Prateeksha Sahu, et. al. International Journal of Engineering Research and Applications
www.ijera.com
ISSN: 2248-9622, Vol. 13, Issue 9, September 2023, pp 90-95

RESEARCH ARTICLE OPEN ACCESS

Efficient Accuracy of Software Defined Network with DOS Attacks using Artificial Neural Network

[1]Prateeksha Sahu, [2]Onkar Nath Thakur, [3]Rakesh Kumar Tiwari
Department of Computer Science and Engineering

Abstract- Distributed Denial of Service (DDoS) attacks have been a serious cybercrime attack for decades and are one of the most disturbing areas of cybersecurity due to their disguising nature. The new opportunities for DDoS attacks have been exposed by the expansion of high-speed networks, the availability of DDoS attack tools, and the shifting of operations online due to the pandemic's new normal. Thus, in the year 2020, reported DDoS attacks crossed the 10 million attack threshold and caused more than 800,000 attacks per month with the reporting of the biggest cyberattack targeted at multinational companies like Amazon and Google. SDN technology is a novel architectural method for creating a responsive network strategy via programmability, abstraction, and centralized controller, which solves legacy networking challenges like communication overhead, rigidity, and failure to maintain global intelligence, etc. It can provide secure, controlled, customized, flexible, faster, and programmable networks according to the demands of the modern network scenario. Various researchers focused on the problem of early detection of DDoS attacks in the SDN environment, but the performance of the actuators in the data plane was not discussed. In this paper, the SDN with DDoS attack using ANN is present. ANN model for identifying different attacks through various extracted feature and data generation is fast. The proposed model is simulated python language with COLAB platform and calculated accuracy, precision, recall, F1-score and loss.

Keywords- Software Define Network (*SDN),* Distributed Denial of Service (***DDoS) Attack, Accuracy, Artificial Neural Network (ANN)***

Date of Submission: 02-09-2023 Date of acceptance: 14-09-2023

I. INTRODUCTION

Cybercriminal activity will be humanity's greatest challenge in the coming decades. Cybersecurity ventures have forecasted that the annual cost of cybercrime damage will rise from $3 trillion in 2015 to $6 trillion by 2021 and $10.5 trillion by 2025 [1]. It also predicted a substantial rise in Internet users, from 6 billion in 2022 to 7.5 billion by 2030 [2]. In today's high-tech world, the Internet could be a dangerous place, as businesses become increasingly adept and reliant on their information systems. The threats to information systems create security concerns to the network to which it is connected. So, network security is one of the identified areas of many organizations that need to be protected as part of their system of internal control, as many cybersecurity threats are largely avoidable. Despite existing traditional solutions, the Denial of Service (DoS) attack is one of the prominent cybercrimes [3, 4, 5]. It is also estimated that the average size of Distributed DoS (DDoS) attacks is four times larger than in previous years, which will double to 14.5 million by 2022 [6]. The prominence of DDoS attacks is increasing with the complexity of the network, security issues with the advancement of Information and Communication Technology (ICT) infrastructure, the usage of highspeed networks, and the availability of free DDoS online tools. So, it is very significant to address a proper security measure against rampant DDoS attacks [4]. According to the Netscout threat intelligence report in 2020 [7], demand for DDoS protection mechanisms has increased by 69%, 50%, and 61% in major firms, mid-tier enterprises, and small to midsize businesses, respectively. Net flow analyzers, nextgeneration firewalls with Intrusion Detection System (IDS) or Intrusion Prevention System (IPS), and inline DDoS detection and mitigation are the main threat detection tools employed to find the security holes in the network. According to Worldwide Infrastructure Security Report (WISR), 75% of DDoS attacks target infrastructure, and the firewall, despite being an effective perimeter security tool used by 62% of enterprises for detecting threats to their network, is not ideal solution for DDoS attacks. Thus, firewall failures in DDoS attacks contribute to 62% in 2019, rising to 83% in 2020. So, it is essential to have an efficient framework and

proper detection strategy for the fast detection of the DDoS attacks before it paralyzes the entire network.

II. BENEFITS OF SDN

The introduction of network programmability, the global network intelligence, the decoupling of data plane and control plane, traffic engineering with dynamic forwarding rules of network traffic in SDN paved a secured and adaptable innovation in the network architecture. But the centralized SDN controller causes potential threats due to its single point failure. As a result, the SDN controller is regarded as the most desired target for DDoS attacks since it provides total visibility and information to the network.

