NATURALISTA CAMPANO
ISSN: 1827-7160
Volume 28 Issue 1, 2024

Analyzing the Impact of DDoS Attack over the Server using ANN Boosting Classifier

Dr.Shaweta Sachdeva,

Assistant Professor Guru Nanak Khalsa College, Yamuna Nagar

Abstract- Throughout the past twenty years, there has been a substantial expansion in the utilization of services and applications reliant on the Internet. Currently, approximately 57% of the world's population engages with the Internet. As a result, there has been a notable increase in the attention directed towards Internet security. Throughout its existence, the Internet has remained susceptible to various security challenges, including threats like Worms, port scans, denial of service attacks, distributed denial of service attacks, and Trojans. Service environment is against high-rate DDoS flooding attacks effectively through the use of a two-level security mechanism. Level-I and Level-II were implemented using an entropy-based mechanism and the ANN based boosting classifier (ANN-BC), respectively. First, traffic instances pass through the Level-I module. Next, the traffic instances pass through the Level-II module. Essential features extracted from the flow information to improve the accuracy rate included the precision, recall, F1-score and loss.

Keywords- *DDoS Attack, Accuracy, Artificial Neural Network (ANN), Boosting Classifier*

I. Introduction

Traditional DDoS defense mechanisms often rely on rule-based techniques and network infrastructure improvements, but these methods often struggle to keep pace with the evolving complexity of attacks. In recent years, the application of machine learning algorithms to DDoS detection and mitigation has collected increasing attention due to their potential to adapt to new attack patterns and enhance the accuracy of detection. Machine learning approaches leverage the power of data analysis and pattern recognition to distinguish between normal network traffic and attack traffic, enabling more proactive and precise responses to DDoS threats [6]. This thesis aims to contribute to the field of cyber-security by investigating and proposing a machine learning-based approach for DDoS attack detection and mitigation. By leveraging the massive amounts of data generated during network operations, machine learning models can learn to identify abnormal traffic patterns associated with DDoS attacks, thereby enhancing the ability to detect and respond to these threats in real-time. Additionally, the utilization of machine learning techniques can facilitate the development of adaptive defense strategies that evolve alongside emerging attack vectors [6]. Machine learning leverages the power of data analysis and pattern recognition to distinguish between malicious and legitimate network traffic. [6].This ability to adapt and learn from evolving attack patterns presents a transformative opportunity to enhance the resilience of digital infrastructures against DDoS attacks. By intelligently analyzing vast datasets generated during network operations, machine learning algorithms can identify anomalies and detect subtle indicators of DDoS activity that might otherwise go unnoticed [7]

Software-defined network (SDN) controllers play an important role in managing denial-of-service (DDoS). SDN provides an innovative and distinct technique of network administration. The separation provides a high level of flexibility and scalability in addition to permitting network administrators to make dynamic network adjustments. Because it is the central component of the SDN, the controller must have the highest level of security. The data plane consists of network switching components without control capabilities. Figure 1 depicts the SDN architecture.

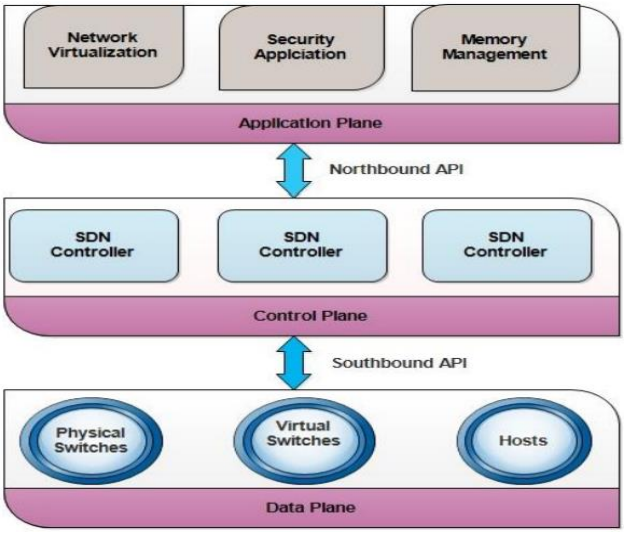

Fig. 1: Architecture of SDN

A clear advantage the SDN offers is the ability to quickly process entire requests from different devices using a programming interface. Centralized intelligence in the SDN transforms the networking function, making it dynamic and powerful. The SDN offers integration with the public cloud. The SDN helps to reduce the overall operating costs by automating and centralizing the administrative process [7, 8]. An added advantage of the SDN is its efficient control of data traffic. Technologies like cloud computing, big data and virtualization increasingly demand dynamic and flexible networks. Such demands have resulted in IT enterprises and the corporate sector switching to SDN services for superior performance, innovations, reduced costs and complexity [10].

