DDOS/DOS Attack Detection Using Machine Learning

Raz Saad , Maor Or , Chen Dahan

February 2024

1. Abstract

The security of the internet is seriously threatened by distributed denial of service (DDoS) attacks. The purpose of a DDoS assault is to disrupt service and prevent legitimate users from using it by flooding the central server with many messages or requests that will cause it to reach its capacity and shut down. Because it is carried out by numerous bots that are managed (infected) by a single botmaster using a fake IP address, this assault is dangerous because it does not involve a lot of work or special tools. For the reasons presented above, this paper will discuss the improvement of previously explored techniques of handling these kinds of attacks, from the paper ”DDOS Attack Identification using Machine Learning Techniques”. As of our improvement techniques, we have considered different algorithms and articles for extracting and deciding on best features for our models. Also, we have expanded on the given data at hand, added from the CIC-IDS2017 an additional relevant portion for our problem. As a result of these innovations, we have managed to improve the scores of identifying’s and detecting’s.

2. Introduction

In recent years, Distributed Denial of Service (DDoS) attacks have caused significant financial losses to industry and governments worldwide, as shown in information security reports [1]. These records are in line with the growing number of devices connected to the Internet, especially driven by the ever-rising usage of computers and mobile. Most of the attacks are targeted to bring down the targeted server and can also lead to a temporary shutdown of the services the organization has to offer. Thus, DDoS attacks may result in the loss of the reputation of an organization, because of the organization's inability to provide service to its customers. To prevent these types of attacks, a detection system needs to be in the organization's network security system to categorize the incoming requests, and separate between the benign traffic of clients' usage and the malicious incoming DDoS attack traffic.

3. Related work

- Subhashini Peneti, Hemalatha E ” DDOS Attack Identification using Machine Learning Techniques” [2]:
  This is the original article that we are improving upon, we explained our methods differences in the workflow section.

- Chandra Sekhar Reddy N, Vemuri P, Govardhan A “An empirical study on support vector machines for intrusion detection”. International Journal of Emerging Trends in Engineering Research (2019) [3]:
  This article has used SVM model for detecting various attacks, DDoS among such. We are using stronger models (Random Forest, XGBoost). Another difference is that our dataset is more recent and so is more relevant.

- Zachariah Pelletier, Munther Abualkibash "Evaluating the CIC IDS-2017 Dataset Using Machine Learning Methods and Creating Multiple Predictive Models in the Statistical Computing Language R" [4]:
  This paper has explored the CIC-IDS2017 dataset and learned about the features in there to extract the top features. We have used their conclusions to train our models better.

- Reis, B., Maia, E., Praça, I. (2020). Selection and Performance Analysis of CICIDS2017 Features Importance. In: Benzekri, A., Barbeau, M., Gong, G., Laborde, R., Garcia-Alfaro, J. (eds) Foundations and Practice of Security. FPS 2019. Lecture Notes in Computer Science(), vol 12056. Springer, Cham. https://doi.org/10.1007/978-3-030-45371-8_4 [5]:
  This paper has also explored the CIC-IDS2017 dataset and learned about the features in there to extract the top features. We have used their conclusions to train our models better. This paper has provided the best results for features that helped improving the training of the models.

4. Achieved Contributions

Our first step was to follow the original paper's methods, for us to verify the article's presented results. It is important to note that our results did not match the originals, while fully following and copying the methods, meaning the presented results of the original paper might not be as authentic as it should. Our next step was to try improving the results by changing models and features, using recommended rankings of features importance, explored by different articles. Lastly, we have added and experimented on a portion of our dataset - in the original paper, only the CSV file named "Wednesday workingHours" from the dataset was used, while on our work, we have added "Friday WorkingHours Afternoon DDos" to have a better training of our models.

5. Evaluation

The proposed system is to develop a smart detection system which has a built-in classification model. The first part is to build a classification model. It involves defining a signature for each sample. After the signatures are extracted, the records are classified based on their signature. Features are selected by previous works on this dataset for building signature. Once the classification models are built, they are compared, and the best model is chosen.

