Network Denial of Service Detection

-Sanjoy Basu Denver, CO

What is Network Denial of Service?

Denial of Service attack is a malicious attempt to bring down network , web based applications or services by overwhelming these resources with too much data or impairing in some other way

Impacts of Denial of Service Attack on Business

Revenue Loss:
Average cost of downtime is $5600/minute

Productivity Loss:
Critical network system are shutdown
workforce productivity comes to halt

Reputation Damage:
Brand suffers or become casualties of data
Breach

Theft:
Attacks include stolen fund consumer data and
Intellectual property

-Source: Verisign

Top 3 Industries Targeted

1.  **57%** FINANCIAL SERVICES

2.  **26%** IT SERVICES/CLOUD/SAAS

3.  **17%** TELECOM

-Source Versign

Problem statement :

Detect denial of service connections

Approach to the solution

- Anomaly detection
- Supervised machine learning to classify malicious traffic

Training Data

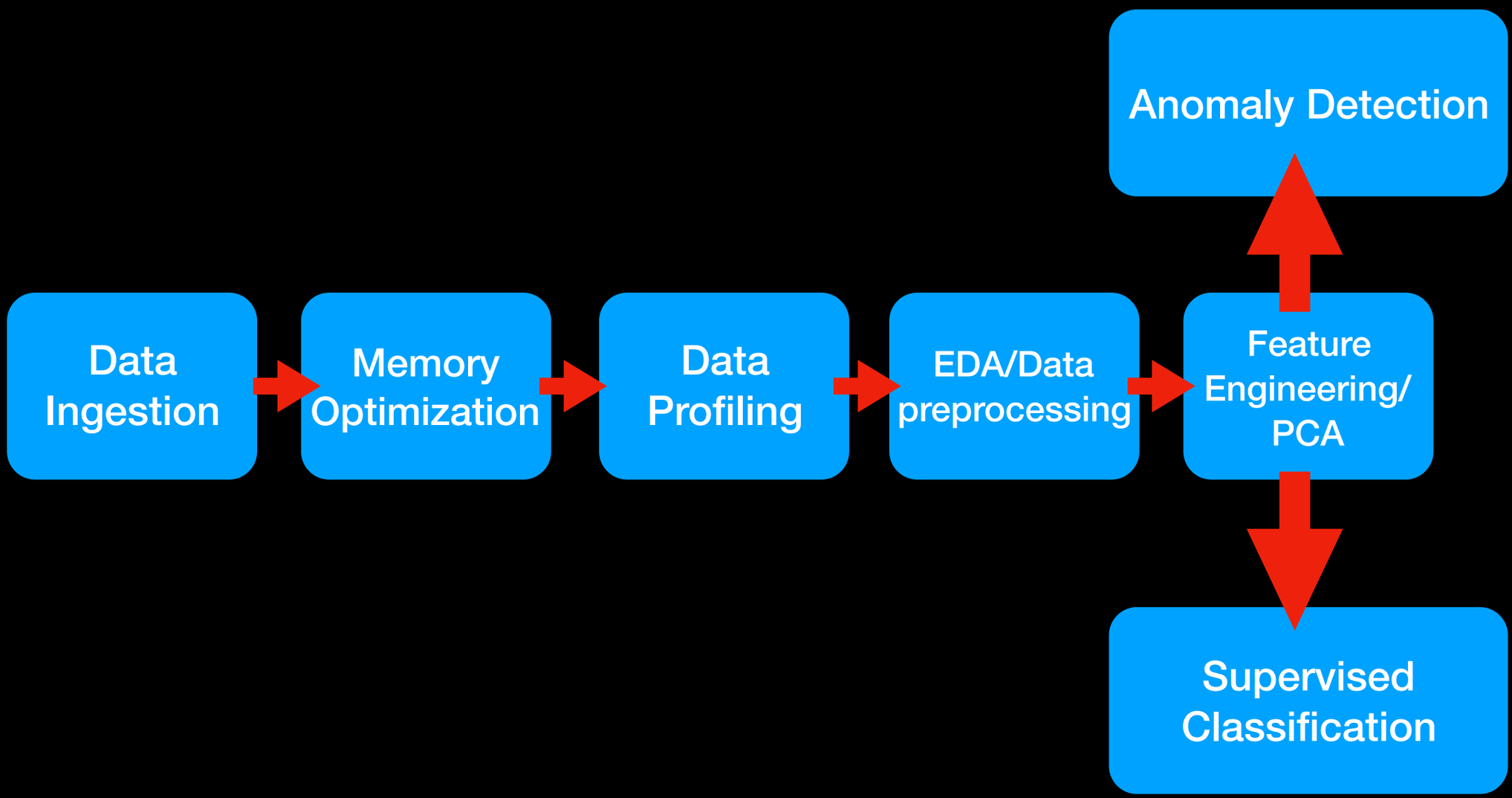
Source:

- Data was generated by MIT Lincoln Laboratory as part of cybersecurity research funded by Defense Advance Research Project Administration(DARPA)
- Data was collected on simulated US Air force Network
- Data was transformed by University of California, Irvine for analytics

Data Profile:

