DDOS Attack Packet Detection and Prevention On a Large-Scale Network Utilising the Bi-Directional Long Short Term Memory Network

[1]Jeevan Pradeep K and [2]Prashanthkumar Shukla
[1,2]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation,Vaddeswaram, India.
[1] jeevanpradeepsappi@gmail.com, [2] prashantshukla2005@kluniversity.in

Correspondence should be addressed to Jeevan Pradeep K : jeevanpradeepsappi@gmail.com.

Article Info
Journal of Machine and Computing (http://anapub.co.ke/journals/jmc/jmc.html)
Doi: https://doi.org/10.53759/7669/jmc202404011
Received 02 June 2023; Revised from 25 August 2023; Accepted 26 October 2023.
Available online 05 January 2024.

Abstract – Security is one of the most challenging conditions for dispersed networks because exclusive threats can damage output overall and can be classified in several ways. At this time, distributed denial-of-service (DDoS) assaults pose the greatest threat to internet security. Rapid identification of communication records for messages referencing DDoS occurrences enables organizations to take preventative action by instantly identifying both positive and negative attitudes in cyberspace. This research suggests a method for locating such assaults. The method includes the use of deep learning models that had been trained on the present dataset using Bi Long Short-Term Memory (Bi LSTM). Our model beats more established machine learning techniques, according to the experimental data.The method includes the use of deep learning models that had been trained on the present dataset using Bi Long Short-Term Memory (Bi LSTM). Our model beats more established machine learning techniques, according to the experimental data. Experimental results showed that the proposed technique could achieve an accuracy of 96.7%, making it the best option for use in the detection of breaches applications.

Keywords – Long Short Term Memory, DDoS attack, SVM, Random Forest.

I. INTRODUCTION

A DDoS attack could be a challenging attempt to interrupt routine operations by deluging a sudden increase in Web traffic from a particular system, service, or network, hitting the intended system or its surrounding systems. DDoS attacks become stronger by utilizing a few vulnerable programming frameworks as their attack operation sources. Computer exploitation, IoT devices, and Many additional automated devices are feasible.DDoS detection is one of the main DDoS security tools. In any event, it is difficult to naturally recognize DDoS attacks because, in most circumstances, attack activity is quite similar to real activity and attackers try to imitate Striking clustersIn the early phases, Having minimal or no action during an aggressive operation could be interpreted as significant [1]. There is no actual host for attack activities since the DDoS attack has improved the common peer-to-peer attack technique. Rather, the attack makes use of ordinary behaviors and services. It is difficult to distinguish between an attack and normal behavior based solely on the protocols and facilities used. A distributed denial-of-service assault is difficult to detect [2]. The attack's features[3],were determined by taking into account three variables, including traffic density, the quantity of ports to be targeted, and the quantity of source IP addresses, which permitted many-to-one attacks during the denial of service (DDoS) assault phase.

These systems are capable of making decisions when the majority of attack flows are logical, but they only use a limited amount of message information—the majority of them relying just on source IP address and destination port information—and their rate of detection is poor.DDoS attacks are categorized in **Fig 1** in accordance with the attack's method, flow, operation, and deploymentWhen an intruder attempts to overwhelm a target with an overwhelming number of packets that originate from the attacker's computer, this is known as a direct assault. A deceptive assault employs a forged IP address to appear as though it is coming from a different computer, much like an indirect attack, which poses a threat to overwhelm the victim's machine. Attackers use a flood attack to transmit a massive amount of data to a system, preventing it from analyzing and approving authorized network activity.

