Application of LSTM for Anomaly Detection in Fog-Cloud Computing

Mir Khalid Iqbal[1], Gagan Sharma[2]
[1]MTech Scholar, [2]Assistant Professor
Sri Satya Sai College of Engineering
RKDF University, Bhopal, Madhya Pradesh, India

Abstract. The cloud computing paradigm and services are extended by Fog computing over edge network to solve and reduce the existing problems associated with cloud computing. Fog computing was introduced to reduce latency wastage and to support the mobility of nodes independent of their location. It is a decentralized architecture to process incoming network traffic. With increasing traffic, there is an increase in network threats. To tackle such threats, fog network implements Intrusion Detection Systems (IDSs) as an integral part of security system to deploy secure network traffic flow. Due to the limitation of resources over the fog network, it is required to design lightweight IDS. In this paper, to provide efficient detection of incoming network traffic, a deep learning technique is proposed and implemented with a sliding window that identifies the intrusion. The result analysis was performed by increasing the window size and showed up its efficiency with other existing techniques. The result was evaluated over the NSL-KDD dataset. The testing scenario was performed over multi-classification of data packets. The sliding window helps to identify the type of intrusion over fog nodes. The highest accuracy of approx. 99% was achieved over the NSL-KDD dataset.

Keywords: — Fog Computing, Intrusion, Deep Learning, Cascaded LSTM, Sliding Window.

1 Introduction

Fog Computing has termed an extension of cloud computing which is designed to provide the effective infrastructure that supports IoT. Local processing is performed over Fog Computing as it acts as an intermediary processing unit between end-users and fog nodes. The network traffic enters into fog nodes for local processing and thus there is a need to provide security and authenticity to incoming network traffic. This is quite important to look after this concern as the nodes are vulnerable to malicious network attacks. As Fog nodes process all sensitive information such as financial, medical, etc. Attackers send malicious packets to target nodes to compromise them and to steal their information. So, it is quite necessary to recognize these attacks or intrusion and to provide secure and reliable services to the end-users. For this, this research work focuses on Intrusion Detection System (IDS) for fog nodes without compromising their efficiency [1]. Every day new threats are discovered and the existing database is not feasible. There is a need for regular updation. The anomaly detection technique showed up their proficiency in network intrusion detection. So, this paper is adopting anomaly-based detection for intrusion activities detection over fog computing. The security dynamics have shown that it has changed considerably. Different research works have shown the importance of the security aspect in fog computing along with the application of machine learning over it. Generally, existing techniques are supervised in nature. Initially, the machine learning approach was used for detection but still, there were some drawbacks. So, later the deep learning approaches have emerged as effective detection techniques. In this paper, an introduction about fog computing is provided along with their issues. Further in this paper contributions of researchers are provided based on machine learning as well as deep learning are given. Further, in this paper, we have proposed a sliding window deep learning approach for fog network traffic analysis about the intrusion. The historical data from fog nodes are used for the analysis of network traffic. The proposed method is implemented at fog nodes for historical data analysis. The remaining section of this paper are illustrated to be as follows: Section 2 introduced the existing techniques for network traffic analysis using fog computing along with that issues and challenges are also discussed. In section 3 chapter gives proposes an architecture to enhance performance. In section 4, the result and discussion along with comparative analysis were performed. Finally, in section 5, the conclusion of proposed methodology are discussed.