II. DDoS IMPACT ON NETWORK

Distributed Denial of Service (DDoS) attacks have the potential to wreak havoc on computer networks, causing a cascade of detrimental effects that disrupt services, compromise user experience, and even result in financial losses [11]. DDoS security threats remain a persistent and evolving challenge for computer networks. Understanding the nuances between classic and modern attacks, amplification, and reflection strategies, layered versus multi-vector approaches, and the various motivations behind these attacks is crucial for organizations to build robust defense mechanisms. Implementing a combination of effective mitigation strategies tailored to the specific threat landscape is imperative to maintaining the availability, integrity, and functionality of critical online services [12].

Disruption of Service Availability: DDoS attacks are primarily aimed at overwhelming a target's resources, rendering them inaccessible to legitimate users. This leads to service downtime and frustration among users who rely on those services. For instance, consider a financial institution that falls victim to a DDoS attack during peak trading hours. The attack saturates the network, making online banking and transactions impossible for users, thus eroding trust and potentially leading customers to switch to more reliable competitors.

Degraded Network Performance: Even if a DDoS attack doesn't completely shut down a service, it can significantly degrade network performance. This manifests as slow loading times, delayed responses, and overall sluggishness. Imagine an e-commerce website facing a DDoS attack on Black Friday, the busiest online shopping day. The website's response time becomes unbearably slow due to the influx of malicious traffic, frustrating shoppers and resulting in abandoned carts.

Collateral Damage: DDoS attacks can have ripple effects that extend beyond the primary target. Shared hosting environments or cloud service providers hosting multiple websites can experience collateral damage when one of their clients is targeted. If a malicious hacker attacks a specific website within such an environment, other websites sharing the same resources might also face performance issues or downtime. For instance, a shared hosting server hosting numerous small business websites might witness all its clients being affected due to a single targeted attack [13].

Resource Depletion: DDoS attacks consume valuable resources such as bandwidth, CPU power, and memory, leaving little room for legitimate traffic. This can lead to network congestion and exhaustion of resources. Consider an online gaming platform hit by a DDoS attack during a highly anticipated game launch. The attack diverts network resources away from the game servers, causing severe lag and rendering the gaming experience unplayable for users.

Reputation Damage: A successful DDoS attack can tarnish an organization's reputation. When users experience service disruptions, they often associate the negative experience with the brand, regardless of whether the outage was a result of an attack. A notable example is the 2016 Dyn attack, which targeted a major DNS provider, causing widespread outages across websites and online services. While Dyn wasn't the primary target, its reputation took a hit due to the collateral damage inflicted on its clients[17]-[20].

Financial Losses: The financial consequences of DDoS attacks can be substantial. Loss of revenue due to disrupted services, the cost of implementing mitigation measures, and potential legal liabilities can add up. For instance, a popular online retailer facing a DDoS attack on a major shopping event might lose millions in sales due to the unavailability of its platform during the crucial sales window.

III. Proposed Methodology

In ML, the ANN based boosting classifier (ANN-BC) is the assortment of Supervised ML calculations in light of the renowned Bayes hypothesis of likelihood. It is the group of calculations that shares a typical strategy to appraise the circulation of the information, for example, each pair of traits being arranged is free of one another. Bayes' hypothesis is a methodology used to gauge the realness of convictions (hypothesis, statements, propositions) in light of pertinent realities accessible (presumptions, records, data). It helps in the change of the anticipated probabilities of an event by applying new data. The ANN-BC is used to solve problems involving binary and multi-characterization, such as text categorization, disease analysis, and forestation. A different methodology is necessary depending on the type of data for each trait. The information is mostly used to determine the characteristics of one of three types of probability distributions. The NB classifier primarily incorporates three types of information dissemination based on classification. Multinomial dispersion is used for names and counts, for example. Binomial ANN-BC dissemination is used for binomial. The Gaussian ANN-BC is used for numeric values such as 1/0, yes/no, or true/false.