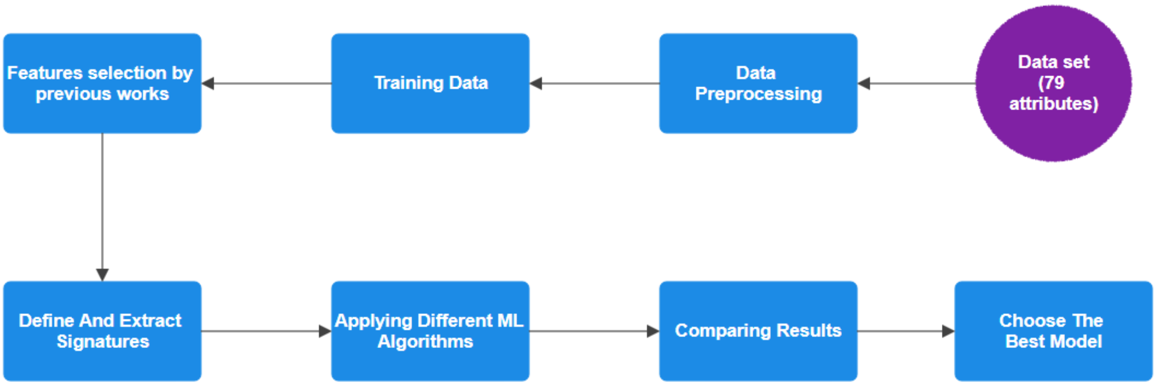
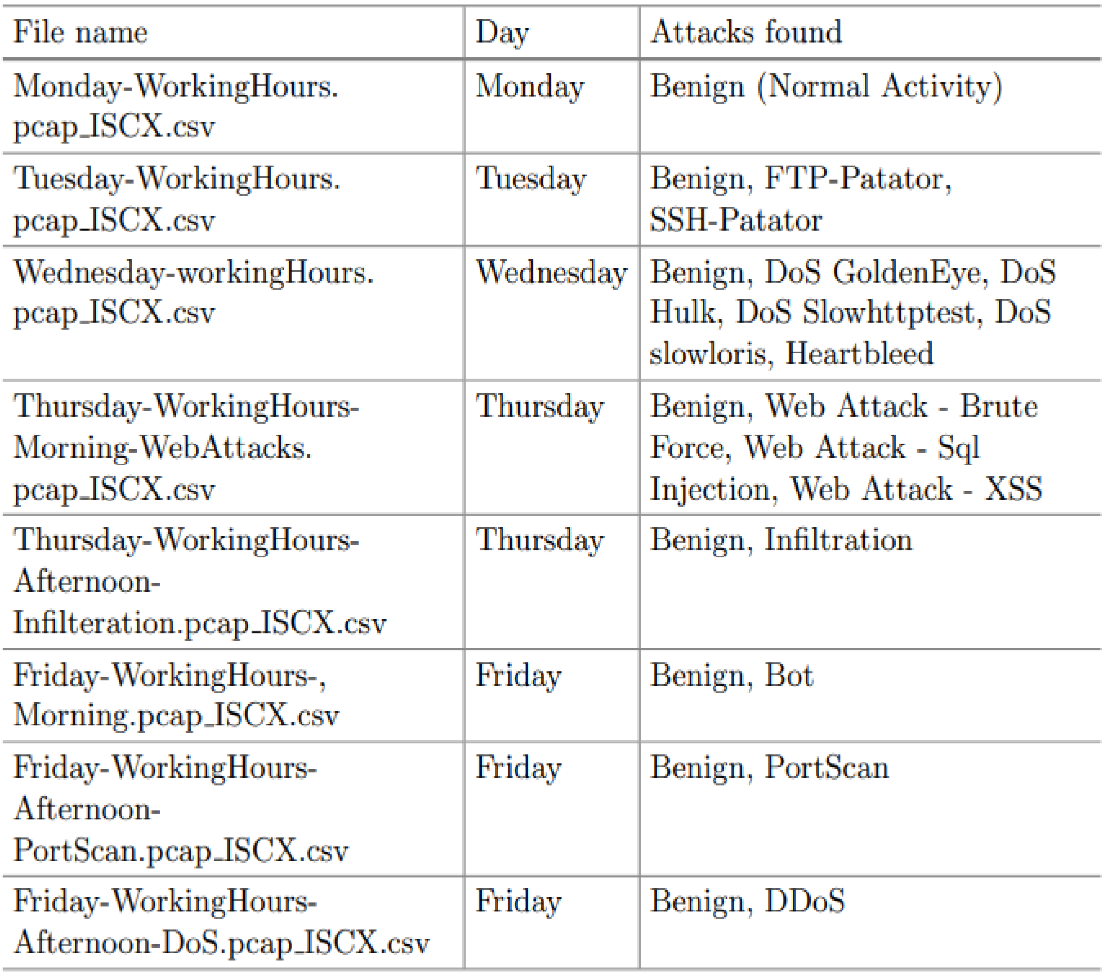