2 Related Work

Fog Computing appears to be a cloud services extension by creating an efficient IoT infrastructure required. Fog computing as a mediator offers personal connectivity for end-users and eliminates contact gaps seen between end-users and Fog systems mostly as network. Thus, the next validity of data traffic on Fog devices is highly dominant. Such resources are vulnerable to malicious attacks. Data among all kinds, in particular the transmission of financial and health data through such apps. The hackers exploit these devices through the sending of malicious data packets. It is crucial to monitor these infringements in order to provide the customer with a convenient and stable service. Efficient IDS (Intrusion Detection System) is therefore important for the safe operation of fog without sacrificing its quality. Sadaf and Sultana [2], using Isolation Forest (IF) and Auto encoder (AE) for Fog environment, introduced the intrusion detection method (auto-IF). This framework seems to be the only entity for the binary nomenclature of fog node that have to categorize attack patterns from packaging. We permit the established NSL-KDD dataset benchmark methodology. In contrast to several other methods of detecting modern intruders, our remote monitoring model is extremely accurate by 95.4%. Yang et al. [3] studied the detection of intrusion for fog computing in the F-RANs (fog radio access networks). Because fog nodes are resource-constrained, a conventional IDS (intrusion detection system) cannot be deployed directly in F-RANs because of the computational complexity and communication overhead. To find this issue, we introduce a skyline scheme based on a query that can inspect the statistics of the IDS log of fog nodes and give a complete flow for data processing. Specifically, a solution with three-step is introduced. Primarily, a strategy of filtering a lightweight fog node is introduced for filtering the raw data, which can decrease the issue of fog-cloud communication. Secondly, a mechanism based on a sliding-window is deployed in a cloud server for processing the asynchronous flow of data efficiently. then, a set of nodes attacked seriously will be identified by the skyline query vie using the pre-processed data. Thirdly, the level of security threat of each fog node is evaluated vie unascertained measures can evaluate the degree of security threat. It is one of the major problems to detect Intrusions that worry company in WSNs (wireless sensor networks). Several discoveries have dealt with this issue and have introduced several methodologies for the detection of various types of intrusions like selective forwarding- a serious attack that may hinder communications in WSNs. So, selective forwarding detection technique has suit as a key demand in the MWSNs (Mobile Wireless Sensor Networks) on spreading the applications of vehicular networks, internet of things (IoT), mobile computing immensely. Yaseen et al. [4] introduced the issues with the use of selective forwarding in MWSNs, and talk about how the techniques available for mitigating this issue in WSNs are not able to apply in problem handling in MWSNs because of the mobility of sensors. So this article introduces a model, provides a capability of global monitoring for detecting malicious ones and tracing moving sensors. The model takes the help of Fog Computing infrastructure for achieving this purpose. Moreover, the article provides a comprehensive discussion, complete algorithm, and experiments for showing the importance and correctness of the introduced approach. Pacheco et al. [5] proposed a Methodology of an Anomaly Behavior Analysis based on Artificial Neural Networks for implementing IDS (adaptive Intrusion Detection System) that are capable of detecting the Fog node while get compromised and then take actions that are required for ensuring communication availability. The outcomes of the experiment show that the introduced approach has capable of characterizing the Fog Nodes the normal behavior resist its complexity because of the adaptive scheme, and also can detect anomalies due to any type of sources like misuses, system glitches, or cyber-attacks, with low false alarms and high rate of detection.

3 Proposed Methodology

In this research, a deep learning-based approach was used to detect attacks based on the similarity of sliding windows that is capable of detecting the known type of attacks at fog nodes. The proposed model consists of two stages i.e., Stage 1: Learning and Stage 2: Detection.

Algorithm: Stage 1 (Learning)

1: Extract the features of incoming traffic.
2: Evaluate the similarities among them.
3: Set the sliding window on incoming data packets.
4: Learn the cascaded LSTM model.
5: Evaluate error
6: Increment the window.

Algorithm: Stage 1 (Detection)

7: Extract the features from incoming traffic

8: Set the sliding window on incoming data packets.
9: Evaluate the deviation from normal behavior.
10: Increment the window.
The data packets are taken as input from the dataset represented as:

$$Data_{sample} = \{(X_1, L_1), (X_2, L_2), (X_3, L_3) \ldots\ldots (X_N, L_N)\} \quad \text{(i)}$$

Where X_N = data packet sample
L_N = data label $\{L_N \epsilon (0,1)\}$ that represents the normal and anomaly traffic at fog nodes.

3.1 Packet Feature Extraction

In this step, the incoming data packets at fog nodes are collected and features are extracted and normalized that contains only numerical value because for effective processing of data samples on the same scale. In this paper, the sliding window is used over features as well as data samples. Every feature showed up its importance.

3.2 CNN Training

During training process of proposed methodology, the extracted features are fed into sliding cascaded models in which features are selected using sliding window. The window size decides the feature set. For anomaly detection cascaded LSTM neural network is used, which is a RNN structure that can be used to forecast time series sequence, to avoid the above impact and obtain considerably greater forecasting accuracy. When the back propagation through time (BPTT) approach is used, the LSTM network can avoid the dispersing gradient problem and record long-term dependence in time series.