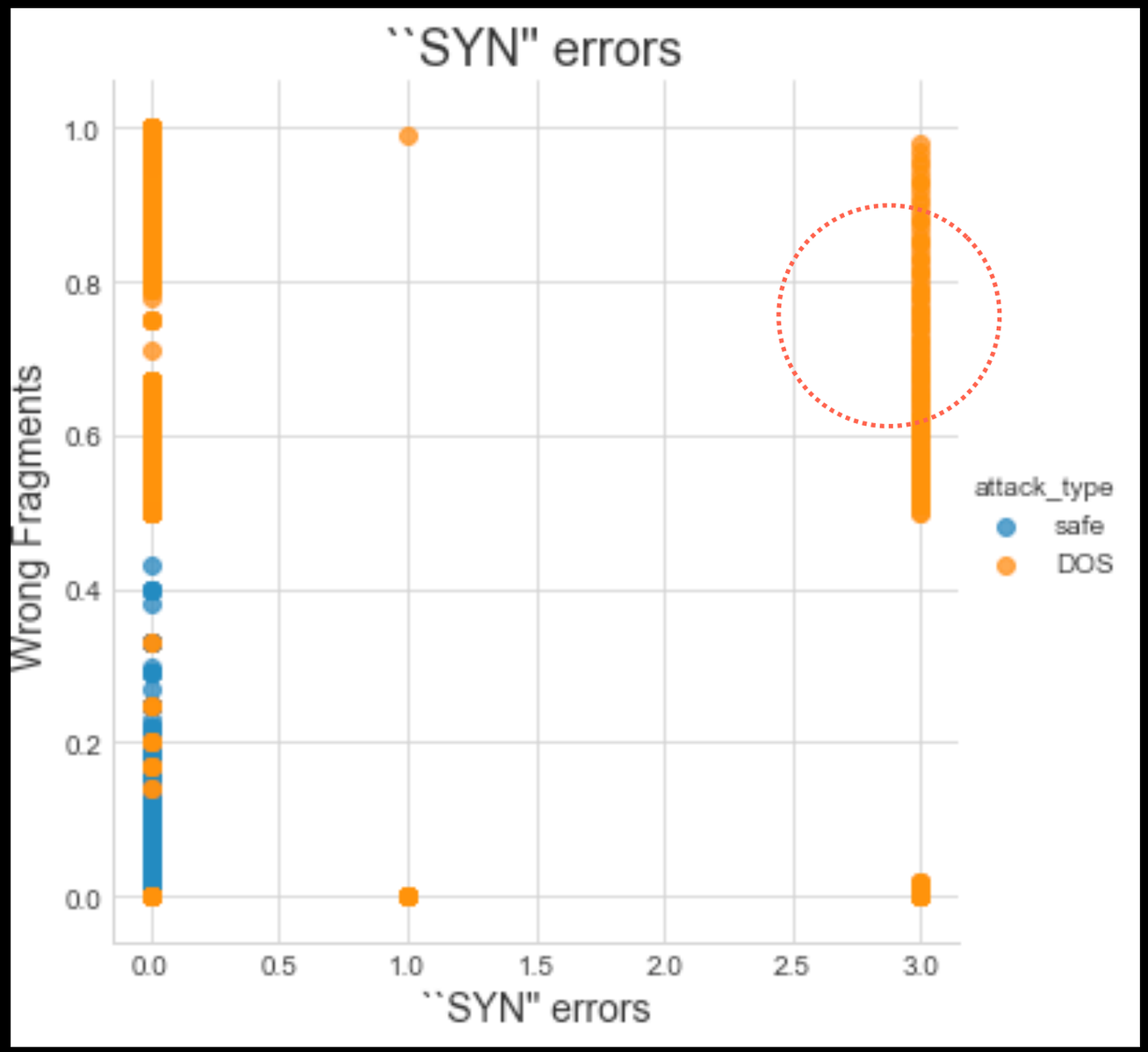
- Number of rows: >4.89 million number of columns:42
- Deep memory usage:>2.5 GB
- Missing data: 0
- Data were not labeled as DOS instead variation of DOS such as *teardrop, neptune, smurf attack* etc