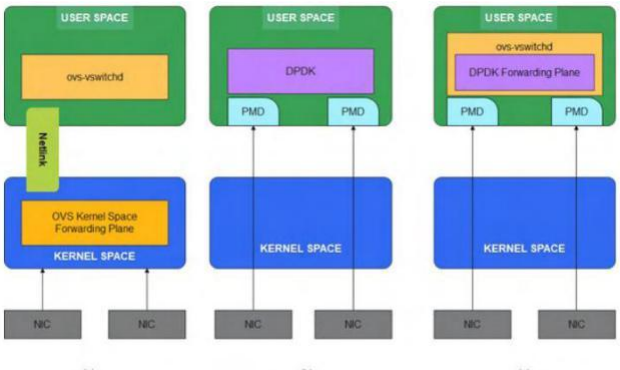


Fig. 1: Packet Processing

The development of technologies like big data, the cloud and virtualization has put pressure on traditional networks typically characterized by a lack of innovation, slow development and production delays.

A clear advantage the SDN offers is the ability to quickly process entire requests from different devices using a programming interface. The software-based controller lets users and administrators manage the traffic flow efficiently with an abstract view of the network. Network administrators increase bandwidth and other resources as needed while simultaneously investing little on additional physical infrastructure.

Centralized intelligence in the SDN transforms the networking function, making it dynamic and powerful. The SDN offers integration with the public cloud, an abstract view of the entire network infrastructure, the management of virtual and physical devices from the centralized controller and low operational costs. 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, which automatically enhances the quality of service (QoS) for multimedia transmissions and Voice over IP (VoIP) [9]. 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].

III. CHALLENGES & ISSUES

3.1 Challenges

The SDN works with strong and dynamic systems administration for IT ventures and correspondence specialist co-ops. However, problems with availability, dependability, scalability, controller placement, security issues like denial-of-service and man-in-the-middle attacks, and vulnerability scans need to be addressed [11]. The SDN regulator is inclined to single-point disappointments and its centralization makes it an obvious objective for assaults. Network manipulation attacks begin when the controller is compromised. Further, compromised information plane gadgets cause a progression of traffic redirections, as well as side-channel and traffic-sniffing assaults. The attacks target disruption because the controller processes each new entry. Regulator disappointment brings about an organization breakdown and administration cuts for a significant length of time, which debases administrations to IT endeavors.

Information plane gadgets and the control plane regulator work freely, imparting data just through the Programming interface. At the point when the quantity of gadgets in the information plane increments, correspondence between the single regulator and various gadgets hits a bottleneck and prompts issues with versatility. Utilizing a legitimate and proficient system to safeguard the SDN regulator guarantees its security, accessibility, dependability and versatility. Our exploration basically centers around security issues in the SDN information plane and control plane [12].

3.2 Issue

The network administrator is responsible for providing security to protect the network from both internal and external intruders. This makes providing and managing security in a computer network difficult. More than 64%, 62%, 59%, and 51% of all businesses were subjected to web-based, phishing, botnet, and DDoS attacks, respectively, according to a survey conducted in 2019
[13]. The purpose of a network intrusion is to disrupt the system resources of the target by stealing confidential information, disabling its functions, and launching various attacks. Due to

the fact that smaller and medium-sized businesses frequently do not have the financial resources to invest in high-level security measures, the threat poses a significant threat to both large and small businesses. Frequently, it brings about assailants zeroing in on weak medium and limited scope associations. Malware, SQL injection, phishing, botnet, cross-site scripting, and DoS and DDoS attacks are the most common threats to networks [14].

Attacks by malware can occur on any device or operating system. Malicious software is created and installed to gain access to a victim's system without their knowledge in these fairly common attacks. The attacks aim to gain access to confidential information and credentials associated with the victim's personal information. harm the framework's assets and gain total access for monetary benefit. Spyware, viruses, and ransomware are the most common types of malware attacks [15, 16].

IV. PROPOSED METHODOLOGY

Techniques or models based on the principles of learning are called as NN. NN are categorized in terms of three basic entities which include the core processing element called as neuron, the interconnection structure and the learning algorithm. Performance of an ANN is defined by its basic architecture which includes various parameters like number of hidden layers in an ANN, the neuron (node) count in each of these layers, transfer function used at each node, weights and parameters of the training algorithm used including their settings too. A general model for software cost estimation based on ANN is shown below:

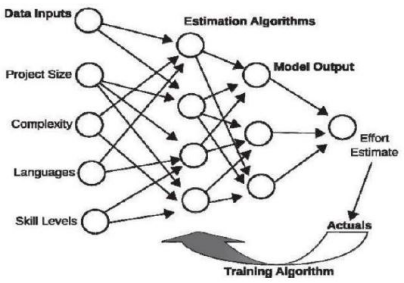


Fig. 2: Neural Network

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 NN Proposed ANN

```
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
_________________________________________________________________
```

Table 1: Training Hyper Parameters

Layers	Dense
Model	Sequential
Neurons	128
Activation function	Relu
Kernal Initializer	Random Uniform
Output Activation Function	Softmax
Training data	80%
Testing data	20%
Loss	Categorical Cross

	Entropy
Optimizer	Adam
Epochs	50

V. RESULT ANALYSIS

Generally, the performance of a classification model is evaluated in terms of accuracy, sensitivity and specificity to calculate which the values of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) need to be considered. A good machine learning model requires high accuracy and low false alarm rates. A confusion matrix is used to determine these parameters. In the confusion matrix, true positive is the number of normal records correctly identified as normal records; false positive is the number of normal records incorrectly identified as attacks; true negative is the number of attack records correctly identified as attacks and false negative is the number of attack records incorrectly identified as normal records.

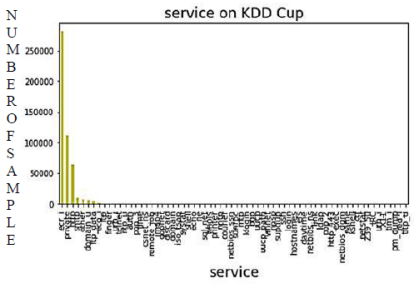


Fig. 3: Different types of Service on KDD Cup

Fig. 3: Different types of Service on KDD Cup

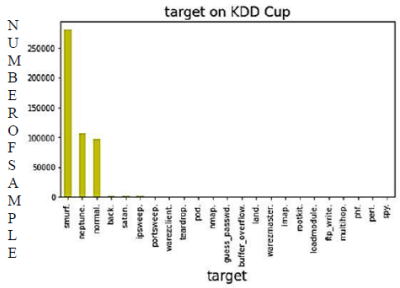


Fig. 4: Different types of target on KDD Cup

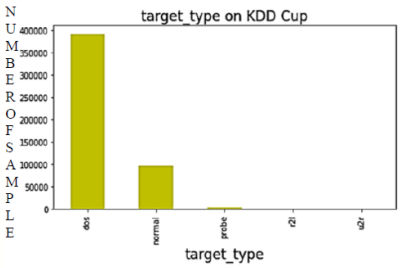


Fig. 5: Target_type on KDD Cup

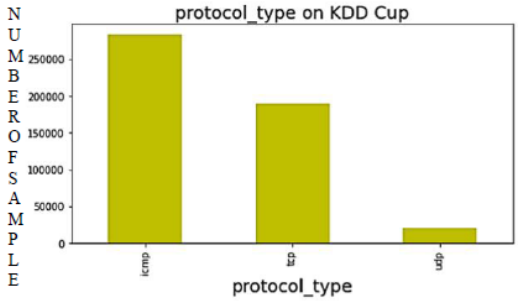
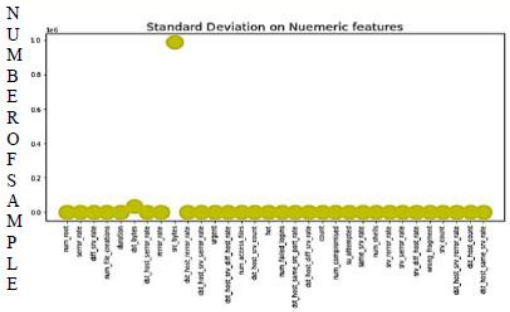
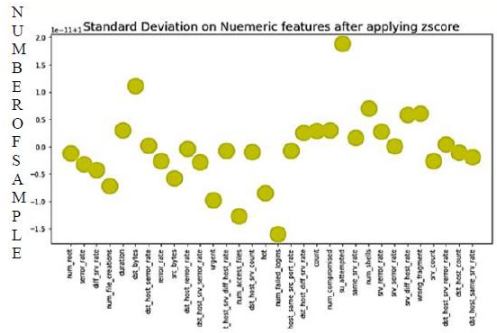


