Network Traffic Classification

Tiago Krebs

Abstract

Using a **classification method** this work pretended to discover if it is possible to **predict if network traffic is a DDoS attack or not** using the **CICIDS2017 Intrusion Detection Evaluation dataset.** In the end, the method reaches a precision of 99% which left us with some doubts about the variance of the dataset and the simulations used to create it.

Motivation

DDoS attacks are one of the most common malignant traffic in the past years. It is mostly used to flood unprotected online services with requests to make them unavailable and exploit for security breaches.

But because of the variance on these attacks, it is still a difficult task to identify them, especially in a distributed environment.

Dataset

The dataset used was the **CICIDS2017 Intrusion Detection Evaluation Dataset** (https://www.kaggle.com/cicdataset/cicids2017).

This dataset contains benign and common attacks. It also includes the results of the network traffic analysis using CICFlowMeter.

Note it does not contain real traffic but a simulation created by the owners. It tries to emulate the behavior of users based on the most common network protocols. Turns out this may be a problem when there is no variation enough.

Data Preparation and Cleaning

Only simple cleaning methods were necessary

- Drop null values (less than 0.1%)
- Rename columns (some column names start with space)

Research Question(s)

Can we predict whether the analyzed network traffic is a malignant DDoS attack through a simplistic classification method?

Methods

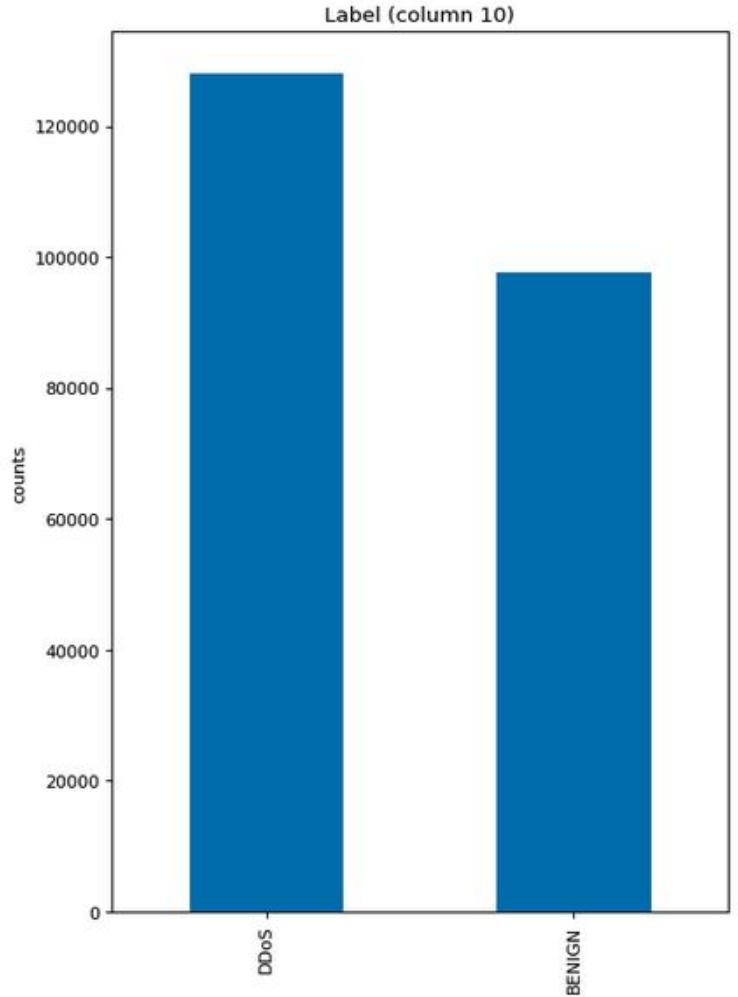
The dataset chosen has a column named `Label`. This column contains only two unique values: **BENIGN** and **DDoS**.

As my wich was identified if a traffic register can be defined as an attack or not and the dataset has this information a supervised method seem a good approach.

So, a **Decision Tree Classifier** was used based on a small set of chosen columns.