Data Pipeline

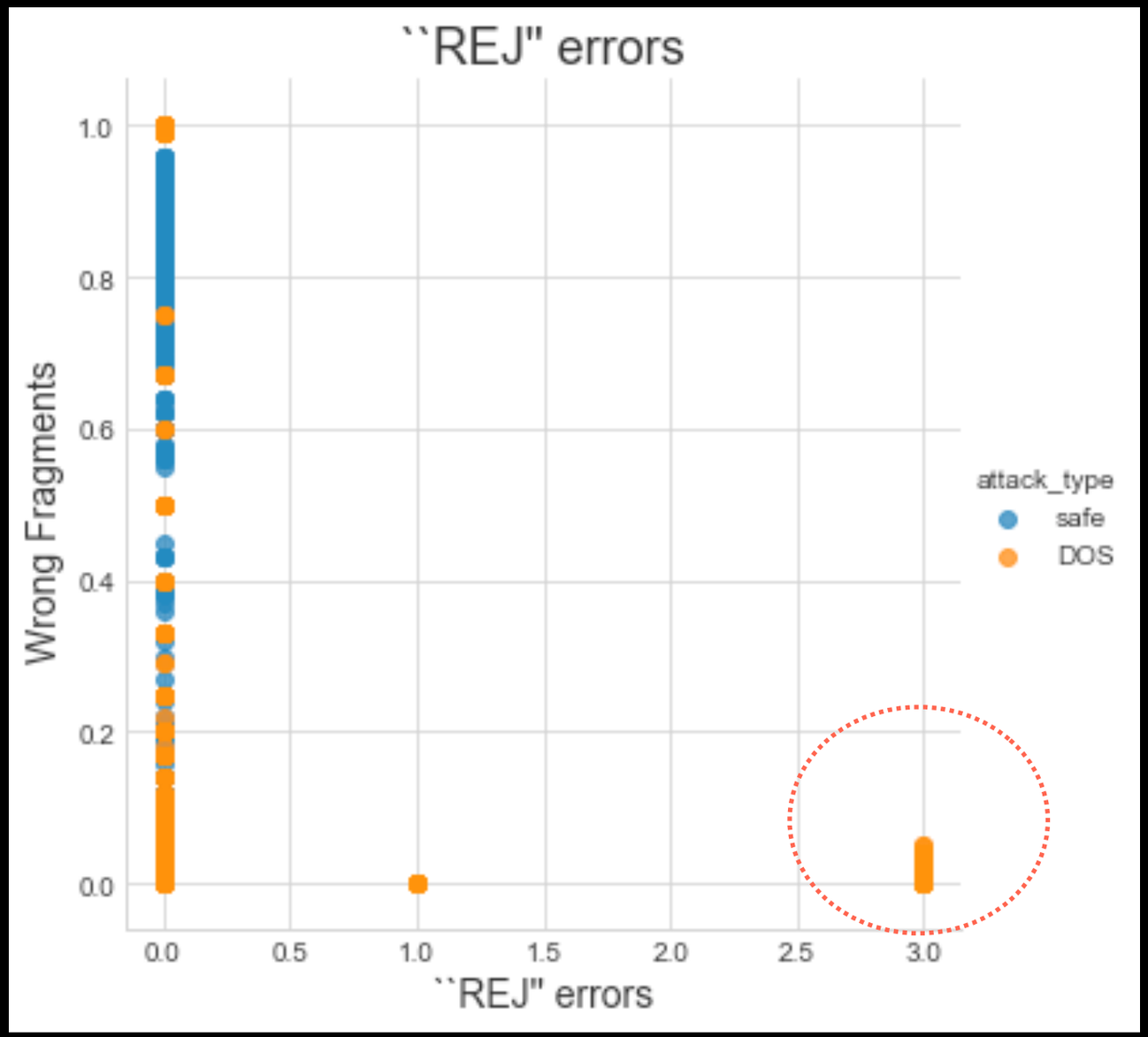


Exploratory Data Analysis
&
Visualization

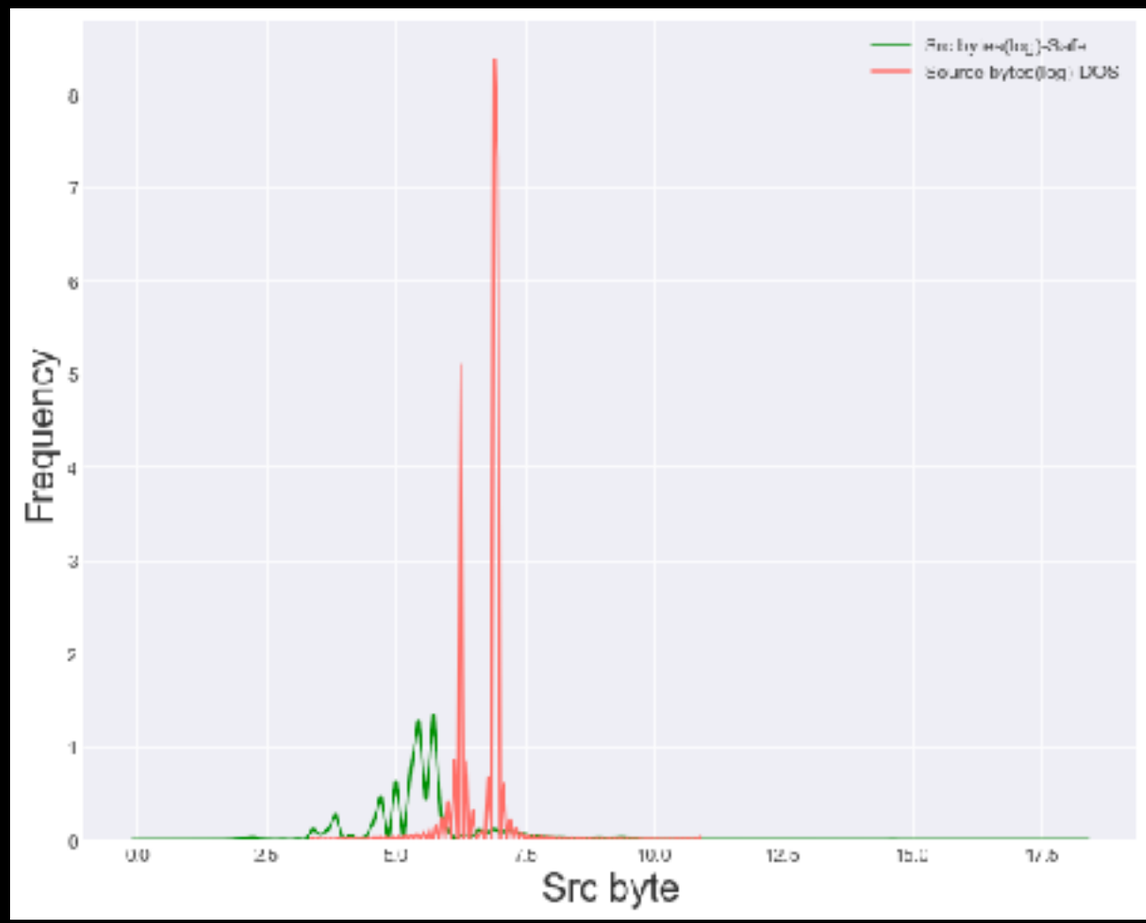
SYN Error vs Wrong Fragments

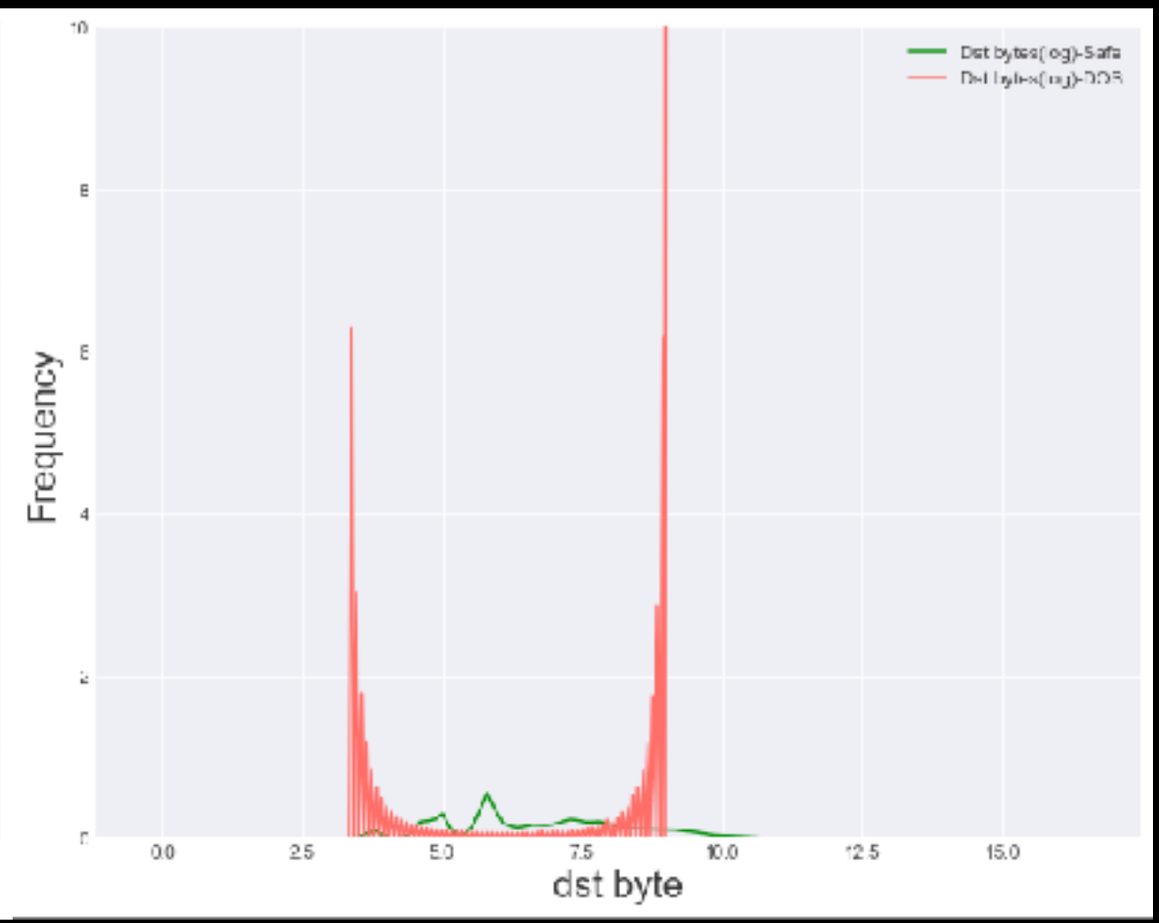


