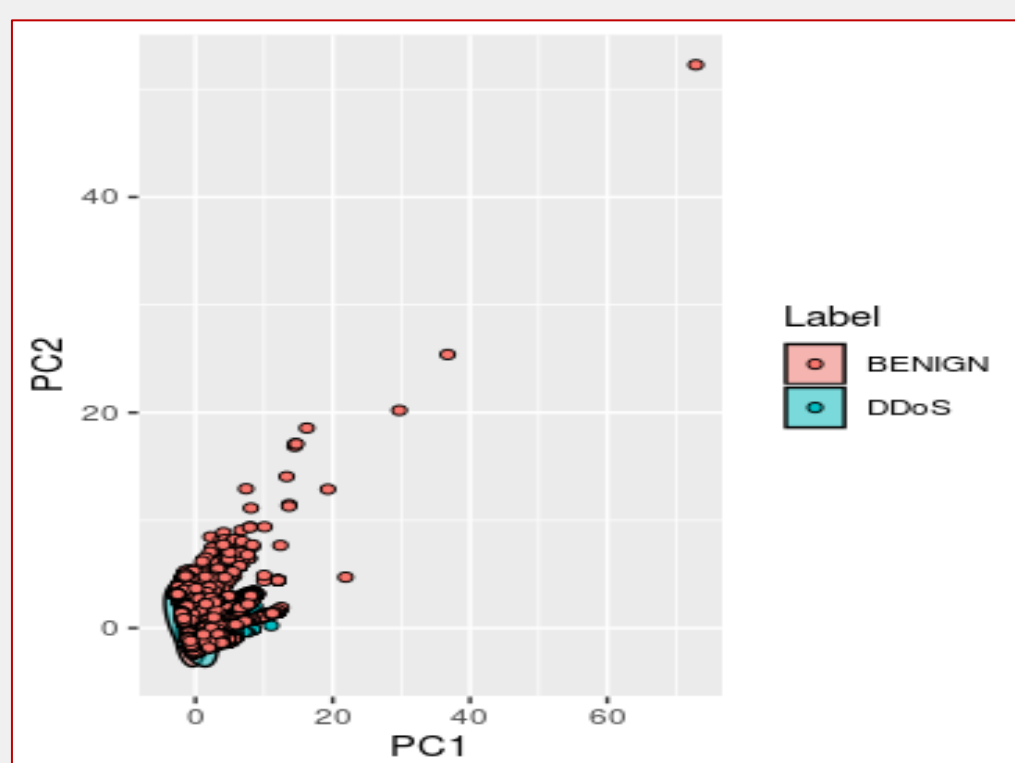
Manually selected the features from the research paper

Principal Component
Analysis

We take PCA1 and PCA2 and try to visualize how it differentiate between benign and malicious traffic



FINDINGS



From this test on the subset data that we choose, we need more than two features to differentiate between non-malicious and malicious traffic.



REFERENCES

Abdulhammed, R., Musafer, H., Alessa, A., Faezipour, M., & Abuzneid, A. (2019). Features Dimensionality Reduction Approaches for Machine Learning Based Network Intrusion Detection. *Electronics*, 8(3), 322.

Khammas, B. M., Monemi, A., Bassi, J. S., Ismail, I., Nor, S. M., & Marsono, M. N. (2015). Feature selection and machine learning classification for malware detection. *Jurnal Teknologi*, 77(1).