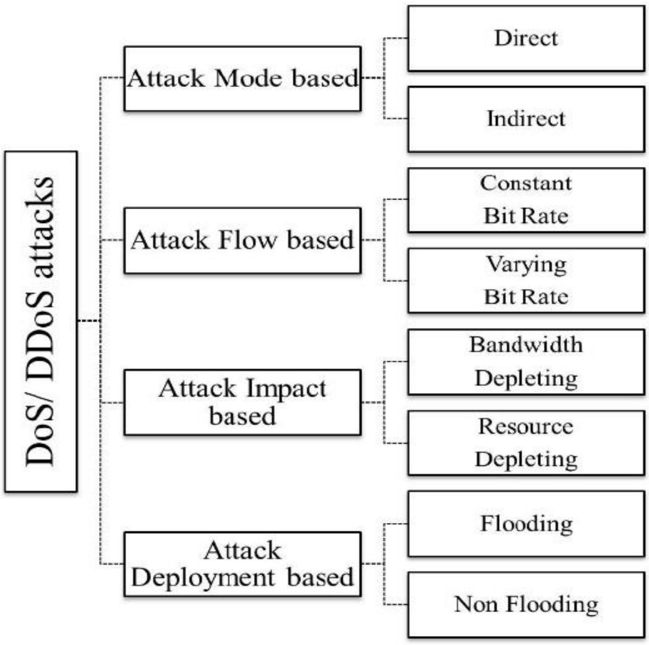


Fig 1. Classification of DDoS Attacks

As a result of the present research trend for machine learning-based DDoS attack detection, the majority of academics copy statistical machine learning techniques to recognizeDDoS attacks. When compared to the current methodologies, it has been discovered that these strategies typically perform better. The attack vector has limitations, the model and threshold value must be changed to account for alterations to equipment,and Attack strategies and a slow pace of attacks need to be considered. These challenges are just some of the ongoing issues[4].The implementation of a technique based on deep machine learning for identifying and thwarting DDoS attack packets on a distributed network using the Bi-Long Short-TermMemory network is the research's suggested remedy for the drawbacks. A Repetitive type that can learn or take on long-term dependencies is the LSTM[5].Our deep learning models are trained [6] using a large data set, the IDS ISCX 2012 dataset, to handle challenging recognition issuesThe remainder of the paper is structured as follows. Section 2 includes earlier investigations into DDoS attack detection with neural networks. We discussed the experimental setup used to train and validate the suggested Bi LSTM network-based DDoS detection in Section 3 of this paper. The details of the experiments carried out and the outcomes for the proposed model and the current best models are provided in Section 4. Results and discussions are given in Section 5. Finally, section 6 provides conclusions and suggestions for the future.

II. ASSOCIATED WORK

The effects of DDoS assaults on the source, intermediary, and destination networksare now being mitigated through research-based solutions that use both proactive and reactive strategies [7]. Common techniques include those for spotting attacks, evading attacks, and retaliating against assaults. Attack anticipation tries to direct entrance and exit traffic up until the attack injures people. Reducing DDoS attack losses is the fundamental objective of the attack response. The techniques used today to identify DDoS assaults that use AI are covered in this section. At the level of anomaly detection, a number of machine learning techniques are created expressly for DDoS avoidance.

To reduce security concerns, it has been proposed to use Deep learning models Educated on prior internet assault results to identify potentially dangerous connections and targets [8].Algorithms employed include Decision Tree, C4.5, Naive Bayes, and Bayesian Network.It was speculated that calculations made with deep learning algorithms might classify individuals who could be harmful to other users on the data plane. The basic channel model, an OSI reference model, is referred to in conventional network knowledge.Using a method for qualitative assessment that is outlined, assaults' varying frequency, particularly DDoS assaults, are calculated.[9].Additionally, a number of elements, like unpredictability and artificial neural networks, are used to identify such assaults[10].To specify the guidelines for routers and controlled data exchange in computernetworks, an innovative OpenFlow-based architecture is adopted. [11]And the proposed procedure for risk assessment [12].

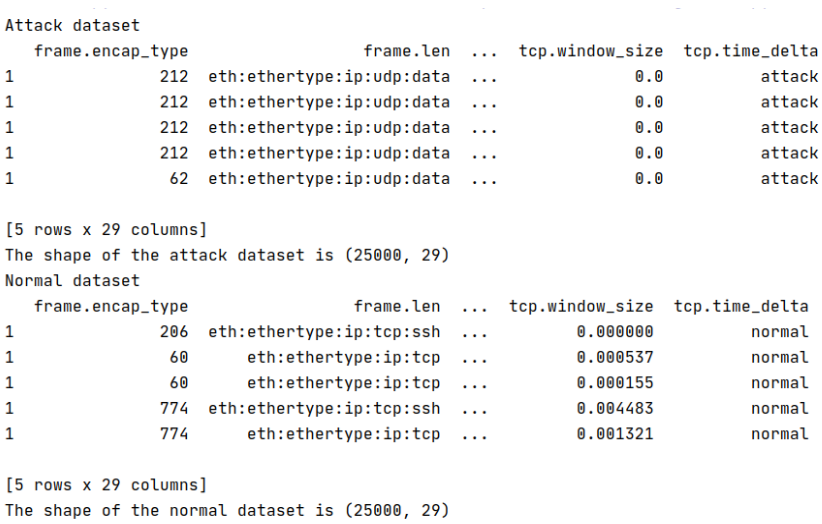
Based on controlled originating IP addresses, reinforcement learning and Hidden Markov Models (HMM) are presented [13] in order to make the distinction between legitimate traffic and DDoS attacks. Agents for detection are placed inside media network nodes or close to the DDoS assault origins. To calculate the likelihood of a particular observation sequence of new IP addresses, HMM is introduced. They assert that the bulk of source IP addresses used in a DDoS attack obscure the harmed party.Support Vector Machinehas proven its capability and proficiency in network classification, making it useful for locating DDoS identifications. [14].In [15], The Support Vector Machine and Genetic Algorithm are used in a method that is provided to recognizeDDoS.They can choose functions depending on GA and have access to more network traffic fields. They then label the packets using SVM in order to detect DDoS attacks.