REJ Error vs Wrong Fragments

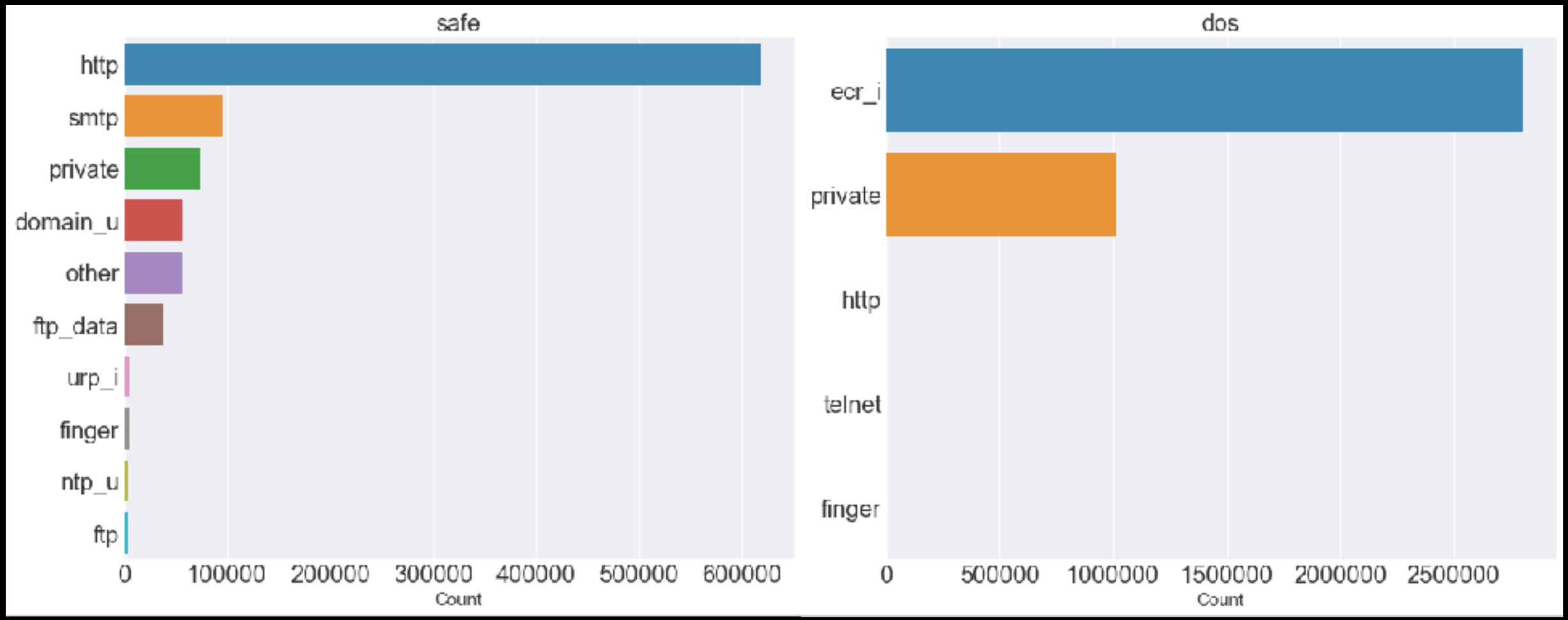


Distribution of Source and Destination bytes size

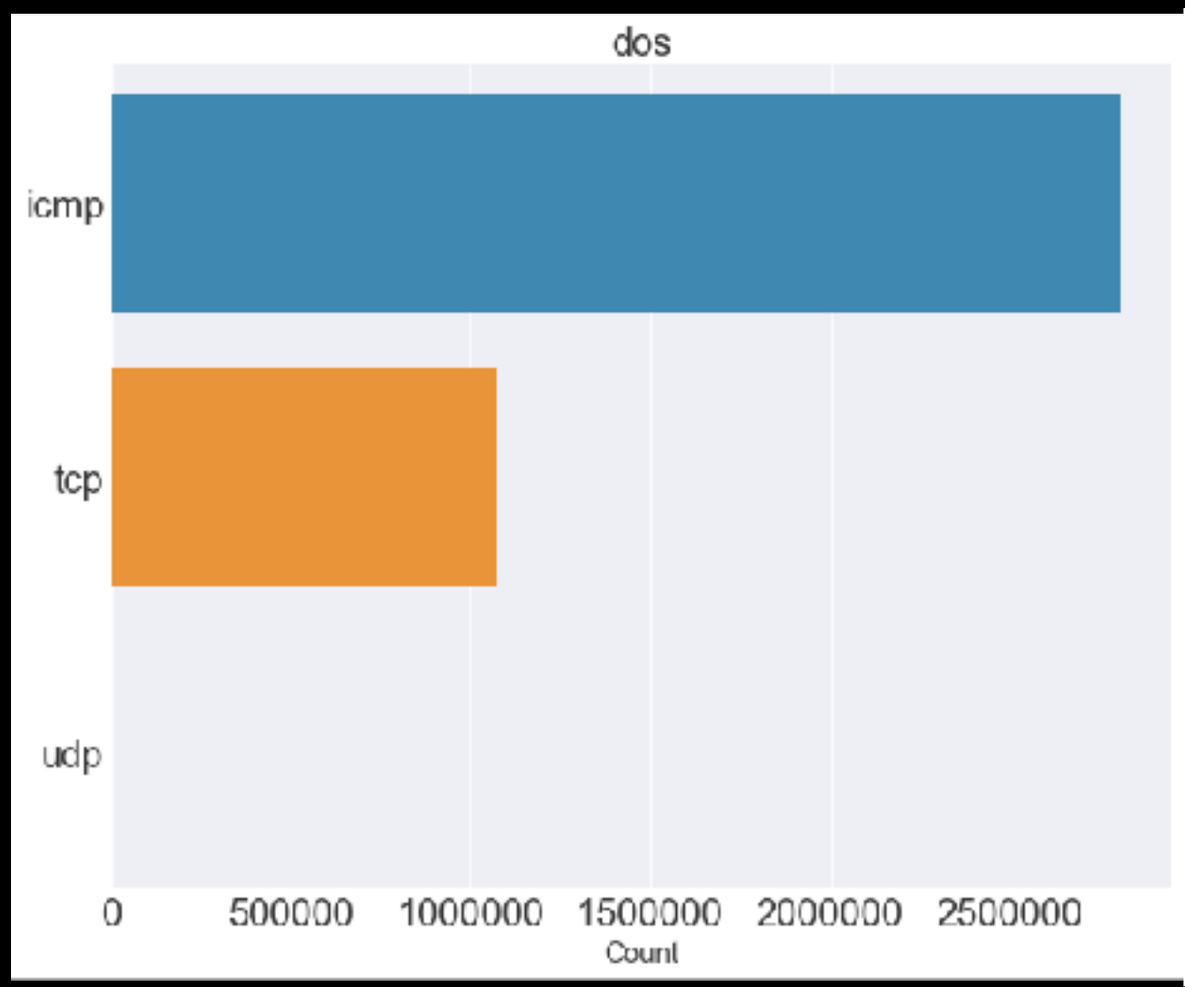


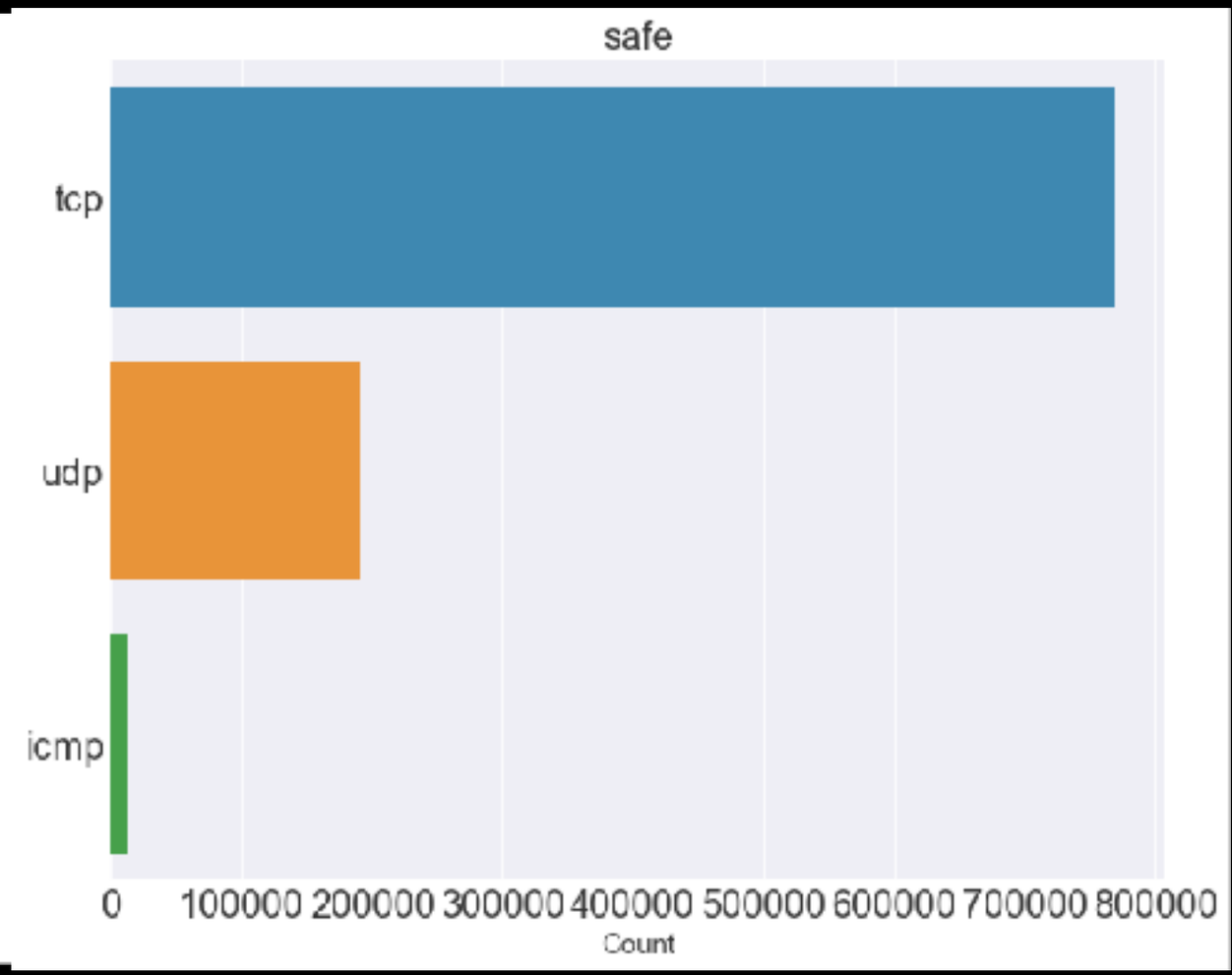