Features selection by previous works
Training Data
Data Preprocessing
Data set (79 attributes)
Define And Extract Signatures
Applying Different ML Algorithms
Comparing Results
Choose The Best Model

Figure 1: Solution Architecture

5.1 Dataset

The dataset used for this paper is the CIC-IDS2017 dataset. According to the author of CIC-IDS-2017 [6], the dataset spanned over eight different files, containing five days of normal and attack traffic data of the Canadian Institute of Cyber security. A short description of all those files is presented in Figure 2. We focused on two parts of the dataset, the files that hold the data of traffic for days Wednesday and Friday, that are relevant to our classification problem (containing DDOS/DOS attacks). The data set consists of 78 different features and a label that is the class name.
There are totally 691,406 records, out of which, 553,124 records are for training and 138,282 for testing. After our previously explained expanding of the dataset, we have grown our dataset to a total of 917,117 records and split it to train/test accordingly.



File name	Day	Attacks found
Monday-WorkingHours. pcap_ISCX.csv	Monday	Benign (Normal Activity)
Tuesday-WorkingHours. pcap_ISCX.csv	Tuesday	Benign, FTP-Patator, SSH-Patator
Wednesday-workingHours. pcap_ISCX.csv	Wednesday	Benign, DoS GoldenEye, DoS Hulk, DoS Slowhttptest, DoS slowloris, Heartbleed
Thursday-WorkingHours-Morning-WebAttacks. pcap_ISCX.csv	Thursday	Benign, Web Attack - Brute Force, Web Attack - Sql Injection, Web Attack - XSS
Thursday-WorkingHours-Afternoon-Infilteration.pcap_ISCX.csv	Thursday	Benign, Infiltration
Friday-WorkingHours-, Morning.pcap_ISCX.csv	Friday	Benign, Bot
Friday-WorkingHours-Afternoon-PortScan.pcap_ISCX.csv	Friday	Benign, PortScan
Friday-WorkingHours-Afternoon-DoS.pcap_ISCX.csv	Friday	Benign, DDoS

Figure 2: Description of CICIDS2017

5.2 Workflow

For the first step, as mentioned before, the workflow of the original article was copied for comparing the results. Afterwards, we have advanced to the second step: In the second step, we have experimented on the same portion of data. Regarding the feature selection - At the first try, we have tried to train our model with all 78 features considered. Our initial model was Random Forest with a hyperparameter of 100 n-estimators. To improve on this, we have tried to change the hyperparameter of n-estimators to 150, this change has not yielded any better results. Another improvement was to the model - we have changed from Random Forest to XGBoost. Later, we have tried to improve the result by experimenting with the same model, with different features - we picked 10 features that were recommended from the article named "Evaluating the CIC IDS-2017 Dataset Using Machine Learning Methods and Creating Multiple Predictive Models in the Statistical Computing Language R" [4] and tested our models with them, this did not improve on the results. Finally, we have tried switching to 15 selected features, recommended from a second article named "Selection and Performance Analysis of CICIDS2017 Features Importance"[5], this has indeed improved our results. Our other approach to improve the current results was to add a portion of the data from CIC-IDS2017 which wasn't used by the original article, as explained in section 4 - Achieved Contributions. In this approach, we first decided to transform all labels of various DoS/DDoS attacks into "DDoS", so we will only have two categories - "DDoS" and Benign". Then, we wanted to select and decide on the best features for the model, so first we used the original article's selected features, which were selected using RFE algorithm in the article. Later, we wanted to further improve on the results, so we investigated further into finding better feature selections, by exploring different relevant articles, which were revolving around the CIC-IDS2017 dataset. Lastly, we have found and used two different feature selections from the articles "Evaluating the CIC IDS-2017 Dataset Using Machine Learning Methods and Creating Multiple Predictive Models in the Statistical Computing Language R" [4] and "Selection and Performance Analysis of CICIDS2017 Features Importance" [5], both had used feature selection algorithms for their selections. We found that the first article's choice of feature selection has improved our results of the current approach, but the latter has improved it even further.