III. EXPERIMENTAL FRAMEWORK

The datasets and measurement presumptions employed within the assessment system are described in this part.

Pre-processing of the Dataset and Data

Several public datasets have been frequently utilized to demonstrate and compare the efficacy and efficiency of different attack detection techniques. The capacity to recognize attacks is evaluated in this work using the IDS-ISCX-2012-dataset, is widely used in scholarly studies. The dataset has to be properly preprocessed in order to train more accurate models. The dataset is preprocessed using min-max normalization. The lowest value of each feature is translated to 0, the highest value to 1, and all other values are translated to a decimal value between 0 and 1. The IDS ISCX 2012 dataset's five rows of attack data and normal dataset were used to train and test the DDoS attack detection algorithm, are shown in **Fig 2.**



```
Attack dataset
   frame.encap_type                 frame.len  ...  tcp.window_size  tcp.time_delta
1               212  eth:ethertype:ip:udp:data  ...              0.0          attack
1               212  eth:ethertype:ip:udp:data  ...              0.0          attack
1               212  eth:ethertype:ip:udp:data  ...              0.0          attack
1               212  eth:ethertype:ip:udp:data  ...              0.0          attack
1                62  eth:ethertype:ip:udp:data  ...              0.0          attack

[5 rows x 29 columns]
The shape of the attack dataset is (25000, 29)
Normal dataset
   frame.encap_type                frame.len  ...  tcp.window_size  tcp.time_delta
1               206  eth:ethertype:ip:tcp:ssh  ...         0.000000          normal
1                60      eth:ethertype:ip:tcp  ...         0.000537          normal
1                60      eth:ethertype:ip:tcp  ...         0.000155          normal
1               774  eth:ethertype:ip:tcp:ssh  ...         0.004483          normal
1               774      eth:ethertype:ip:tcp  ...         0.001321          normal

[5 rows x 29 columns]
The shape of the normal dataset is (25000, 29)
```

Fig 2. IDS ISCX 2012 Dataset Rows

Evaluation Metrics

Accurate threat identification is made possible by properly detecting DDoS, which lowers the number of false alarms. For purposes of rating the DDoS attack detection task, the confusion matrix contains four categories: True Negative (TN), True Positive (TP), False Negative (FN), and False Positive (FP), accordingly. To evaluate the effectiveness of the learning algorithms, the accuracy is calculated using measurements taken from the confusion matrix. The specific calculation formula is displayed as follows.

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

IV. EXPERIMENTS AND RESULTS

Performance Evaluation with Existing Methods

In this part, the dataset is trained and evaluated for DDoS detection using machine learning techniques such as Multi-Layer Perceptron, Random Forest, Logistic Regression, SVM, K-Nearest Neighbours (K-NN), and Decision Tree. The algorithms listed above were chosen from the literature. The following table provides a brief overview of each algorithm as well as the confusion matrix that was obtained for detecting DDoS attacks using the IDS ISCX 2012 dataset.

Logistic Regression

Machine learning has incorporated the statistical approach of logistic regression. The machine learning approach can be used to segment a dataset Having at least one unrestricted variable that affects the result of the final logistic regression prediction. [16].This method categorizes observations into a specific set of classes. Concerns around classification include email spam vs. non-spam, internet payment fraud, and more. It uses the logistic sigmoid function and gives the output transformation's likelihood rating. It takes its name from the logistic equation, which serves as the main objective of the system. Like linear regression, logistic regression uses an equation as a representation. Comparable to linear regression, logistic regression analyses data using an equation. The created matrix of confusion for detecting DDoS assaults using logistic regression using the dataset is shown in **Fig 3.**

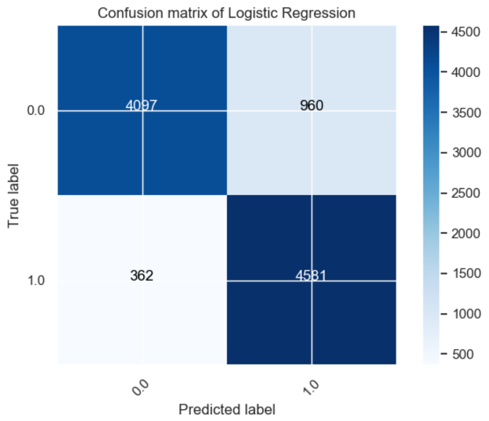


Fig 3. Logistic Regression Confusion Matrix Obtained.