Service type malicious vs safe

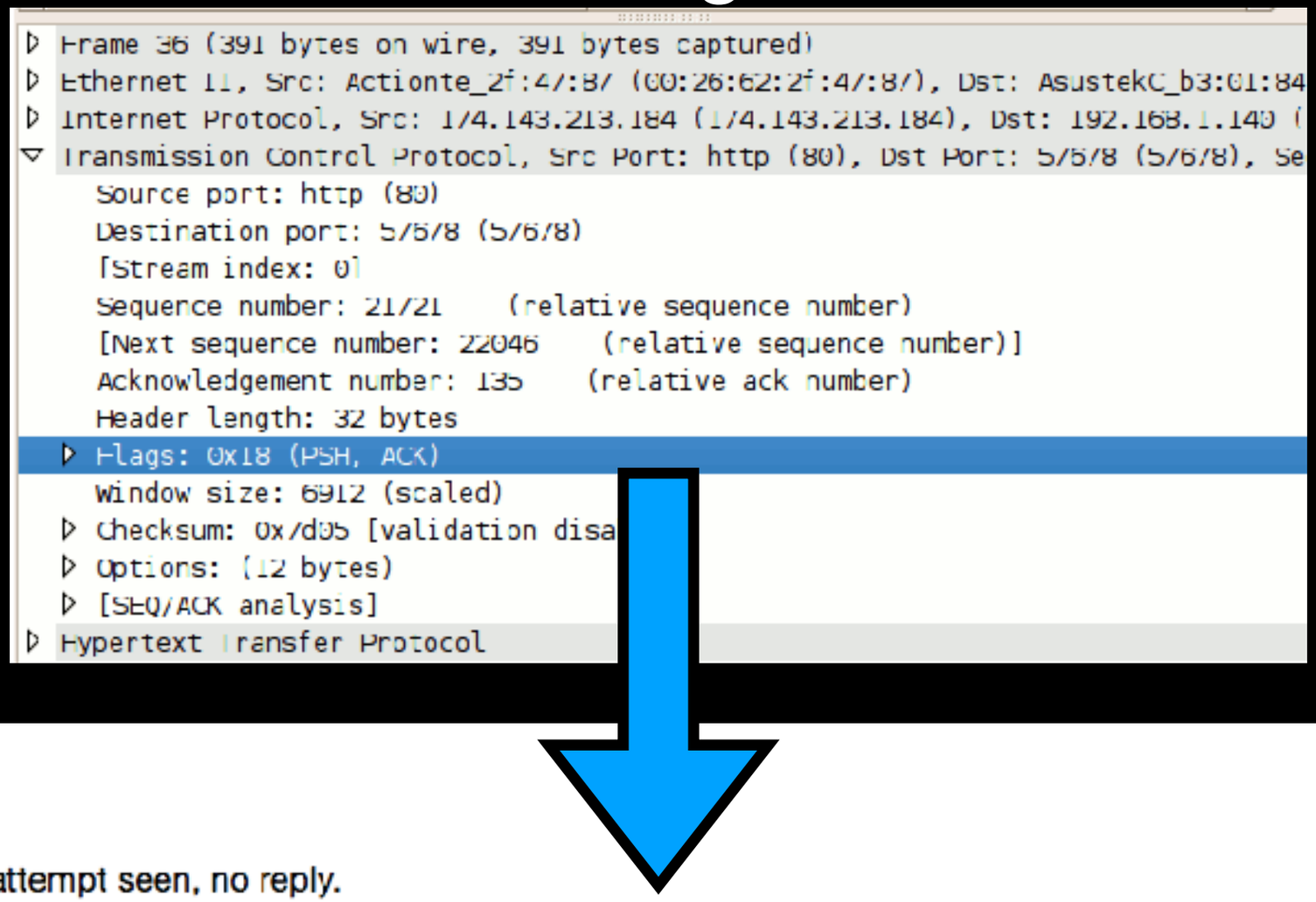


Protocol malicious vs safe





TCP Flags



TCP Flags

S0 Connection attempt seen, no reply.

S1 Connection established, not terminated.

