Disposition - DDOS Anomaly Detection

Blue Data (Mario Blauensteiner, Stefan Diener, Heyu David Zhang)

26/03/2021

Motivation

Distributed Denial of service (DDoS) attacks are a growing threat to Internet Service Providers (ISPs). An increasing availability of DDoS-for-hire services and of unsecured IoT devices and botnets cause these attacks to grow in in magnitude, frequency, and sophistication. With this project, we aim to test different approaches to develop a machine learning based anomaly detection algorithm, that is able to flag DDoS attacks among a stream of benign web traffic.

Data Set

We are using the ddos-dataset from Kaggle (https://www.kaggle.com/devendra416/ddos-datasets).

The data set is a collection of DDoS and "benign" webtrafic flows from different years and different DDoS traffic generation tools.

The dataset contains 12'794'627 data points, where each point corresponds to one flow, either in the forward (source to destination) or backward (destination to source) directions.

In addition to the DDoS / benign label, it contains 83 statistical features such as Duration, Number of packets, Number of bytes, Length of packets, etc. that are also calculated separately in the forward and reverse direction.

Methodology

To detect the anomalies, we aim to employ different classification models, including Naive Bayes as a more simple approach baseline model, k-Nearest Neighbors (kNN) as clustering approach, and tree-based algorithms like decision trees and random forest.

The models will then be evaluated on the common classification metrics accuracy, precision, recall and f1-score.