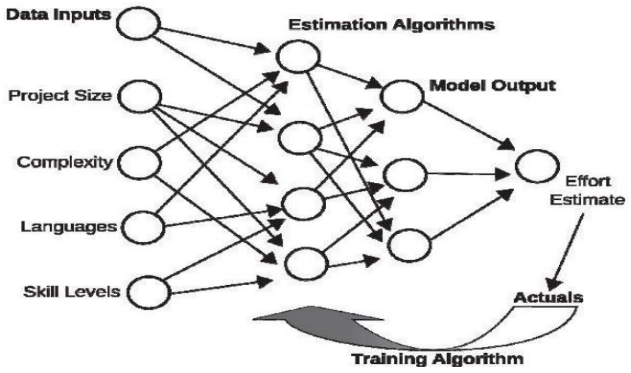

Fig. 2. Architecture of ANN

Step 1: Data set collection

According to dataset attribute information

- target column 'Normal' represents Good Connection
- Bad connection attack types are
  - DoS(Denial of Service)
  - User to Root(U2R)
  - Remote to Local(R2L)
  - Probe

Files used kddcup.data_10_percent.gz, kddcup.names, training_attack_types

```
# map actual type to another column called 'target_type'
df['target_type'] = df.target.apply(lambda x : attack_dict[x[0:-1]] )
df.target_type.value_counts()
```

```
dos       391458
normal     97277
probe       4107
r2l         1126
u2r           52
Name: target_type, dtype: int64
```

Step 2: Categorical Features Exploration and Analysis

```
[ ] # Identifying categorical features
    numeric_cols = df._get_numeric_data().columns # gets all the numeric column names

    categorical_cols = list(set(df.columns)-set(numeric_cols))
    categorical_cols
```

```
['service', 'target', 'flag', 'target_type', 'protocol_type']
```

Step 3: Split Data into training and testing purpose into 80: 20 ratio

```
[ ] X_train, X_test, y_train, y_test = train_test_split(x, y, test_size=0.20, random_state=101)
    print('Shape of Independent features Train data : ' + str(X_train.shape))
    print('Shape of Dependent features Train data : ' + str(y_train.shape))
    print('Shape of Independent features Test data: ' + str(X_test.shape))
    print('Shape of Dependent features Test data: ' + str(y_test.shape))
```

```
Shape of Independent features Train data : (395216, 122)
Shape of Dependent features Train data : (395216, 5)
Shape of Independent features Test data: (98804, 122)
Shape of Dependent features Test data: (98804, 5)
```

Step 4: Defining sequential model

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense (Dense)               (None, 122)               15006