5.3 AI metrics

Evaluation metrics: The following metrics are used for evaluating the results:

- Accuracy:

  Accuracy refers to the percentage of records which are correctly classified. It is used to evaluate the classification models. The formula to calculate accuracy is: Accuracy = TP + TN / Total samples.

- Precision and recall:

  Precision refers to the number of records that are correctly identified as positive in the total number of records that are identified as positive. Recall refers to the number of records that are correctly identified as positive in the total records that are positive. Precision and recall are effective evaluation metrics in case of classification problems where the density of class labels varies largely.

  Precision = TP / (TP + FP)

  Recall = TP / (TP + FN)

- F1 score:
  F1 score is derived from precision and accuracy. F1 score is high only when both are high. So, it is not biased. It takes equal weightage of recall and precision.
  F1 score = (2 × Precision × Recall) / (Precision + Recall)

6. Dataset Exploration

In our dataset, CIC-IDS2017, there are 78 different features. There are totally 691,406 records in the original dataset, out of which, 439,683 Benign records and 251,723 Dos records. After data expanding, there are totally 917,117 records in the dataset, out of which, 537,369 Benign records and 379,748 DDOS/DOS records.
We have explored the correlation of two groups of features - the first is the group of features that were chosen by the original article for the training of the models. The second was the group of features that we have chosen and that were proved to be the most effective.

First group features:
[' Destination Port', ' Bwd Packet Length Mean', ' Bwd Packet Length Std',' Bwd Packets/s', ' Packet Length Mean', ' Max Packet Length', ' Avg Bwd Segment Size', 'Init_Win_bytes_forward']

Second group features:

[' Fwd IAT Min', 'Init_Win_bytes_forward',' Destination Port', ' Bwd Packet Length Min', ' Init_Win_bytes_backward', ' Subflow Fwd Bytes', ' Total Fwd Packets', ' Total Length of Bwd Packets', ' Bwd Packet Length Mean', ' Fwd Packet Length Min']

