Journal of Discrete Mathematical Sciences & Cryptography
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 26 (2023), No. 5, pp. 1317–1325
DOI : 10.47974/JDMSC-1747

A hybrid methodology with learning based approach for protecting systems from DDoS attacks

G. Ramesh [*]
Department of Computer Science and Engineering
Gokaraju Rangaraju Institute of Engineering & Technology
Hyderabad
Telangana
India

Venkata Ashok K Gorantla [†]
Senior Technical Product Manager
Verizon
U.S.A.

Venkataramaiah Gude [§]
GP Technologies LLC
U.S.A.

Abstract

Distributed Denial of Service (DDoS) attacks still prevailing in Internet based and cloud based applications. To detect such attacks and mitigate their effect, many approaches came into existence. There are signature based methods, metrics based methods and machine learning (ML) based methods. With the availability of training data, ML based solutions, of late, became popular. However, there is need for evaluation of different ML models for real time usage in distributed applications. We proposed a ML based framework that has mechanisms, including feature selection, to have supervised learning for threat detection. The framework enables workflow required to pre-process data, select essential features, train ML classifiers and detect the DDoS attack and classify it. We also proposed an algorithm known as DDoS Attack Detection for Critical Services Protection (DAD-CSP) that takes dataset and ML pipeline as input, exploits the ML models and evaluates them. Feature selection has resulted in dimensionality reduction for improving quality in training.

[*] *E-mail:* `ramesh680@gmail.com (Corresponding Author)`
[†] *E-mail:* `ashok.gorantla@gmail.com`
[§] *E-mail:* `gvramaiah.se@gmail.com`

The ML models such as Decision Tree, Naïve Bayes and Random Forest showed different capabilities in attack classification. RF exhibited highest performance with 92% accuracy when compared with other two models.

Subject Classification: *68M25.*

Keywords: *Cyber security, Machine learning, DDoS attack detection, Artificial intelligence.*

1. Introduction

In service oriented distributed applications, it is important that the service being rendered to genuine users is uninterrupted. If there is intentional interruption by adversaries through an attack, it is usually known as DDoS attack. Unless DoS attack which result in less damage to the system, DDoS has high level of victimization and damage to distributed applications. Therefore, it is important to see that DDoS attacks are detected and the attack effects are minimized. Their intention is to deny service of specific application have monetary and other gains from such attacks. Big enterprises in the real world are found victims of such attacks. With the emergence of ML models as part of AI, now it is possible to have continuous training to models and detect DDoS attack in real time.

Many ML techniques based approaches are found in the literature. From the review of literature, valuable insights are ascertained. First, ML algorithms have potential to detect DDoS attacks provided training data. Second, feature selection plays an important role in improving prediction performance. Third, there is need for novel approaches to protect critical services and infrastructure from DDoS attacks. Towards a solution, we proposed a ML framework wit underlying supervised learning mechanisms to detect DDoS attacks accurately. We proposed an algorithm known as DDoS Attack Detection for Critical Services Protection (DAD-CSP) to ensure that crucial services in a given distributed application are not denied due to DDoS attacks. We implemented the system using Python data science platform and evaluated with various performance metrics. Rest of the paper has Section 2 with literature review, Section 3 with proposed work, Section 4 with results and Section 5 conclusions.

2. Related Work

Filho *et. al*. [1] proposed a smart detection method for detecting DDoS attacks. They exploited different ML approaches with feature engineering. It is an online approach towards automatic detection of attacks. Alrehan *et. al*. [2] explored the usage ML methods to detect DDoS attacks in Vehicular Ad Hoc Network (VANET). They found different solutions in the literature with simulation and empirical studies. In the process, they used both unsupervised and supervised learning methods. Zekri *et. al*. [3] proposed ML based approaches to detect DDoS attacks in cloud computing environment. Robinson *et. al*. [4] investigated different ML models and ranked them based on their performance in detecting DDoS attacks. Bindra and Soon [5] applied ML models to detect DDoS attacks in supervised learning approach.

Rios *et. al*. [6] combined ML models with fuzzy logic to reduce the effectiveness of DDoS attacks. Saini *et. al*. [7] proposed an IDS that detect DDoS attacks using ML models. In [8], [9], [10], [11] authors focused on cloud environment and proposed a DDoS detection methodology from the source side of distributed environment. Elsayed *et. al*. [12] proposed a system to protect SDN infrastructure by detecting attacks using ML techniques. Since SDN plays crucial role in controlling a network, attacks on SDN have potential to collapse the entire network. Wani *et. al*. [13] explored the usage of ML models in cloud computing for secure environment. Sofi *et. al*. [14] investigated on the modern kinds of DDoS attacks. Accordingly, they proposed a ML based framework to detect such attacks. Wehbi *et. al*. [15] explored on different approaches used to detect DDoS attacks in IoT environment.