 dense_1 (Dense)             (None, 1)                 123

 dense_2 (Dense)             (None, 5)                 10

=================================================================
Total params: 15,139
Trainable params: 15,139
Non-trainable params: 0
_________________________________________________________________
```

IV. Result Analysis

So, the accuracy can be measured according to Eq. 1

$$Accurancy = \frac{TN + TP}{TN + TP + FN + FP} \tag{1}$$

The additional metrics encompass Precision, Sensitivity (Recall), and Specificity, each computed using formulas denoted as Eq. 2 and Eq. 3.

$$\Pr ecision = \frac{TP}{TP + FP} \tag{2}$$

$$\mathrm{Re}\,call = \frac{TP}{TP + FN} \tag{3}$$

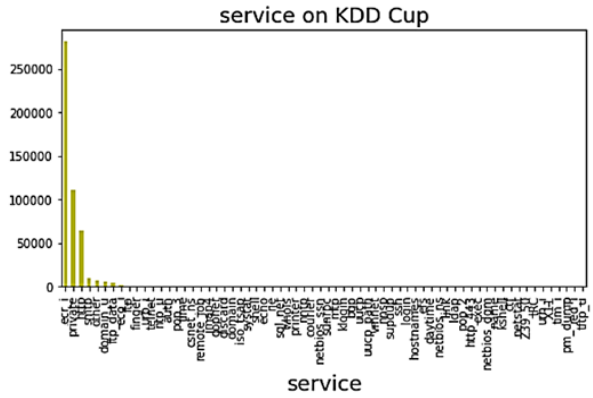

Fig. 3: Different types of Service on KDD Cup

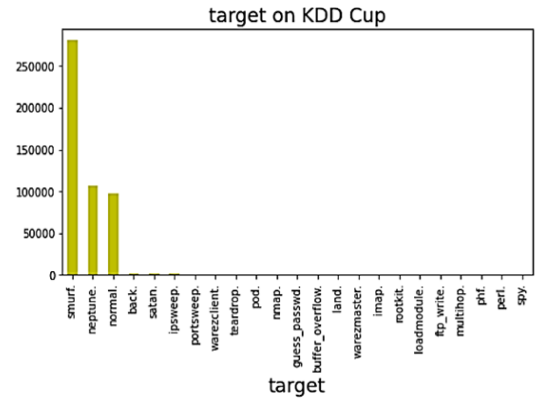

Fig. 4: Different types of target on KDD Cup

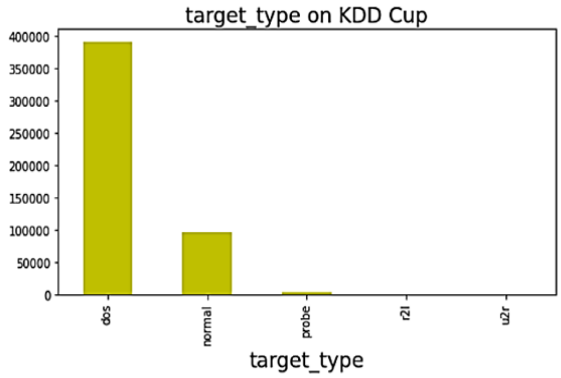

Fig. 5: Target_type on KDD Cup

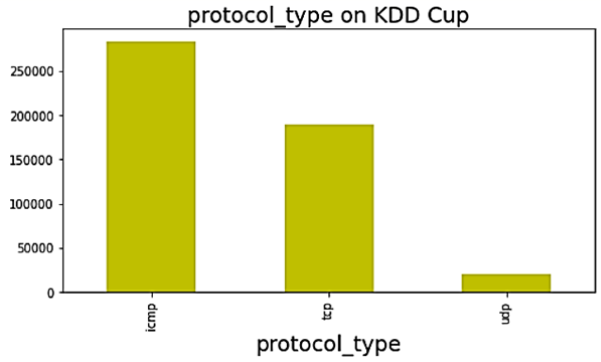

Fig. 6: Protocol_type on KDD Cup

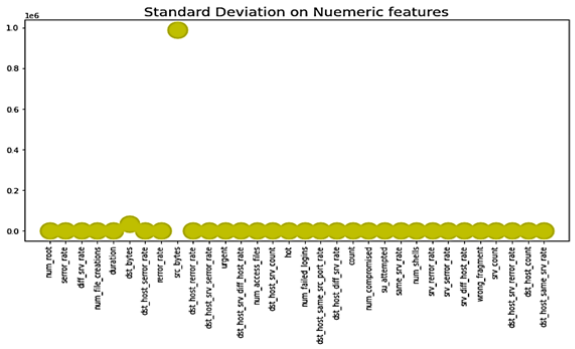

Fig. 7: Standard Deviation on Numerical Features

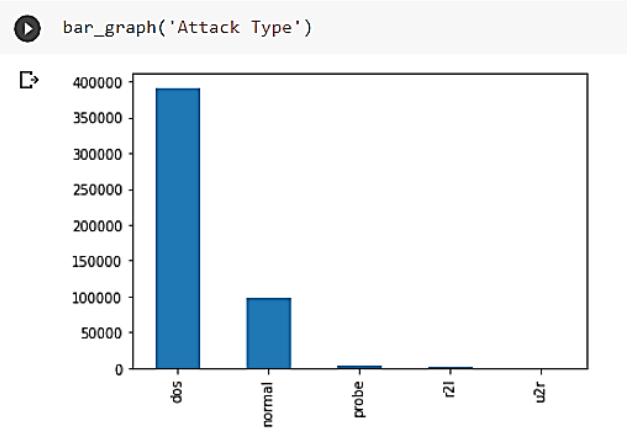

Fig. 8: Different Attack Types

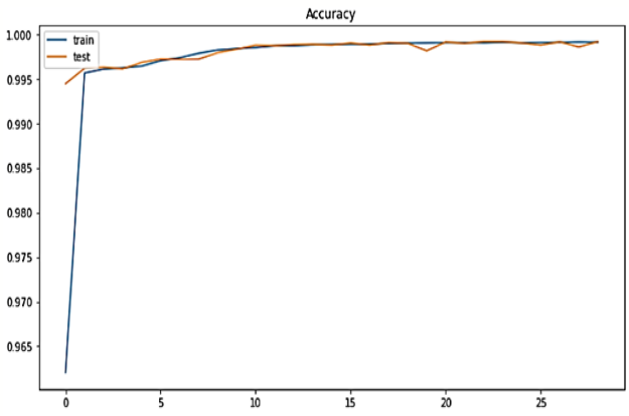

Fig. 9: Accuracy for Test and Training

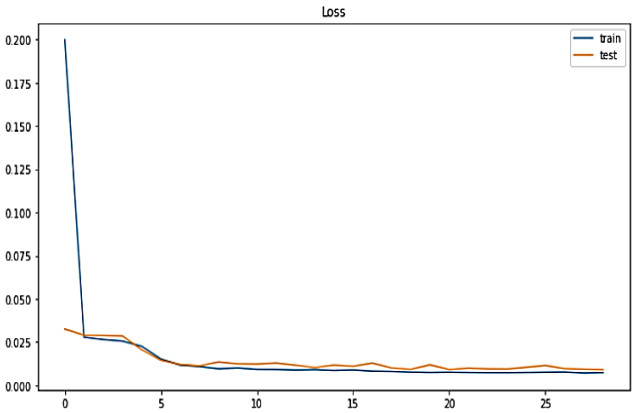

Fig. 10: Loss for Test and Training

Table I: comparison Result

Algorithms	Precession	Recall	F1_Score	Accuracy	Loss
Decision Tree	92.76%	82.34%	86.46%	84.14%	0.045
SVM	93.56%	88.46%	88.80%	88.46%	0.024
Boosting Classifier	98.78%	93.45%	96.78%	97.78%	0.001

Table 1 displays the results of implemented method in terms of precision, recall, accuracy, loss and F1-score. Decision tree (DT) gives a precision of 92.76%, a recall of 82.34%, an accuracy of 84.14%, a loss 0.045 and a F1-score of 77.52%. SVM gives a precision of 93.56%, a recall of 88.46%, an accuracy of 88.46%, loss 0.024 and a F1-score of 88.80%. ANN-BC gives a precision of 98.78%, a recall of 93.45%, an accuracy of 97.78%, loss 0.001 and a F1-score of 96.78%.

V. Conclusion