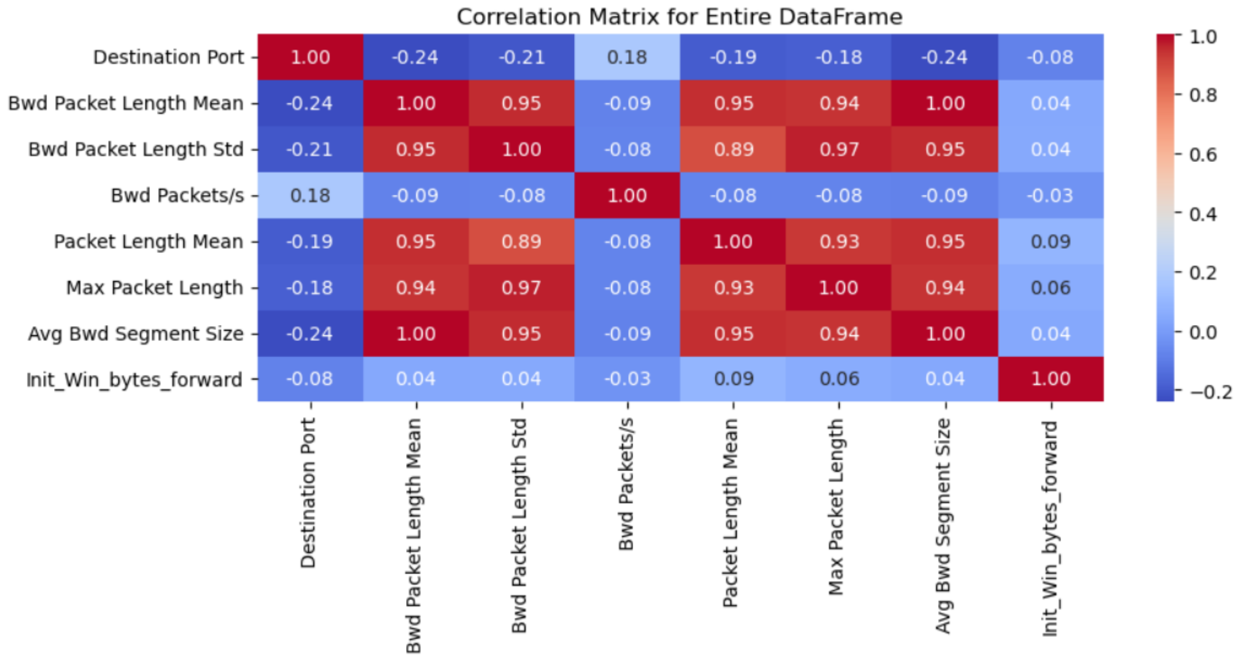


Figure 3: Correlation matrix original features

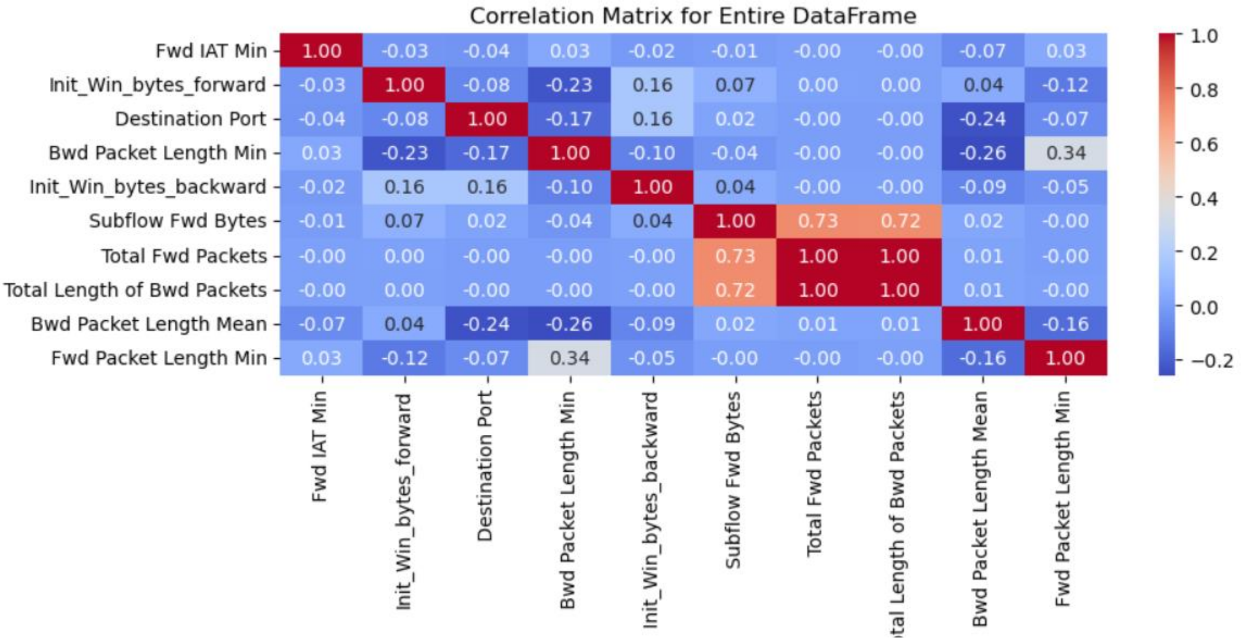
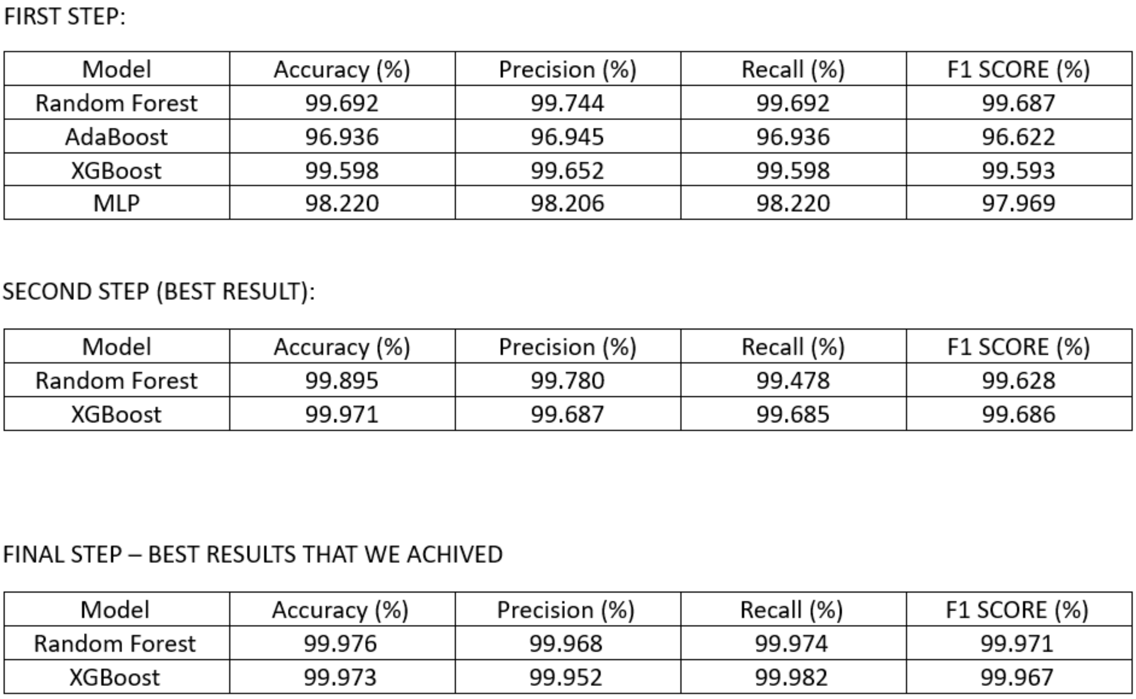