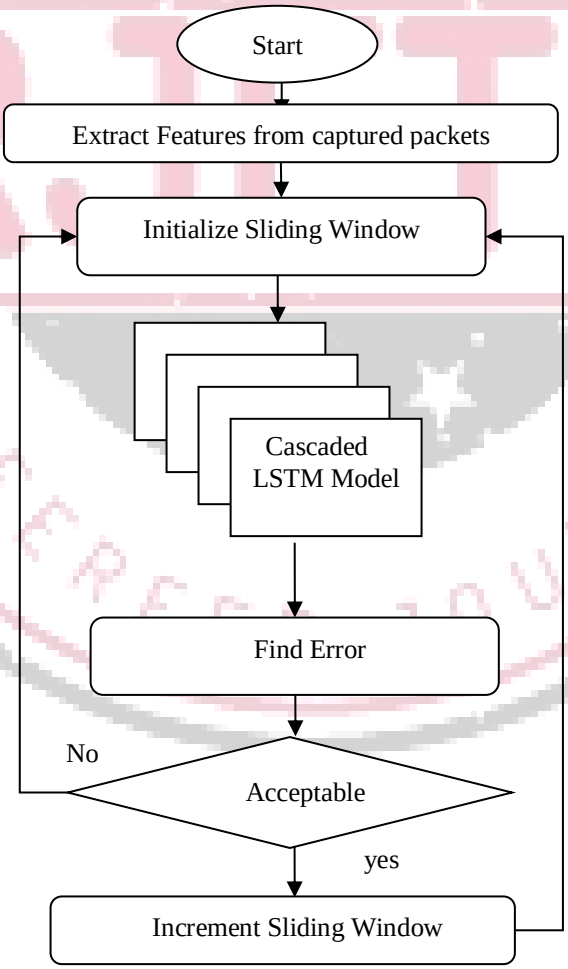

Fig. 1. Learning Flow Chart

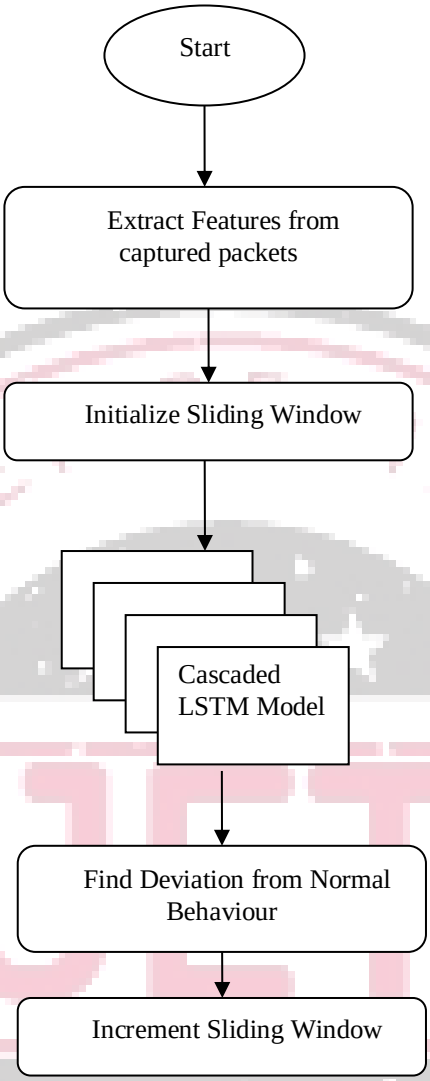

Fig. 2. Detection Flow Chart

4 Result Analysis

In this work, the NSL-KDD dataset is considered to be a benchmark for performance evaluation. Many recent works are performed by researchers for fog node intrusion detection and the NSL-KDD dataset was taken as a benchmark. For training, 10% of the dataset was taken and the other 10% are used for testing purposes. The entire simulation was performed on the MATLAB platform using a deep learning toolbox. The proposed work was evaluated on the basis of following parameters:

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN)*100 \quad \text{(ix)}$$

$$\text{Precision} = TP/(TP+FP)*100 \quad \text{(x)}$$

$$\text{Recall} = TP/(TP+FN)*100 \quad \text{(xi)}$$

$$\text{F\_Measure} = 2*\text{Precision}*\text{Recall}/(\text{Precision} + \text{Recall}) \quad \text{(xii)}$$

Where,

TP stands for true positive that means if actual and predicted data samples are an anomaly in nature, then TP is evaluated.

TN stands for true negative that means if actual and predicted data samples are not an anomaly in nature, then TN is evaluated.

FP stands for false positive that means if actual and predicted data samples are normal and anomaly in nature respectively, then FP is evaluated.

FN stands for false negative that means if actual and predicted data samples are an anomaly and normal respectively, then FN is evaluated.

Table 1 shows the performance evaluation of the proposed fog node intrusion detection system on the NSL-KDD dataset with variable sliding window size.