DDoS attacks represent a grave and ever-evolving threat to computer networks, targeting their availability, functionality, and overall stability. This comparative and detailed exploration delves into the various facets of DDoS security threats, highlighting their distinct characteristics, potential impacts, and mitigation strategies. Normal detection processes, like IP address monitoring are unhelpful in detecting DDoS attacks. The use of advanced ANN models will enable automatic feature extraction. The proposed ANN-BC system improves precision by about 8%, recall by 10-12%, F1-Score by 12-16%, accuracy by 10-15% and loss by 5-10% in the previous system.

References

[1] K. Muthamil Sudar, M. Beulah and P. Deepalakshmi, “Detection of Distributed Denial of Service Attacks in SDN using Machine learning techniques”, International Conference on Computer Communication and Informatics (ICCCI), Jan. 27 – 29, 2021, Coimbatore, INDIA.

[2] Muthamil Sudar, K., & Deepalakshmi, P. (2020). A two level security mechanism to detect a DDoS flooding attack in software-defined networks using entropy-based and C4. 5 technique. Journal of High Speed Networks, (Preprint), 1- 22.

[3] Dong, S., & Sarem, M. (2019). DDoS Attack Detection Method Based on Improved KNN With the Degree of DDoS Attack in Software-Defined Networks. IEEE Access, 8, 5039-5048.

[4] Dong, S., Abbas, K., & Jain, R. (2019). A survey on distributed denial of service (DDoS) attacks in SDN and cloud computing environments. IEEE Access, 7, 80813- 80828.

[5] Gu, Y., Li, K., Guo, Z., & Wang, Y. (2019). Semisupervised K-means DDoS detection method using hybrid feature selection algorithm. IEEE Access, 7, 64351- 64365.

[6] A. Raghavan, F. D. Troia, and M. Stamp, ``Hidden Markov models with random restarts versus boosting for malware detection," *J. Comput. Virol. Hacking Techn.*, vol. 15, no. 2, pp. 97107, Jun. 2019.

[7] T. Young, D. Hazarika, S. Poria, and E. Cambria, ``Recent trends in deep learning based natural language processing [review article]," *IEEE Comput. Intell. Mag.*, vol. 13, no. 3, pp. 5575, Aug. 2018.

[8] X. Lei and Y. Xie, ``Improved XGBoost model based on genetic algorithm for hypertension recipe recognition," *Comput. Sci*, vol. 45, pp. 476481, 2018.

[9] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, ``Deep learning for visual understanding: A review," Neurocomputing, vol. 187, pp. 2748, Apr. 2016.