SF Normal establishment and termination. Note that this is the same symbol as for state

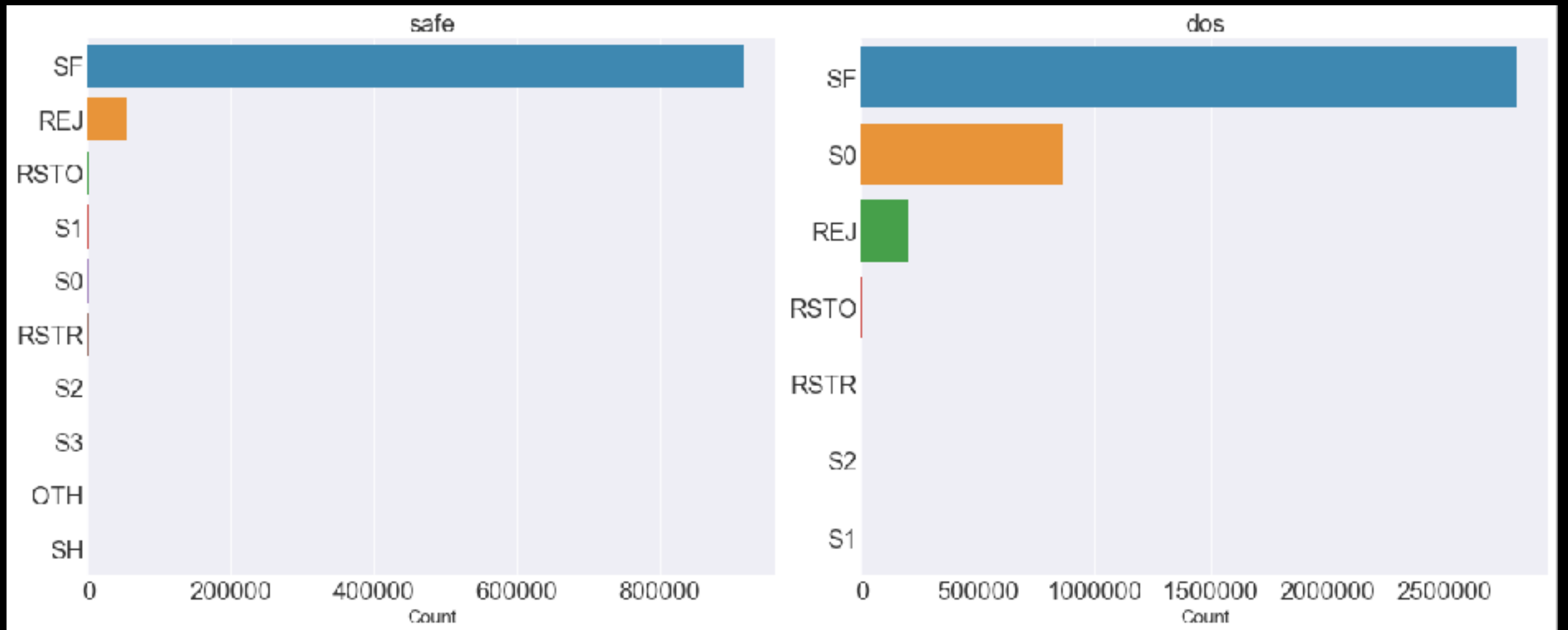
S1. You can tell the two apart because for S1 there will not be any byte counts in the summary, while for SF there will be.

REJ Connection attempt rejected.

S2 Connection established and close attempt by originator seen (but no reply from responder).

S3 Connection established and close attempt by responder seen (but no reply from originator).

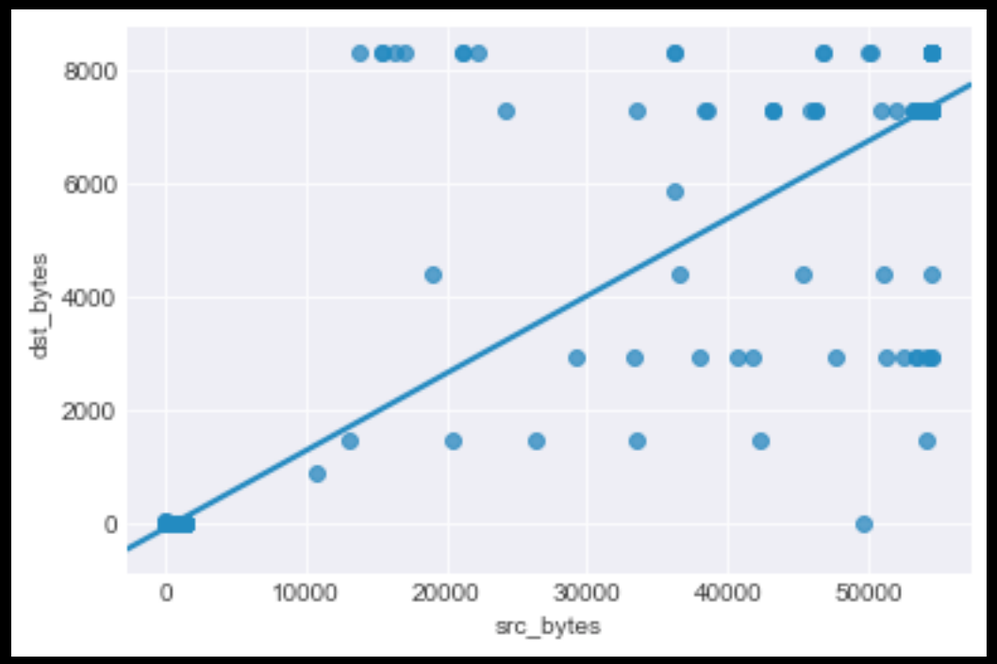
TCP Flags



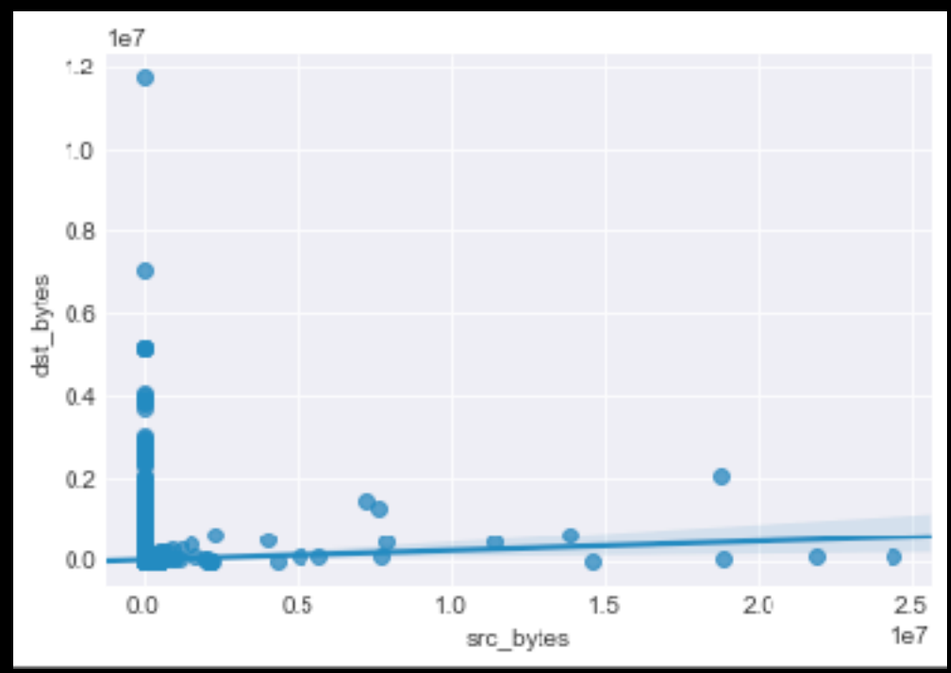
	Safe	DOS
SF	0.9428	0.7239
S0	0.0004	0.2233