SVM

In an N-dimensional space, Support Vector Machine (SVM) locates a hyperplane for classifying data values. To separate the two groups of data points, one could choose from a variety of possible hyperplanes. Our goal is to locate a plane that displays the greatest margin, or the greatest distance between the points of all classes. Maximizing the margin gap provides some support, allowing for increased trust in additional data points. The hyperplane will be produced iteratively by SVM in order to reduce error. SVM divides the datasets into groups with the goal of locating a maximal marginal hyperplane (MMH).The confusion matrix for SVM-based Detection of DDoS Attacks was created using the dataset is displayed in **Fig 4.**

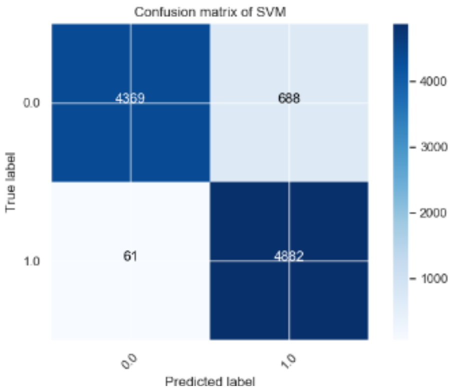


Fig 4. Confusion Matrix for SVM Obtained

Random Forest

Random Forest, as its name suggests, is an assembly of various independent decision trees that works as a single entity. When each individual tree spits out a class prediction, the class with the most votes becomes the model's prediction. The number of trees in the forest and the results it will produce are clearly and directly related; the more trees, the more precise the result. However, bear in mind that building a forest is not the same as making a choice based on gains through data or an index strategy. It is a commonly employed technique due to all of its benefits, which are stated below. Both classification and regression jobs can use it. But if there are adequate trees for the Random Forest method, the classifier would not overfit the model, which is a severe problem that will damage the results.Additionally, the Random Forest classifier can manage the absence of values. **Fig 5** displays the confusion matrix generated by the dataset for Random Forest-based DDoS attack detection.

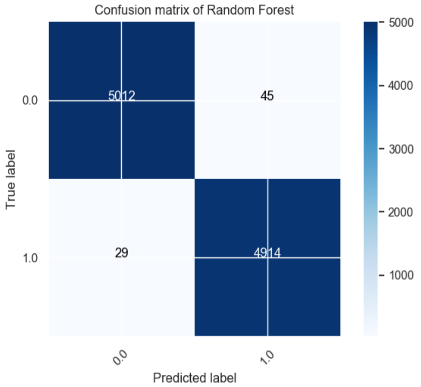


Fig 5. Confusion Matrix for Random Forest Obtained

K-Nearest Neighbors (K-NN)

It isthe most frequently Employed core machine learning algorithm for classification and regression. Data of user attributes in network packets are grouped based on user activity using the most basic version of kNN clustering.[17]. New data points are categorized based on measures of similarity (such as distance function). The shortest distance determines how close or far away a neighbor is. Data is distributed to the class with the nearest neighbors. A higher value of k will boost accuracy as the number of neighbors gets closer. This is done by using the class identifiers of the K-closest training examples in the test case. The confusion matrix, which was produced using the dataset for DDoS attack detection using K-NN, is shown in **Fig 6**.

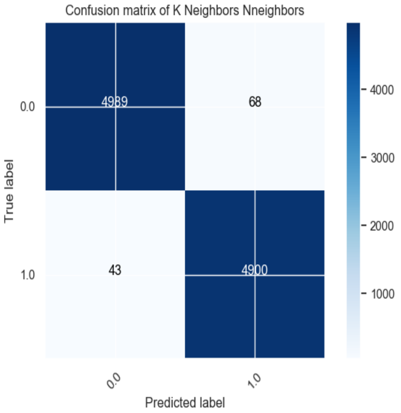


Fig 6. K-NN Confusion Matrix Obtained

Multi Layer Perceptron

A layer of input and a layer of output make up the two layers of a perceptron. Here, these layers are identical, but between the levels listed above, there can be further layers that are hidden. It has considerably increased computer processing speed when utilized to address classification and regression problemsIt is well known that it is possible for an input vector and its corresponding output vector to map in a nonlinear manner.Using the dataset, **Fig 7** displays the confusion matrix for Multi-Layer Perceptron-based DDoS assault detection.