Figure 4: Correlation matrix our features

As we can see, our group of features has less correlation than the first.

7. Algorithm and results

In the results below, we can see all the relevant results:



FIRST STEP:

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 SCORE (%)
Random Forest	99.692	99.744	99.692	99.687
AdaBoost	96.936	96.945	96.936	96.622
XGBoost	99.598	99.652	99.598	99.593
MLP	98.220	98.206	98.220	97.969

SECOND STEP (BEST RESULT):

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 SCORE (%)
Random Forest	99.895	99.780	99.478	99.628
XGBoost	99.971	99.687	99.685	99.686

FINAL STEP – BEST RESULTS THAT WE ACHIVED

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 SCORE (%)
Random Forest	99.976	99.968	99.974	99.971
XGBoost	99.973	99.952	99.982	99.967

Figure 5: Results

As we can observe, our approach has improved upon the results of the original paper's. We can see that our Random Forest model in step 3 has gotten better results in all parameters than the original's Random Forest (in step 1): accuracy: +≈ 0.3%, precision: +≈ 0.22%, recall: +≈ 0.3%, F1-Score: +≈ 0.3%. This is due to our novel and innovative approach: Addition of data, aggregation of the DoS/DDoS types to a singular "DDoS" label and finding better features for training from other papers that has explored the CIC-IDS2017 dataset.

8. Summary

The goal of this research was to try to improve on the article's "DDOS Attack Identification using Machine Learning Techniques" [2] results of identifying DoS/DDoS attacks. This was done by adding data to the dataset, using other feature for the training of the models, and changing the label types in the data to a general label. As we observed, we did manage to improve the results.

9. Bibliography:

[1] D. Anstee, C. F. Chui, P. Bowen, and G. Sockrider, Worldwide Infrastructure Security Report, Arbor Networks Inc., Westford, MA, USA, 2017.

[2] Subhashini Peneti, Hemalatha E ” DDOS Attack Identification using Machine Learning Techniques”

[3]Chandra Sekhar Reddy N, Vemuri P, Govardhan A “An empirical study on support vector machines for intrusion detection”. International Journal of Emerging Trends in Engineering Research (2019)

[4] Zachariah Pelletier, Munther Abualkibash "Evaluating the CIC IDS-2017 Dataset Using Machine Learning Methods and Creating Multiple Predictive Models in the Statistical Computing Language R"

[5] Reis, B., Maia, E., Praça, I. (2020). Selection and Performance Analysis of CICIDS2017 Features Importance. In: Benzekri, A., Barbeau, M., Gong, G., Laborde, R., Garcia-Alfaro, J. (eds) Foundations and Practice of Security. FPS 2019. Lecture Notes in Computer Science(), vol 12056. Springer, Cham. https://doi.org/10.1007/978-3-030-45371-8_4

[6] Canadian Institute for Cybersecurity - Intrusion Detection Evaluation Dataset (CIC-IDS2017) - https://www.unb.ca/cic/datasets/ids-2017.html