Relational plot between dst_bytes and src_bytes

Malicious



Safe

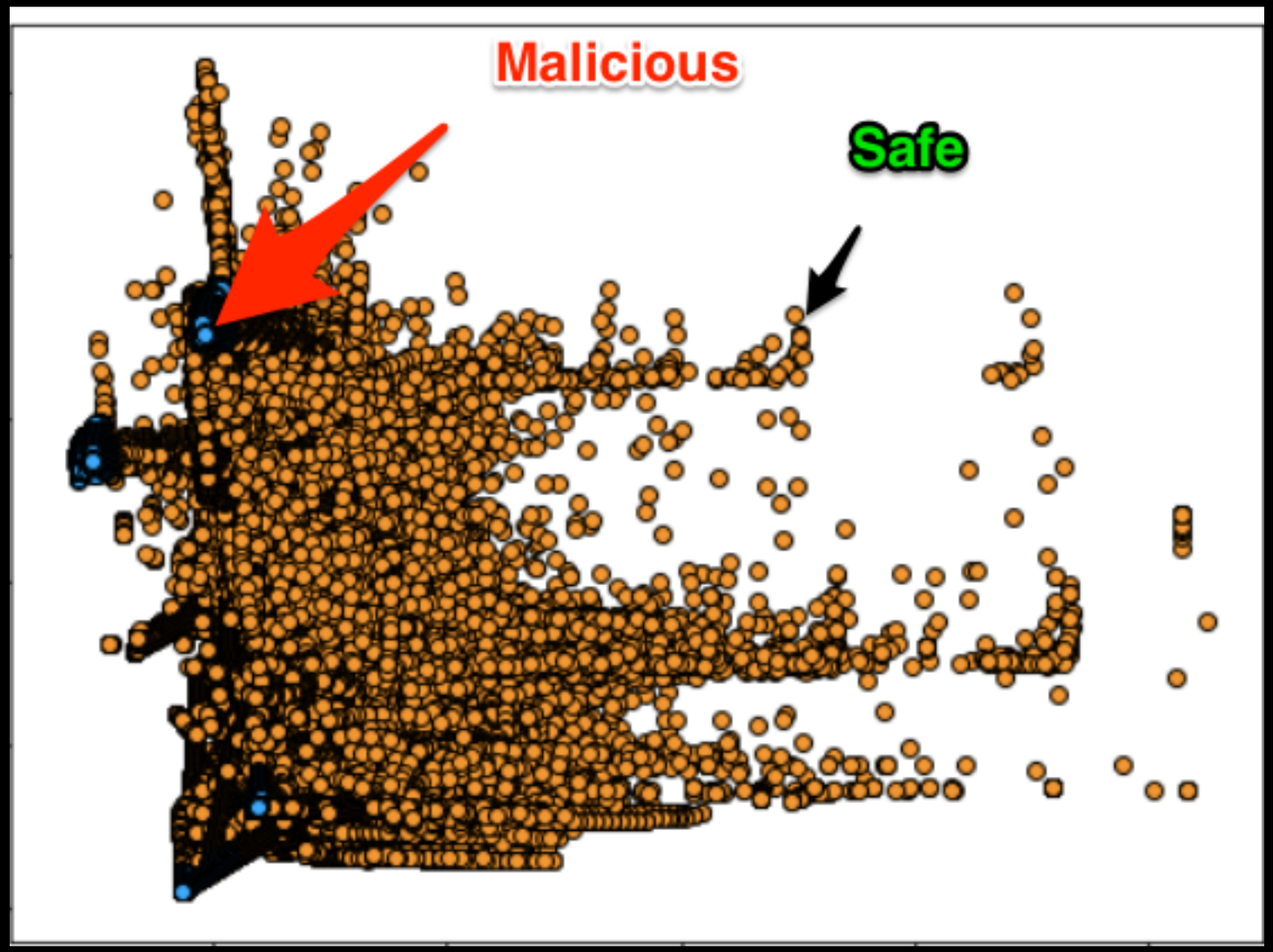


Insight from exploratory data analysis

- Malicious connections are likely to have both wrong fragments and SYN error
- Malicious connections are likely to have both wrong fragments and REJ error
- ecr_i types of service are more likely to be malicious
- Malicious connections are most likely to have flag S0 set indicating “connection attempt seen but no reply”
- Safe packets either has high destination packet size OR high source packet size but never both
- Malicious packets may have high destination packet size AND high source packet size and they may occur simultaneously
- 72.39% of malicious packets have TCP flags set to SF indicating normal establishment and termination of connection
- Packets with ZERO SYN or REJ error can be malicious