3. Proposed Method

DDoS attacks still prevailing in Internet based and cloud based applications. To detect such attacks and mitigate their effect, many approaches came into existence. There are signature based methods, metrics based methods and machine learning (ML) based methods. With the availability of training data, ML based solutions, of late, became popular. However, there is need for evaluation of different ML models for real time usage in distributed applications. Towards this end, we proposed a ML framework to deal with the detection of DDoS attacks. Since ML models have provision for supervised learning and there is availability of training data for efficient detection of DDoS attacks, this kind of ML is preferred in our methodology. We have chosen three ML models that are widely used and high performing models. They are known as RF, DT and Naïve Bayes. The proposed framework is show in Figure 1. It has mechanisms to have pre-processing, exploratory data analysis and visualization besides prediction of DDoS attacks and threats using aforementioned ML techniques.

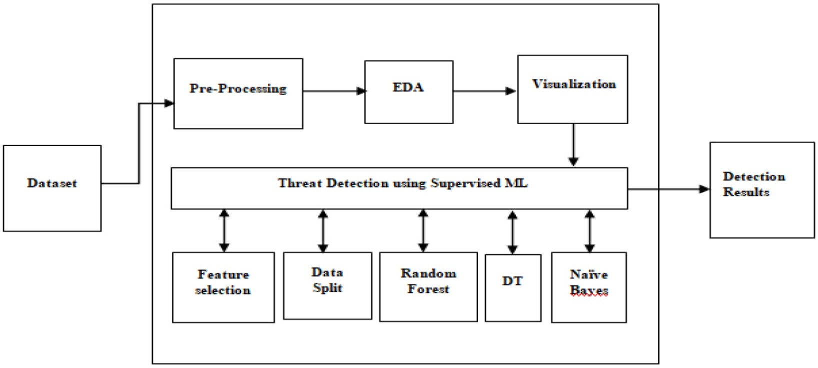

Figure 1

Overview of the proposed ML based framework for DDoS attack detection

As presented in Figure 1, the given dataset is pre-processed before moving further. The pre-processing approach takes care of missing values. The data after then is subjected to exploration for understanding data dynamics. The exploratory data analysis has visualization of different facts including the attack distribution in the data. Feature selection has resulted in identification of contributing features.

Different ML models showed their capability in prediction process. It shows different attack models in data distribution and the data is subjected to agglomerative clustering, t-SNE technique and PCA for visualizing the attack distribution in the form clusters, visual map and principal components respectively. The given dataset is used by ML models in training and testing phases respectively.

Algorithm 1: DDoS Attack Detection for Critical Services Protection (DAD-CSP)

```
Algorithm: DDoS Attack Detection for Critical Services Protection (DAD-CSP)
Inputs: ML models pipeline M, dataset D
Output: DDoS detection results
1. Start
2. D'←Pre-Process(D)
3. F←Feature Selection(D')
4. (T1, T2)←SplitData(D')
5. For each m in pipeline M
6. Train m using F and T1
```

```
7.  End For
8.  For each m in pipeline M
9.  Fit the model m on T2 with model knowledge
10. Evaluation
11. Results
12. End For
13. End
```

Algorithm 1 takes dataset and different ML models' pipeline as input. The dataset is subjected to pre-processing. It improves the dataset by filling missing values and treating null values properly. Then the dataset is subjected to feature selection. The feature selection process results in identification of contributing features instead of using all features. The training set is used in the training phase of each prediction model. There is an iterative process to train each model in M. Each model gains intelligence and that is used in the testing phase. Once all the models are trained, there is another iterative process to deal with the prediction of DDoS attacks. Evaluation of the prediction process is made using standard metrics.

4. Experimental Results

Experimental results are observed in terms of exploratory data analysis, feature selection and prediction of DDoS attacks using ML models like DT, RF and Naïve Bayes.

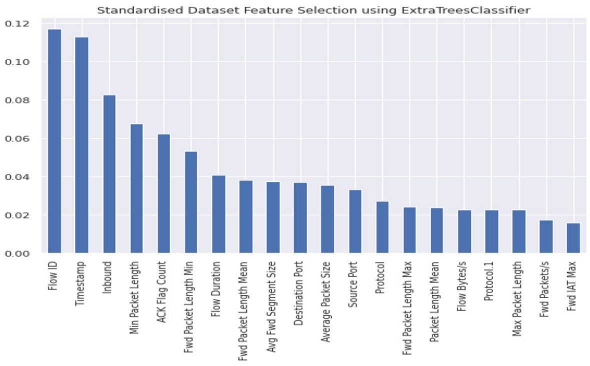

Figure 2

Shows the selected features and their importance

4.1 *Feature Selection Results*

Feature engineering is made on the given dataset with the help of ExtraTree classifier. It is used in order to identify features that are important for detection and classification of DDoS attacks. Since feature engineering is crucial for improving quality in training, feature selection is implemented to identify features. Figure 2 shows the selected features.