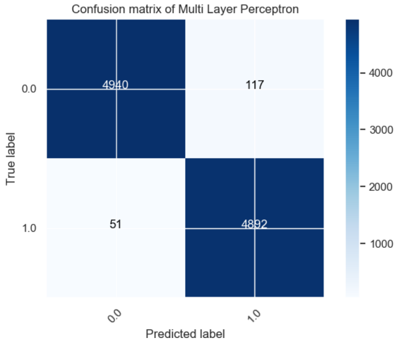


Fig 7. Confusion Matrix for Multi Layer Perceptrons Obtained

Decision Tree

Information is dissected by decision trees based on the features' If-then-else clauses. A decision node, a leaf node, and a branch are the three main parts. [18]. We deploy this method to predict a text's class grade by working our way down the tree from the top. A root attribute's values are compared to those of the record attributes.We continue to follow the branch that corresponds to that value based on similarity and the jump to the next node. **Fig 8** displays the confusion matrix for Decision Tree-based DDoS Attack Detection using the dataset.

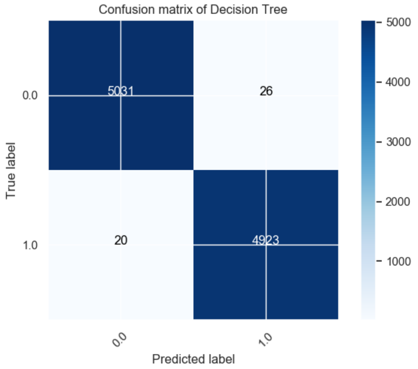


Fig 8. Confusion matrix for Decision Tree obtained Performance assessment using the proposed Bi-LSTM network

An effective deep neural network called an RNN employs loops to interact with sequence input by accessing its internal memory. RNNs struggle with issues like the vanishing gradients problem and the inflating gradients problem while learning long-term dependencies. An upgrade that takes care of the aforementioned problems is the LSTM-based models for RNNs. A long-term storage unit (LSTM) model collects significant input features. The assigned weights are used to determine whether to delete or keep the data. Thus, an LSTM model determines which data should be kept or deleted.

Recurrent neural networks (RNNs) that are bidirectional are just two distinct RNNs combined. This technique enables the networks to deliver both backward and forward knowledge about the sequence at any point in time. The bidirectional Long Short Term Memory is an enhancement of the Long Short Term Memory models and is created by adding input data to two LSTMs. An LSTM (also known as a forward layer) is initially added to the input sequence. In the second round (also known as the reverse layer), the LSTM model is given the inverse shape of the input sequence. By using Long Short Term Memory two Times, the model becomes more accurate because it is more dependent on long-term learning.**Fig 9.** shows the Bi-LSTM network architecture in action. The following is how the Bidirectional-Long Short Term Memory network is represented mathematically:

$$x = (x_1, x_2, \dots\dots, x_n) \tag{2}$$

$$\overrightarrow{h_t} = (\overrightarrow{h_1}, \overrightarrow{h_2}, \dots\dots, \overrightarrow{h_n}) \tag{3}$$

$$\overleftarrow{h_t} = (\overleftarrow{h_1}, \overleftarrow{h_2}, \dots\dots, \overleftarrow{h_n}) \tag{4}$$