[10] Abduvaliyev, A., Pathan, A.-S. K., Zhou, J., Roman, R., and Wong, W.-C. "On the Vital areas of Intrusion Detection Systems in Wireless Sensor Networks", IEEE Communications Surveys & Tutorials, Vol. 15, Issue 3, pp. no. 1223–1237, 2015.

[11] Abubakar, A. I., Chiroma, H., Muaz, S. A., and Ila, L. B. "A Review of the Advances in Cyber Security Benchmark Datasets for Evaluating Data-driven based Intrusion Detection Systems", Procedia Computer Science, Vol. 62, pp. no. 221–227, 2015.

[12] Bay, S. D., Kibler, D., Pazzani, M. J., and Smyth, P. (2015), "The UCI KDD archive of Large Data Sets for Data Mining Research and Experimentation", ACM SIGKDD Explorations Newsletter, Vol. 2, Issue 2, pp. no. 81–85, 2015.

[13] Aburomman, A. A. and Reaz, M. B. I. "A novel SVM-kNN-PSO ensemble method for Intrusion Detection System. Applied Soft Computing", Vol. 38, pp. no. 360–372, 2015.

[14] Pedro Casas, JohanMazel and Philippe Owezarski "Unsupervised Network Intrusion Detection Systems: Detecting the Unknown without Knowledge", Elsevier Computer Communications, Vol. 35, Issue 7, pp. no. 772 – 783, 2012.

[15] Carlos A. Catania, Facundo Bromberg and Carlos García Garino "An Autonomous Labeling Approach to Support Vector Machines Algorithms for Network Traffic Anomaly Detection", Elsevier Expert Systems with Applications, Vol. 39, Issue 2, pp. no. 1822–1829, 2012.

[16] Xie, B & Zhang, Q, "Application-layer anomaly detection based on application-layer protocols' keywords", Computer Science and Network Technology (ICCSNT), 2nd International Conference on, pp. 2131-2135, 2012.

[17] Sachdeva S, Ali A (2021) A Hybrid approach using digital Forensics for attack detection in a cloud network environment", International Journal of Future Generation Communication and Networking 14(1):1536–1546.

[18] Sachdeva, S., Ali, A. Machine learning with digital forensics for attack classification in cloud network environment. Int J Syst Assur Eng Manag. Vol.13, 156–165 (2022). https://doi.org/10.1007/s13198-021-01323-4

[19] S.Sachdeva,Aleem Ali,Shahnawaz Khan," Secure and Privacy Issues in Telemedicine: Issues, Solutions, and Standards,pp:321-331, Chapter · January 2022 DOI: 10.1007/978-3-030-99457-0_19.

[20] Sachdeva,Aleem Ali,Salman Khalid," Telemedicine in Healthcare System: A Discussion Regarding Several Practices",pp:295-310, Chapter · January 2022 DOI: 10.1007/978-3-030-99457-0_19.