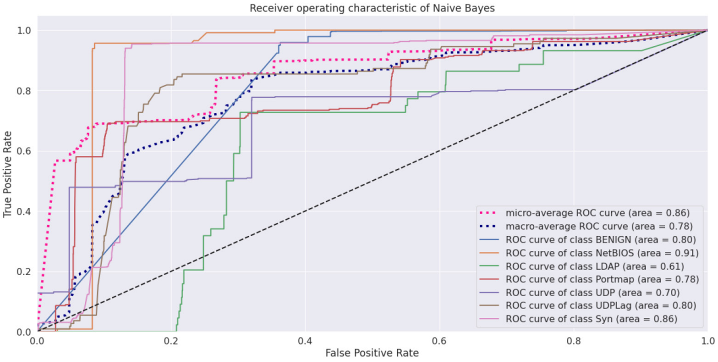

Figure 3

Shows ROC curve for different attack classes with Naïve Bayes model

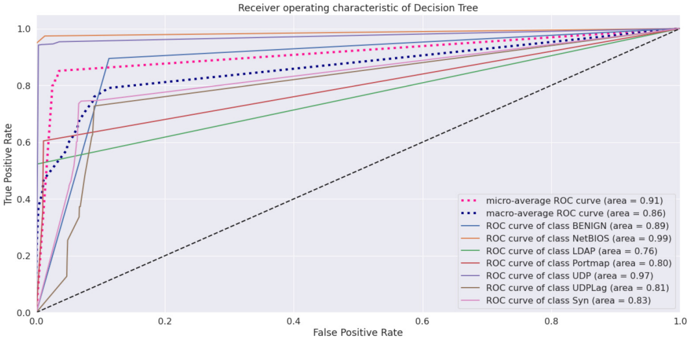

Figure 4

Shows ROC curve for different attack classes with DT model

4.2 *Classification Results*

DDoS attack detection and classification are made using different ML approaches that are pipelined in the proposed algorithm. Receiver Operative Curve (ROC) of the three ML models is computed and presented in this section.

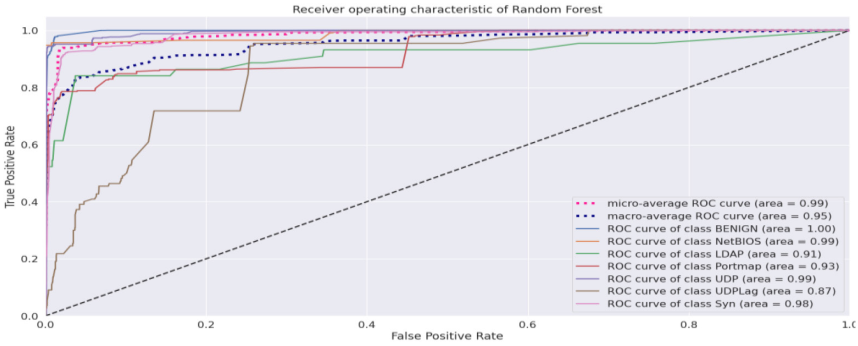

Figure 5

Shows ROC curve for different attack classes with RF model

As presented in Figure 3, Figure 4 and Figure 5, there is visualization of different ML models in terms of various attack classes identified. The ROC curve with area is presented. The results showed that there is acceptable performance by the three models.

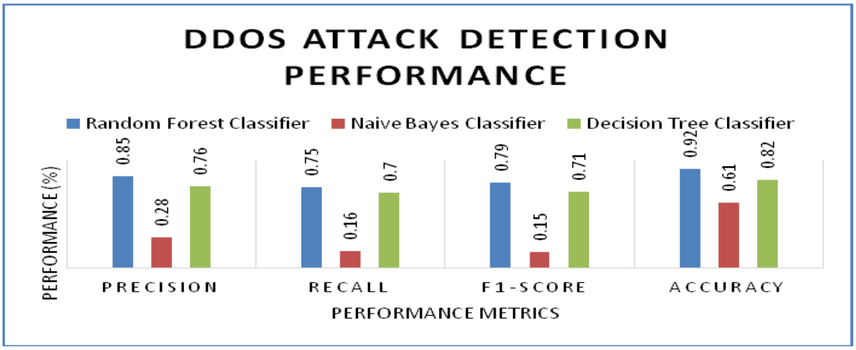

Figure 6

DDoS attack detection performance comparison

As presented in Figure 6, the detection performance of DT, RF and Naïve Bayes is visualized. The performance is measures using different metrics including accuracy. The prediction performance of each model is different due to their underlying mechanisms. However, the highest performance is achieved by RF classifier with 85% precision, 75% recall, 79% F1-score and 92% accuracy