$$y_t = [\overrightarrow{h_t}, \overleftarrow{h_t}] \tag{5}$$

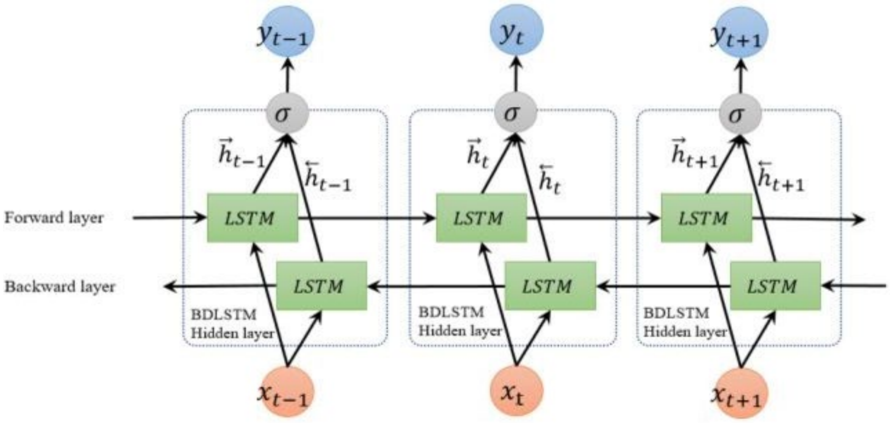


Fig 9. Bidirectional-Long Short Term Memory Architecture

For the same data set used for cutting-edge algorithms, The suggested network educating and verifying process is shown to be successful in **Fig 10** and **Fig 11**, respectively.

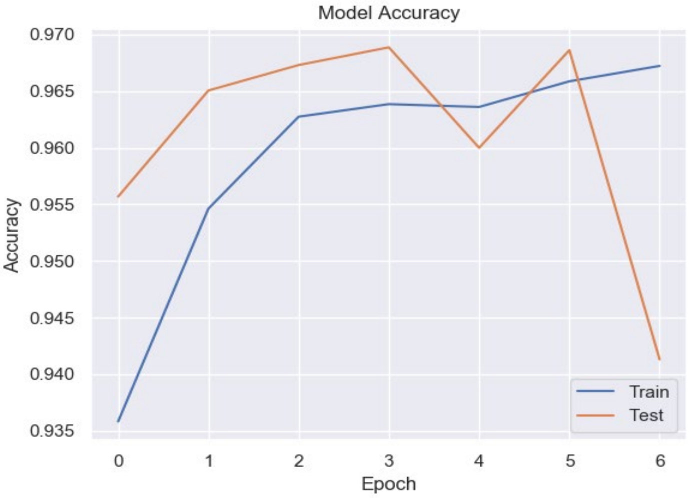


Fig 10. The Proposed Network's Model Accuracy During Educating and Verifying

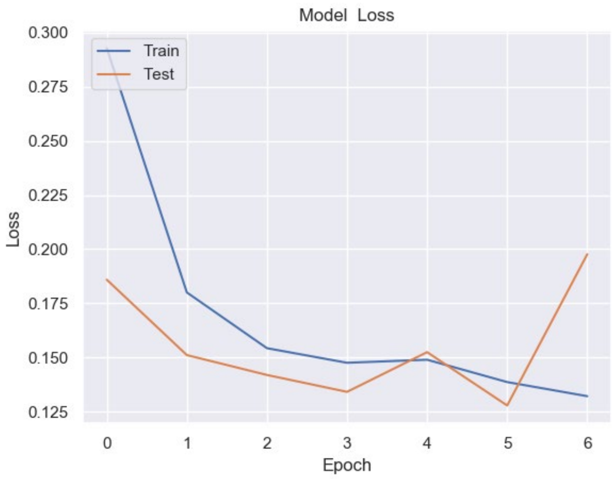


Fig 11. The Proposed Network's Failure Rate During Educating and Verifying

The proposed DDoS detection based on the confusion matrix of the Bidirectional-Long Short-Term Memory network is presented in **Fig 12** and **Fig 13** displays the Bi-LSTM network model's accuracy, which is 94.36 percent.

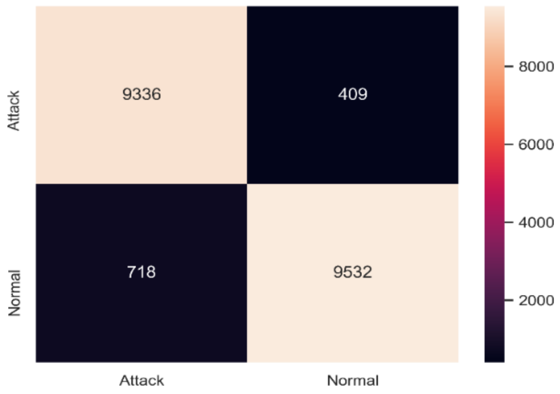
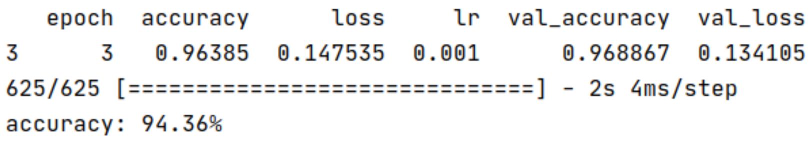