Table 1. Evaluation of Proposed Methodology

Window Size	Accuracy	Precision	Recall	F_Measure
1-5	99.47	99	99.5	99.4
5-10	99.2	94.9	100	97.4
10-15	99.2	99	98.51	98.8
15-20	99.9	99.5	100	99.8
20-25	99.9	100	99.9	99.9
25-30	99.5	99.5	99	99.3
30-35	99.5	98.5	100	99
35-40	99.3	99.4	98	98

4.1 State-of-Art Comparison

The fog nodes are vulnerable to network attacks, and network connectivity enables malware injection from the internet. In most of the attack detection learning model, vanishing gradient problem occurs and faces overfitting issues during the latter stages of training. Nowadays it has become one of the most promising research areas for researchers as daily new attacks are introduced in the network. This section is dedicated to exploring the work of other researchers in the field of intrusion detection. A comparative state-of-art with other existing work is illustrated over the NSL-KDD dataset. Table 2 and fig 3 show that the proposed algorithm has efficiency over existing techniques.

Table 2. Comparative Performance Evaluation

Ref	**Techniques**	**Accuracy (in %)**
C. Yin et al. [6]	RNN	83
Hawawreh et al. [7]	DNN	98
Sudqi et al. [8]	MLP	77
Sadaf et al. [1]	Autoencoder	95
Ours		99

5 Conclusion

The rising amount of network traffic poses a security risk for security breaches such as DoS attacks, etc. It calls for a safety solution to avoid such attacks. First, we need to consider a method for detecting attacks to prevent all kinds of attack. A deep learning-based approach (Cascaded LSTM) was used in this study to prevent threats groups on the basis of feature detection that can identify the form of threat known. This study designed to prevent responses based on a resemblance of the sliding windows, which can identify the recognized kinds of actions on fog nodes. In this paper, sliding window feature based Cascaded LSTM model is proposed for fog node prevention from network attacks. In this methodology, the incoming traffic on fog nodes are captured and analyzed using Cascaded LSTM model. The proposed approach was implemented for multi-classification in which anomaly was detected according to their types. The analysis of data packets was evaluated over fog nodes. For anomaly type analysis feature sliding window results in more effectively. For testing the methodology, NSL-KDD dataset is used for validation of proposed methodology over other state-of-art.

References

1. Abbasi, B. Z., & Shah, M. A. 2017. Fog computing: Security issues, solutions and robust practices. ICAC 2017 - 2017 23rd IEEE International Conference on Automation and Computing: Addressing Global Challenges through Automation and Computing. https://doi.org/10.23919/IConAC.2017.8082079
2. K. Sadaf and J. Sultana. 2020. Intrusion Detection Based on Autoencoder and Isolation Forest in Fog Computing. *IEEE Access*, 8, pp. 167059-167068.
3. Yang, S. 2017. IoT Stream Processing and Analytics in the Fog. IEEE Communications Magazine, 55(8), 21–27. https://doi.org/10.1109/MCOM.2017.1600840
4. Yaseen, Q.; Albalas, F.; Jararweh, Y.; Al-Ayyoub, M. 2016. A Fog Computing Based System for Selective Forwarding Detection in Mobile Wireless Sensor Networks. In *Proceedings - IEEE 1st International Workshops on Foundations and Applications of Self-Systems, FAS-W 2016*; Institute of Electrical and Electronics Engineers Inc. 256–262. https://doi.org/10.1109/FAS-W.2016.60.
5. Pacheco, J., & Hariri, S. (2016). IoT security framework for smart cyber infrastructures. Proceedings - IEEE 1st International Workshops on Foundations and Applications of Self-Systems, FAS-W 2016, 242–247. https://doi.org/10.1109/FAS-W.2016.58
6. C. Yin, Y. Zhu, J. Fei, and X. He, ‘‘A deep learning approach for intrusion detection using recurrent neural networks,’’ IEEE Access, vol. 5, pp. 21954–21961, 2017.
7. M. AL-Hawawreh, N. Moustafa, and E. Sitnikova, ‘‘Identification of malicious activities in industrial Internet of Things based on deep learning models,’’ J. Inf. Secur. Appl., vol. 41, pp. 1–11, Aug. 2018, doi: 10.1016/j.jisa.2018.05.002.
8. B. Sudqi Khater, A. W. B. Abdul Wahab, M. Y. I. B. Idris, M. Abdulla Hussain, and A. Ahmed Ibrahim, ‘‘A lightweight perceptron-based intrusion detection system for fog computing,’’ Appl. Sci., vol. 9, no. 1, p. 178, Jan. 2019, doi: 10.3390/app9010178.