5. Conclusion and Future Work

In this paper, we proposed a ML based framework that has mechanisms, including feature selection, to have supervised learning for threat detection. The framework enables workflow required to pre-process data, select essential features, train ML classifiers and detect the DDoS attack and classify it. We also proposed an algorithm known as DDoS Attack Detection for Critical Services Protection (DAD-CSP) that takes dataset and ML pipeline as input, exploits the ML models and evaluates them. Feature selection has resulted in dimensionality reduction for improving quality in training. The ML models such as DT, NB and RF showed different capabilities in attack classification. RF exhibited highest performance with 92% accuracy when compared with other two models. This research has resulted in many significant results. First, it has proved that ML approaches are suitable for DDoS attack detection. Second, feature selection approach has potential to improve classification accuracy of ML models. As future work, we will exploit hybrid models that combine linear approaches (ML models) and deep learning (non-linear) models for improving prediction performance.

References

[1] Lima Filho, F.S. de, Silveira, F.A.F., de Medeiros Brito Junior, A., Vargas-Solar, G. and Silveira, L.F. Smart Detection: An Online Approach for DoS/DDoS Attack Detection Using Machine Learning, [online] Security and Communication Networks, P1-16 (2019).

[2] Alrehan, A.M. and Alhaidari, F.A. Machine Learning Techniques to Detect DDoS Attacks on VANET System: A Survey, *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, P1-6 (2019).

[3] Marwane Zekri, Said El Kafhali, Noureddine Aboutabit and Youssef Saadi. DDoS Attack Detection using Machine Learning Techniques in Cloud Computing Environments, *2017 2nd IEEE*, 1-7 (2017).

[4] Rejimol Robinson, R.R. and Thomas, C. Ranking of machine learning algorithms based on the performance in classifying DDoS attacks. [online] IEEE Xplore, p185-190 (2015).

[5] Naveen Bindra and Manu Sood. Detecting DDoS Attacks Using Machine Learning Techniques and Contemporary Intrusion Detection Dataset., Automatic Control and Computer Sciences, 53(5), pp. 419–428 (2019).

[6] Rios, V. de M., Inácio, P.R.M., Magoni, D. and Freire, M.M. Detection of reduction-of-quality DDoS attacks using Fuzzy Logic and machine learning algorithms, *Computer Networks*, 186, p.107792 (2021).

[7] Saini, P.S., Behal, S. and Bhatia, S. Detection of DDoS Attacks using Machine Learning Algorithms. *2020 7th International Conference on Computing for Sustainable Global Development (INDIACom)*. P 16-21 (2020).

[8] He, Z., Zhang, T. and Lee, R.B. Machine Learning Based DDoS Attack Detection from Source Side in Cloud. [online] IEEE Xplore.p 114-120 (2017).

[9] A.DivyaRani, G. Ramesh, K. Madhavi, An Efficient and Effective Framework to Track, Monitor, and Orchestrate Resource Usage in an Infrastructure as a Service, *International Journal of Recent Technology and Engineering* (IJRTE), Volume-8 Issue-3 (September 2019).

[10] Thirupathi, N., Madhavi, K., Ramesh, G., Sowmya Priya, K. Data Storage in Cloud Using Key-Policy Attribute-Based Temporary Keyword Search Scheme (KP-ABTKS). In: Smys, S., Bestak, R., Rocha, Á. (eds) Inventive Computation Technologies. ICICIT 2019. Lecture Notes in Networks and Systems, Vol 98 (2020). Springer, Cham.

[11] Reddy, N.M., Ramesh, G., Kasturi, S.B. et. al. Secure data storage and retrieval system using hybridization of orthogonal knowledge swarm optimization and oblique cryptography algorithm in cloud. *Appl Nanosci* (2022).

[12] Elsayed, M.S., Le-Khac, N.-A., Dev, S. and Jurcut, A.D. Machine-Learning Techniques for Detecting Attacks in SDN. [online] IEEE Xplore. P 277-281 (2019).

[13] Wani, A.R., Rana, Q.P., Saxena, U. and Pandey, N. Analysis and Detection of DDoS Attacks on Cloud Computing Environment using Machine Learning Techniques, *Amity International Conference on Artificial Intelligence (AICAI)*. P 870-876 (2019).

[14] Irfan Sofi, Amit Mahajan and Vibhakar Mansotra. Machine Learning Techniques used for the Detection and Analysis of Modern Types of DDoS Attacks, *International Research Journal of Engineering and Technology (IRJET)*. 4, p 1086-1092 (2017).

[15] Wehbi, K., Hong, L., Al-salah, T. and Bhutta, A.A. A Survey on Machine Learning Based Detection on DDoS Attacks for IoT Systems. [online] IEEE Xplore. P 1-6 (2019).