Fig. 6: Protocol_type on KDD Cup



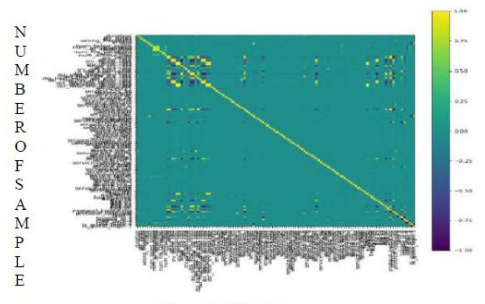
Numerical Feature

Fig. 7: Standard Deviation on Numerical Features



Numerical Feature

Fig. 8: Standard Deviation on Numerical Features after applying z-score



Numerical Feature

Fig. 9: Exploratory Data Analysis

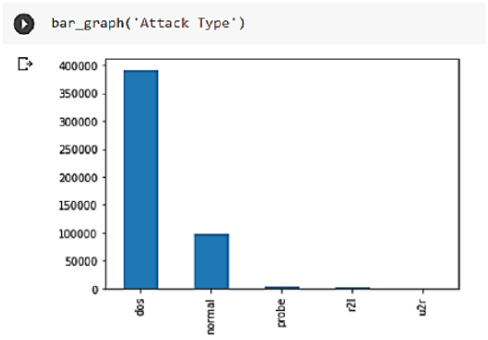


Fig. 10: Different Attack Types

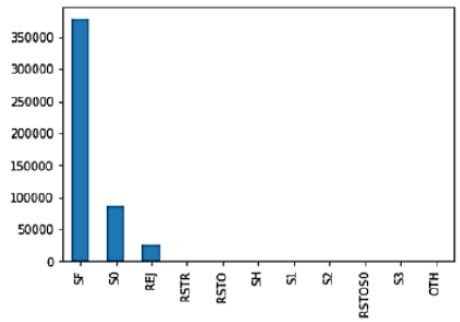


Fig. 11: Different Flag Types

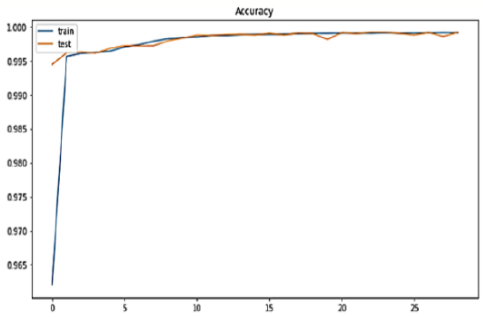


Fig. 12: Accuracy for Test and Training

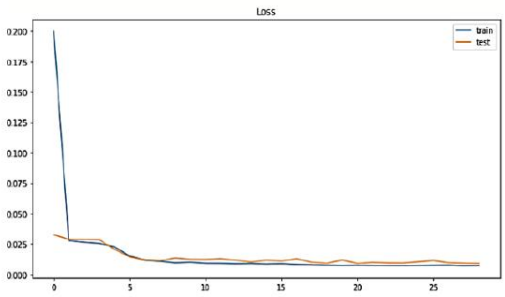


Fig. 13: Loss for Test and Training

Table II: comparison Result

Algorithms	Precession	Recall	F1_Score	Accuracy	Loss
Decision Tree	92	9	95	97	3
SVM	100	72	84	76	33
Proposed ANN	100	100	100	99	0.001

VI. CONCLUSION

The early detection of DDoS attacks is challenging in the cyber world due to their distributed nature, which is difficult to track. The fast detection of DDoS attacks is the need in today's world. The ADE technique is proposed in this study as a novel entropybased lightweight solution for DDoS detection in the SDN environment utilising the D3 framework. Since it is implemented in the D3 framework, it provides an efficient packet capturing mechanism with fast detection in a high-speed network without requiring any hardware resources. Furthermore, by implementing the detection approach in the data plane, the controller overhead is reduced. Although entropy-based detection is the widely accepted lightweight strategy for DDoS detection in the SDN environment, it is less efficient at detecting false alarms.

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
[11]. in Wireless Sensor Networks", IEEE Communications Surveys & Tutorials, Vol. 15, Issue 3, pp. no. 1223–1237, 2015.
[12]. 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.
[13]. 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.
[14]. Aburomman, A. A. and Reaz, M. B. I. "A novel SVM-kNN-PSO ensemble method for Intrusion Detection System. Applied
[15]. Soft Computing", Vol. 38, pp. no. 360–372, 2015.
[16]. PedroCasas,JohanMazelandPhilippeOwezarski "Unsupervised Network Intrusion Detection Systems: Detecting the Unknown without Knowledge", Elsevier Computer Communications, Vol. 35, Issue 7, pp. no. 772 – 783, 2012.
[17]. 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.
[18]. 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.