Fig 12. Confusion Matrix for Bi-LSTM Network-based DDoS Detection

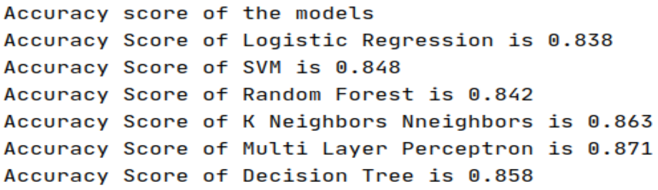


```
   epoch  accuracy      loss     lr  val_accuracy  val_loss
3      3   0.96385  0.147535  0.001      0.968867  0.134105
625/625 [==============================] - 2s 4ms/step
accuracy: 94.36%
```

Fig 13. Achieved Accuracy with Bi-LSTM Network

V. RESULTS AND DISCUSSION

The Proposed approach trains and tests for DDoS attack detection using the Bidirectional-Long Short-Term Memorynetwork. The state-of-the-art traditional machine learning methods are contrasted with the Bi-LSTM-basedDDoS detection model. The IDS-ISCX-2012 dataset is used to train a Bidirectional-Long Short Term Memory network with 10,000 inputs, and 20,000 inputs are used for testing. **Fig 14** displays the accuracy rating for the cutting-edge machine learning methods. It is clear that the Bi-LSTM network's accuracy score exceeds that of conventional machine learning techniques. The proposed method's accuracy is compared to state-of-the-art approaches in **Fig 15.**



```
Accuracy score of the models
Accuracy Score of Logistic Regression is 0.838
Accuracy Score of SVM is 0.848
Accuracy Score of Random Forest is 0.842
Accuracy Score of K Neighbors Nneighbors is 0.863
Accuracy Score of Multi Layer Perceptron is 0.871
Accuracy Score of Decision Tree is 0.858
```

Fig 14. Accuracy Score for The Cutting Edge Algorithms

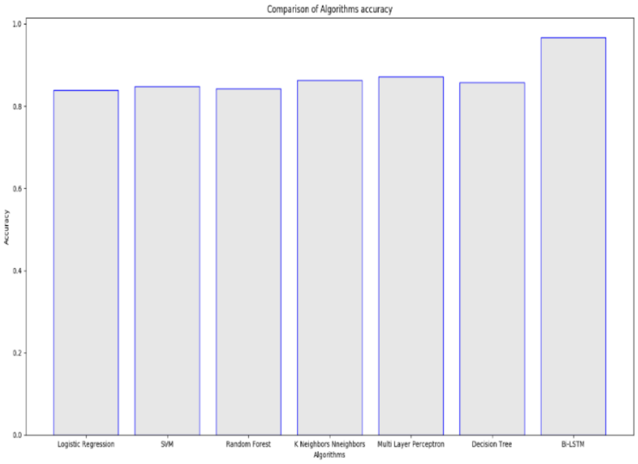


Fig 15. Comparison of the Suggested and the Existent Approaches' Accuracy.

VI. CONCLUSION

In this investigation, a Bi-LSTM-based DDoS detection method is suggested. The framework fixes a critical problem in conventional machine learning techniques by doing away with the necessity for manual function engineering. The system's capacity to automatically pick up on complex representations and define and attack traffic flows with accuracy is empirically tested.10,000 samples from the IDS-ISCX-2012 dataset are used to train the Bi-LSTM network, and 20,000 samples are used to complete testing. The results demonstrate that our suggested Bi-LSTM system outperforms the currently used learning methodologies in obtaining flawless performance in test situations. Without the inclusion of flow-level statistical data, the suggested technique would not only have competitive performance but also higher efficiency.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflicts of Interests

The author(s) declare(s) that they have no conflicts of interest.

Funding

No funding agency is associated with this research.

Competing Interests

There are no competing interests.

References

[1]. T. Peng, C. Leckie, and K. Ramamohanarao, “Survey of network-based defense mechanisms countering the DoS and DDoS problems,” ACM Computing Surveys, vol. 39, no. 1, p. 3, Apr. 2007, doi: 10.1145/1216370.1216373.

[2]. J. Mirkovic and P. Reiher, “A taxonomy of DDoS attack and DDoS defense mechanisms,” ACM SIGCOMM Computer Communication Review, vol. 34, no. 2, pp. 39–53, Apr. 2004, doi: 10.1145/997150.997156.