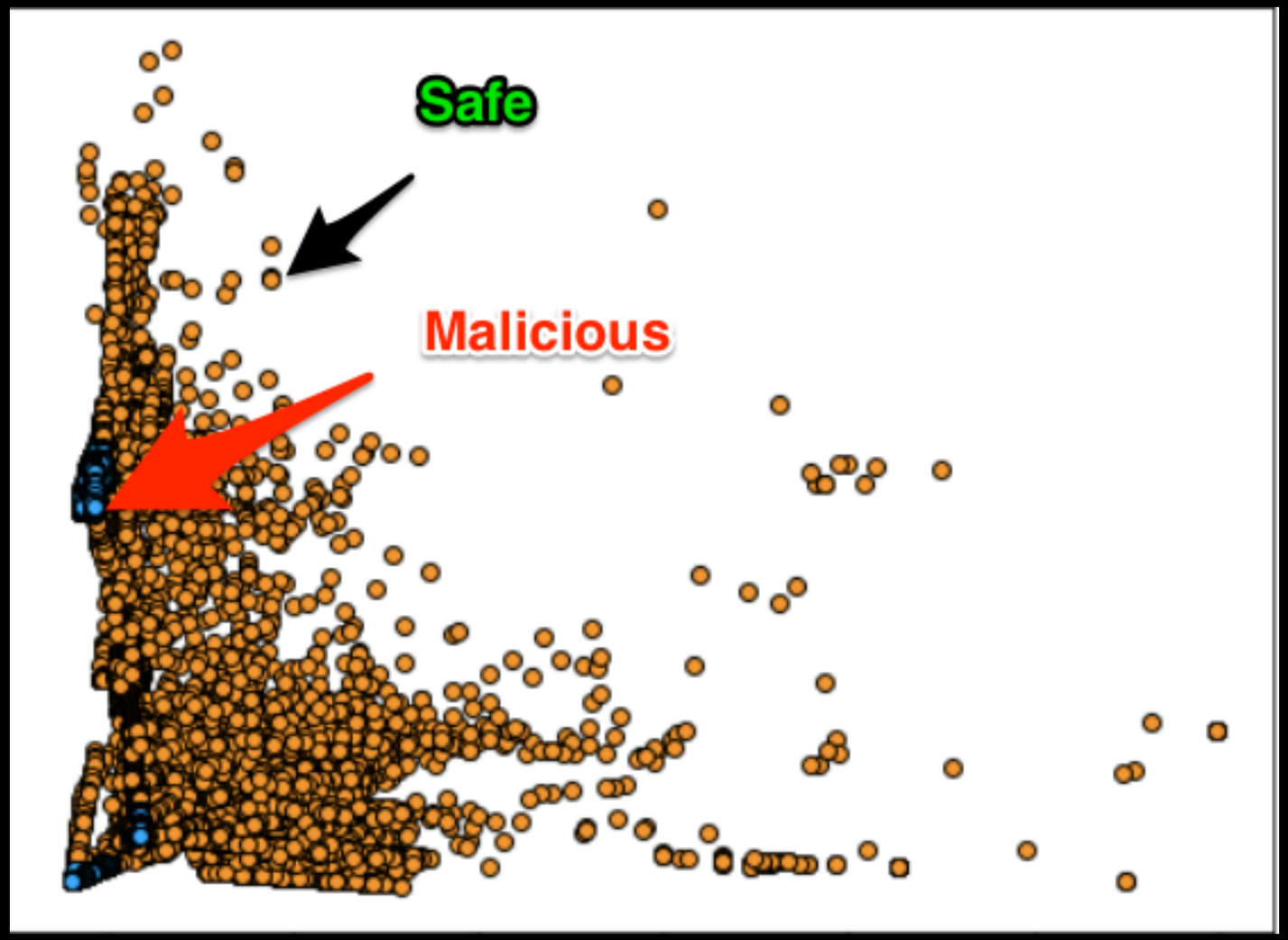
Anomaly Detection
Using
Gaussian Mixture Model

Anomaly Detection on training data



	Detection
Malicious	0.9783
Safe	0.8148

Anomaly Detection on validation data



	Detection
Malicious	0.959
Safe	0.7149

Supervised Machine Learning

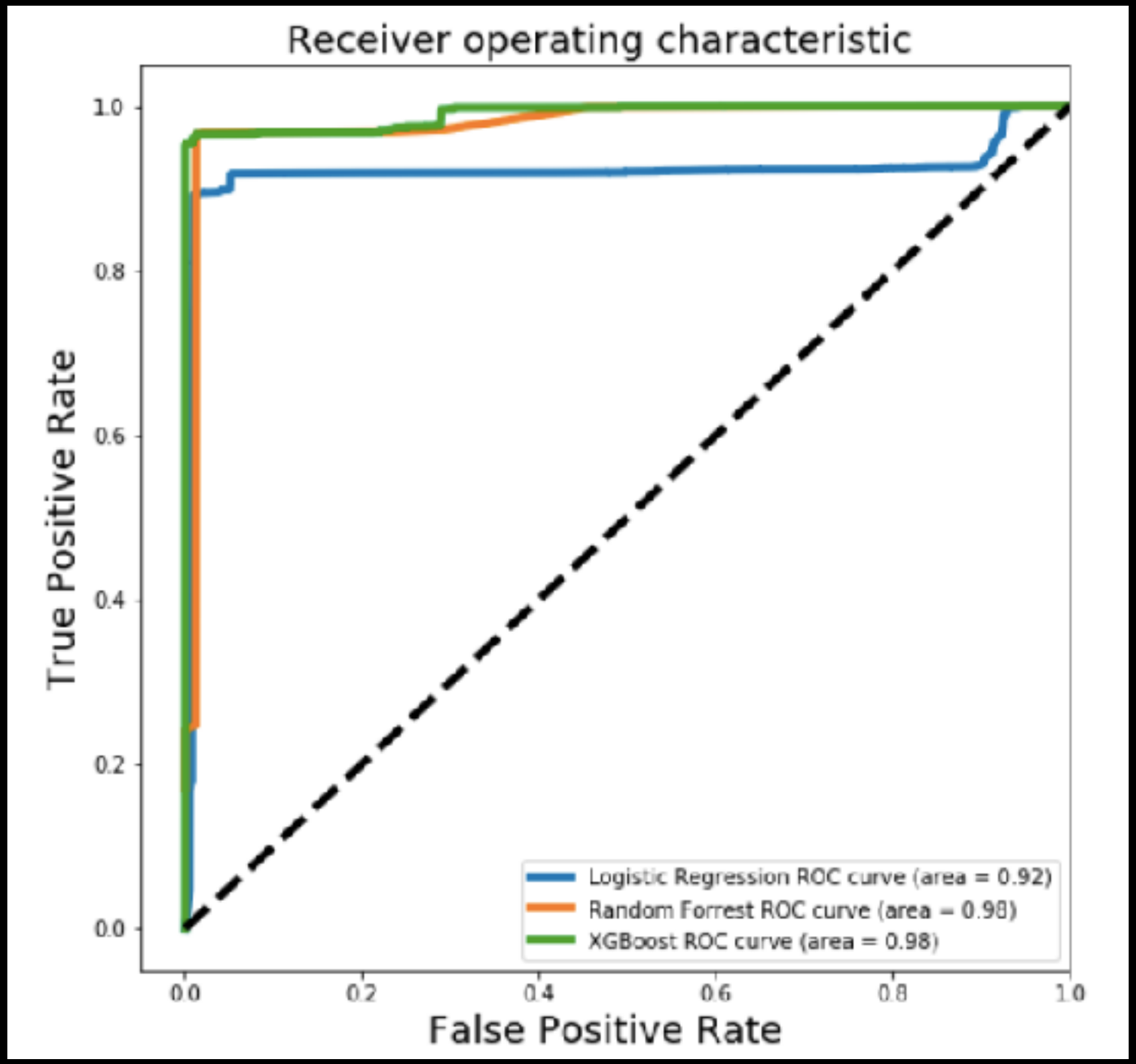
K-Fold Classification Models Accuracy Train/Test

	1	2	3
Logistic Regression L1	0.9983	1	0.9953
Logistic Regression L2	0.9982	1	0.9954
Random Forest	0.9986	1	0.8734
XGBoost	0.9986	1	0.9413

Validation data

- Validation data is completely different data set
- Validation data has additional variation of Denial of Service connection not seen during training and testing
- The data set was never used either entirely or partially during model training
- Validation data set was ingested only after model was trained and persisted (pickled)

Receiver Operating Characteristics



Questions