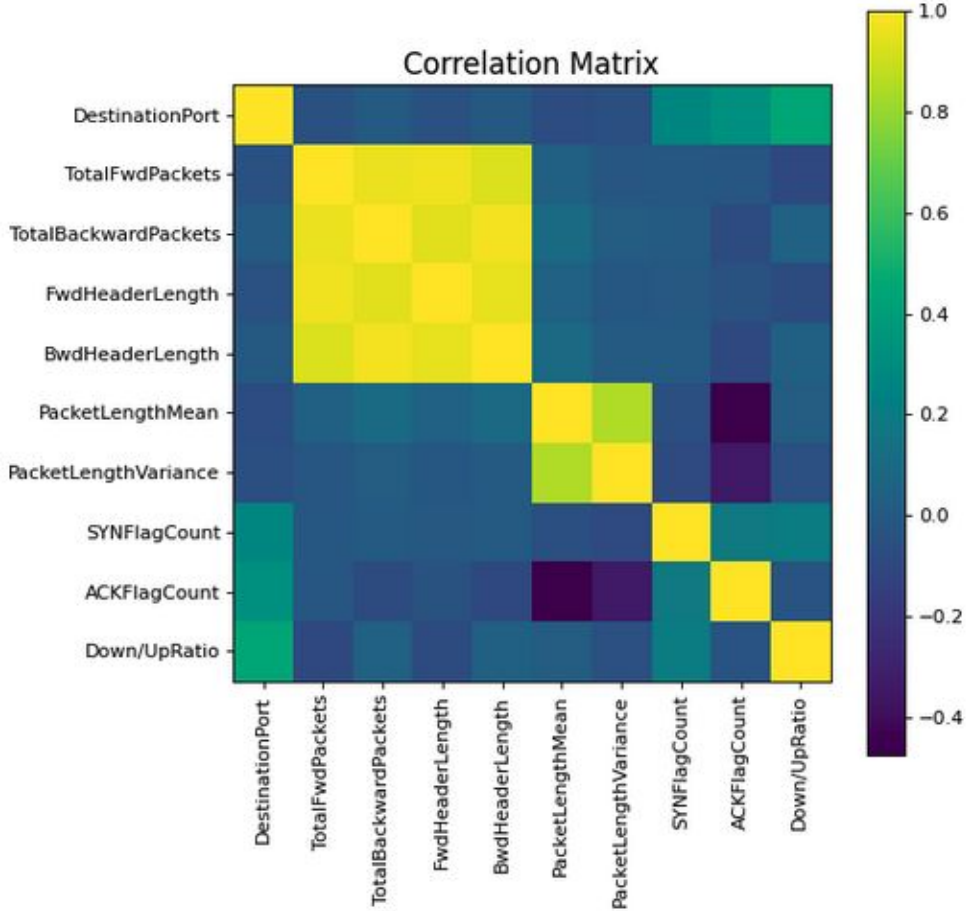
Findings

- The dataset is composed by 57% of DDoS and 43% benign traffic



Findings

- Most correlated or negatively correlated columns were founded



Findings

- Based on these most correlated columns and the Decision Tree Classifier the method reach an accuracy score of 99,4%

- This score is too good to be real so doubts raised regarding the quality of the dataset and the methods used to build it.

Limitations

The dataset description doesn’t have enough details about the methods used to simulate the traffic, needer what type of DDoS attacks were made. This can lead to a poor classifier capable to have a good accuracy only on this specific minimal dataset.

Also, there not too much open DDoS datasets available with this information available. It is a fact that to simulate and predict all types of denial attacks is quite impossible though.

Conclusions

- 99,4% of accuracy? It is too good to be true.

Some hypotheses to be answered on future works:

- Evaluate if the traffic and the DDoS generated have variation enough.
- Input some real traffic flow dataset and add DDoS variance.
- Find my error. I wasn't able to find if there is something wrong with my method.

References

https://www.kaggle.com/cicdataset/cicids2017

https://www.kaggle.com/kerneler/starter-cicids2017-3f12f887-1

https://scikit-learn.org/stable/modules/generated/sklearn.metrics.accuracy_score.html

https://scikit-learn.org/stable/auto_examples/text/plot_document_classification_20newsgroups.html#sphx-glr-auto-examples-text-plot-document-classification-20newsgroups-py