[3]. K. Sonar, and H. Upadhyay, “A survey: DDOS attack on Internet of Things,” International Journal of Engineering Research and Development, vol. 10, no. 11, pp.58-63, 2014.

[4]. X. Yuan, C. Li, and X. Li, “DeepDefense: Identifying DDoS Attack via Deep Learning,” 2017 IEEE International Conference on Smart Computing (SMARTCOMP), May 2017, doi: 10.1109/smartcomp.2017.7946998.

[5]. K. SaiSravani and P. Raja Rajeswari, “Prediction Of Stock Market Exchange Using LSTM Algorithm,” International Journal of Scientific and Technology Research, vol. 9, no. 3, pp.417-421, 2020.

[6]. https://gitlab.com/santhisenan/ids_iscx_2012_dataset

[7]. Y.-S. Choi, J.-T. Oh, J.-S. Jang, and J.-C. Ryou, “Integrated DDoS Attack Defense Infrastructure for Effective Attack Prevention,” 2010 2nd International Conference on Information Technology Convergence and Services, Aug. 2010, doi: 10.1109/itcs.2010.5581263.

[8]. L. Sun, Z. Li, Q. Yan, W. Srisa-an, and Y. Pan, “SigPID: significant permission identification for android malware detection,” 2016 11th International Conference on Malicious and Unwanted Software (MALWARE), Oct. 2016, doi: 10.1109/malware.2016.7888730.

[9]. M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, “An empirical evaluation of information metrics for low-rate and high-rate DDoS attack detection,” Pattern Recognition Letters, vol. 51, pp. 1–7, Jan. 2015, doi: 10.1016/j.patrec.2014.07.019.

[10]. K. Johnson Singh, K. Thongam, and T. De, “Entropy-Based Application Layer DDoS Attack Detection Using Artificial Neural Networks,” Entropy, vol. 18, no. 10, p. 350, Oct. 2016, doi: 10.3390/e18100350.

[11]. I. Mihai-Gabriel and P. Victor-Valeriu, “Achieving DDoS resiliency in a software defined network by intelligent risk assessment based on neural networks and danger theory,” 2014 IEEE 15th International Symposium on Computational Intelligence and Informatics (CINTI), Nov. 2014, doi: 10.1109/cinti.2014.7028696.

[12]. T. Chin, X. Mountrouidou, X. Li, and K. Xiong, “Selective Packet Inspection to Detect DoS Flooding Using Software Defined Networking (SDN),” 2015 IEEE 35th International Conference on Distributed Computing Systems Workshops, Jun. 2015, doi: 10.1109/icdcsw.2015.27.

[13]. X. Xu, Y. Sun, and Z. Huang, "Defending DDoS Attacks Using Hidden Markov Models and Cooperative Reinforcement Learning," Lecture Notes in Computer Science, pp. 196–207, doi: 10.1007/978-3-540-71549-8_17.

[14]. K. V S S R Murthy and K. V V Satyanarayana, "Intrusion detection mechanism with machine learning process A case study with FMIFSSVM, FLCFSSVM, misuses SVM, anomaly SVM and Bayesian methods," International Journal of Engineering &amp; Technology, vol. 7, no. 2.7, p. 277, Mar. 2018, doi: 10.14419/ijet.v7i2.7.10597.

[15]. Toeshik Shon, Yongdae Kim, Cheolwon Lee, and Jongsub Moon, "A machine learning framework for network anomaly detection using SVM and GA," Proceedings from the Sixth Annual IEEE Systems, Man and Cybernetics (SMC) Information Assurance Workshop, 2005., doi: 10.1109/iaw.2005.1495950.

[16]. V. Ramani Varanasi, "A Comparative Evaluation of supervised and unsupervised algorithms for Intrusion Detection," International Journal of Advanced Trends in Computer Science and Engineering, vol. 9, no. 4, pp. 4834–4843, Aug. 2020, doi: 10.30534/ijatcse/2020/9394202.

[17]. A. D. Jadhav and V. Pellakuri, "Intrusion Detection System Using Machine Learning Techniques for Increasing Accuracy and Distributed &amp; Parallel Approach for Increasing Efficiency," 2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA), Sep. 2019, doi: 10.1109/iccubea47591.2019.9128620.

[18]. D. Bhavana, K. Kishore Kumar, V. Chilakala, H. G. Chithirala and T. R. Meka, "A Comparison Of Various Machine Learning Algorithms In Designing An Intrusion Detection System," International Journal of Scientific and Technology Research, vol. 8, no. 12, pp.2407-2413, 2019.