Complete dataset preview:

```
df %>% head(6)
```

```
##         V1                                Flow ID         Src IP Src Port
## 1: 1739476  172.31.69.25-18.219.193.20-80-37882-6  18.219.193.20    37882
## 2: 1822666     172.31.69.28-18.219.9.1-80-63287-6   172.31.69.28       80
## 3:  905739  172.31.69.28-52.14.136.135-80-63095-6  52.14.136.135    63095
## 4: 1143064 172.31.69.28-18.216.200.189-80-52341-6 18.216.200.189    52341
## 5: 1934016  172.31.69.28-18.218.55.126-80-57459-6   172.31.69.28       80
## 6: 1916750     172.31.69.28-18.219.9.1-80-56276-6   172.31.69.28       80
##           Dst IP Dst Port Protocol              Timestamp Flow Duration
## 1:  172.31.69.25       80        6 16/02/2018 11:27:29 PM          8660
## 2:    18.219.9.1    63287        6 22/02/2018 12:13:52 AM          5829
## 3:  172.31.69.28       80        6 22/02/2018 12:14:02 AM          3396
## 4:  172.31.69.28       80        6 22/02/2018 12:28:04 AM          2390
## 5: 18.218.55.126    57459        6 22/02/2018 12:19:45 AM         17362
## 6:    18.219.9.1    56276        6 22/02/2018 12:18:50 AM        201316
##    Tot Fwd Pkts Tot Bwd Pkts TotLen Fwd Pkts TotLen Bwd Pkts Fwd Pkt Len Max
## 1:            1            1               0               0               0
## 2:            4            3             935             298             935
## 3:            1            1               0               0               0
## 4:            1            1               0               0               0
## 5:            4            3             935             314             935
## 6:            4            3             935             274             935
##    Fwd Pkt Len Min Fwd Pkt Len Mean Fwd Pkt Len Std Bwd Pkt Len Max
## 1:               0             0.00             0.0               0
## 2:               0           233.75           467.5             298
## 3:               0             0.00             0.0               0
## 4:               0             0.00             0.0               0
## 5:               0           233.75           467.5             314
## 6:               0           233.75           467.5             274
##    Bwd Pkt Len Min Bwd Pkt Len Mean Bwd Pkt Len Std Flow Byts/s Flow Pkts/s
## 1:               0          0.00000          0.0000       0.000   230.94688
## 2:               0         99.33333        172.0504  211528.564  1200.89209
## 3:               0          0.00000          0.0000       0.000   588.92815
## 4:               0          0.00000          0.0000       0.000   836.82008
## 5:               0        104.66667        181.2880   71938.717   403.17936
## 6:               0         91.33333        158.1940    6005.484    34.77121
##    Flow IAT Mean Flow IAT Std Flow IAT Max Flow IAT Min Fwd IAT Tot
## 1:      8660.000        0.000         8660         8660           0
## 2:       971.500     2104.125         5260            7        5822
## 3:      3396.000        0.000         3396         3396           0
## 4:      2390.000        0.000         2390         2390           0
## 5:      2893.667     6770.578        16711            7       17355
## 6:     33552.667    67513.713       168848            6      201308
##    Fwd IAT Mean Fwd IAT Std Fwd IAT Max Fwd IAT Min Bwd IAT Tot Bwd IAT Mean
## 1:        0.000       0.000           0           0           0            0
## 2:     1940.667    3119.412        5541          46        5540         2770
## 3:        0.000       0.000           0           0           0            0
## 4:        0.000       0.000           0           0           0            0
## 5:     5785.000    9779.680       17077          38       17074         8537
## 6:    67102.667   89590.874      168848          30       32432        16216
##    Bwd IAT Std Bwd IAT Max Bwd IAT Min Fwd PSH Flags Bwd PSH Flags
## 1:       0.000           0           0             0             0
## 2:    3521.392        5260         280             0             0
## 3:       0.000           0           0             0             0
## 4:       0.000           0           0             0             0
## 5:   11559.782       16711         363             0             0
## 6:   22575.091       32179         253             0             0
##    Fwd URG Flags Bwd URG Flags Fwd Header Len Bwd Header Len Fwd Pkts/s
## 1:             0             0             32             32  115.47344
## 2:             0             0             92             72  686.22405
## 3:             0             0             20             20  294.46408
## 4:             0             0              20              20  418.41004
## 5:             0             0              92              72  230.38820
## 6:             0             0              92              72   19.86926
##    Bwd Pkts/s Pkt Len Min Pkt Len Max Pkt Len Mean Pkt Len Std Pkt Len Var
## 1:  115.47344           0           0        0.000      0.0000         0.0
## 2:  514.66804           0         935      154.125    332.3064    110427.6
## 3:  294.46408           0           0        0.000      0.0000         0.0
## 4:  418.41004           0           0        0.000      0.0000         0.0
## 5:  172.79115           0         935      156.125    333.3426    111117.3
## 6:   14.90195           0         935      151.125    330.9275    109513.0
##    FIN Flag Cnt SYN Flag Cnt RST Flag Cnt PSH Flag Cnt ACK Flag Cnt
## 1:            0            0            0            0            1
## 2:            0            1            0            0            0
## 3:            0            0            0            0            1
## 4:            0            0            0            0            1
## 5:            0            1            0            0            0
## 6:            0            1            0            0            0
##    URG Flag Cnt CWE Flag Count ECE Flag Cnt Down/Up Ratio Pkt Size Avg
## 1:            0              0            0             1       0.0000
## 2:            0              1            1             0     176.1429
## 3:            0              0            0             1       0.0000
## 4:            0              0            0             1       0.0000
## 5:            0              1            1             0     178.4286
## 6:            0              1            1             0     172.7143
##    Fwd Seg Size Avg Bwd Seg Size Avg Fwd Byts/b Avg Fwd Pkts/b Avg
## 1:             0.00          0.00000              0              0
## 2:           233.75         99.33333              0              0
## 3:             0.00          0.00000              0              0
## 4:             0.00          0.00000              0              0
## 5:           233.75        104.66667              0              0
## 6:           233.75         91.33333              0              0
##    Fwd Blk Rate Avg Bwd Byts/b Avg Bwd Pkts/b Avg Bwd Blk Rate Avg
## 1:                0              0              0                0
## 2:                0              0              0                0
## 3:                0              0              0                0
## 4:                0              0              0                0
## 5:                0              0              0                0
## 6:                0              0              0                0
##    Subflow Fwd Pkts Subflow Fwd Byts Subflow Bwd Pkts Subflow Bwd Byts
## 1:                1                0                1                0
## 2:                4              935                3              298
## 3:                1                0                1                0
## 4:                1                0                1                0
## 5:                4              935                3              314
## 6:                4              935                3              274
##    Init Fwd Win Byts Init Bwd Win Byts Fwd Act Data Pkts Fwd Seg Size Min
## 1:                -1               225                 0                 0
## 2:                -1             32768                 1                 0
## 3:                -1             32738                 0                 0
## 4:                -1             32738                 0                 0
## 5:                -1             32768                 1                 0
## 6:                -1             32768                 1                 0
##    Active Mean Active Std Active Max Active Min Idle Mean Idle Std Idle Max
## 1:           0          0          0          0         0        0        0
## 2:           0           0           0           0          0         0         0
## 3:           0           0           0           0          0         0         0
## 4:           0           0           0           0          0         0         0
## 5:           0           0           0           0          0         0         0
## 6:           0           0           0           0          0         0         0
##    Idle Min Label
## 1:        0  ddos
## 2:        0  ddos
## 3:        0  ddos
## 4:        0  ddos
## 5:        0  ddos
## 6:        0  ddos
```