Anomaly-based Intrusion Detection System using Bidirectional Long Short-Term Memory for Internet of Things

1st P. Jagdish Kumar
Department of CSE
Bharath Institute of Higher Education and Research
Chennai, Tamil Nadu, India
pjkcse123@gmail.com

2nd S. Neduncheliyan
School of Computing
Bharath Institute of Higher Education and Research
Chennai, Tamil Nadu, India
dean.cse@bharathuniv.ac.in

3rd Myasar Mundher Adnan
Department of Computers Techniques Engineering, College of Technical Engineering
The Islamic University
Najaf, Iraq
Maiser.monther@iunajaf.edu.iq

4th Sudhakar K
Dept. of AI & DS
Nitte Meenakshi Institute of Technology
Bengaluru, India
sudhakar.k@nmit.ac.in

5th A.V.V.Sudhakar
Department of Electrical and Electronics Engineering
SR University
Warangal, India
avv.sudhakar@sru.edu.in

2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE) | 979-8-3503-1860-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICDCECE60827.2024.10549281

***Abstract*—The developing use of Internet of Things (IoT) applications in several aspects created a large number of data which is need the occurrence of existing techniques like fog and cloud computing. The Intrusion Detection System (IDS) is recognized several cybersecurity problems and its resources. In this paper, the Bidirectional Long Short-Term Memory (BiLSTM) is proposed for anomaly-based IDS in IoT network. The BoT-IoT and ToN-IoT datasets are used for identifying IDS in IoT networks. The normalization is used as preprocessing and Gain Ratio is used for selecting optimal features and the BiLSTM is used for classification of IDS in network which enables the model to scale large dataset and enhance the accuracy. The performance of BiLSTM is estimated by accuracy, f1score, recall and precision. The BiLSTM achieves better accuracy of 98.76% and 96.84% for BoT-IoT and ToN-IoT dataset respectively which is better when compared to existing techniques like Group Theory Binary Spring Search-Hybrid Deep Neural Network (GTBSS-HDNN) and DNN.**

***Keywords*—bidirectional long short-term memory, deep neural network, internet of things, intrusion detection system, normalization**

I. Introduction

Nowadays, the development of computer system and Internet has affected by serious problems such as confidentiality, privacy and security issues during data transmission. It is definitely minimum network glitch trigger in regular operations of world [1]. Because of the convenience development in communication network industry, the gadgets are treated as physical body extension. The dependency makes a better possibility to enhance the attempts to illegal access for delicate privacy data in network [2, 3]. The recent data displays that most attackers targets a huge organization for gain the profit in minimum duration. The Intrusion Detection System (IDS) is used for device security and recognize the network intrusions [4]. To secure the communication, IDS is developed in which malicious network behavior is detected and managed to transfer alarm into system manager [5]. The IoT is an integration of cloud-connected embedded system used through the customer access service used the incorporation of electronic related internet protocols [6]. The IDS are efficient techniques to protect IoT systems from numerous attacks in which various systems are uses IDS for detect the malicious networks [7]. Main contribution of the research as follows:

- The Bidirectional Long Short-Term Memory (BiLSTM) is proposed for anomaly-based IDS in IoT network.
- The BoT-IoT and ToN-IoT datasets are used for identifying IDS in IoT networks. The normalization is used as preprocessing and Gain Ratio is used for selecting optimal features.
- The BiLSTM performance is estimated by accuracy, f1score, recall and precision which enables the model to scale large dataset and enhance the accuracy.

Rest of the research is arranged as follows: Section 2 elaborates literature review, section 3 provides proposed method, section 4 gives the experimental results and section 5 provides conclusion.

II. Literature Review

M. Shobana et al. [8] developed a Group Theory Binary Spring search (GTBSS)-Hybrid Deep Neural Network (HDNN) for IDS in IoT. Primarily, the privacy preserving approach was developed through block-chain based technique. It contained two stages such as blockchain and Modified Independent Component Algorithm (MICA) for avoid intrusion attacks. It was highly effective which detect intricate patterns. However, it leads computational complexity that limits scalability.

Monika Vishwakarma and Nishtha Kesswani [9] implemented a DNN based IDS for IoT network which recognize the malicious packets. The benchmark NetFlow-based dataset was used to train the model and the packet capturing and detection algorithm was developed in attack detection. The implemented model overcomes the overfitting issue which employed dropout layers for regularization. However, it was exposed to manage malicious data that lead misclassification.

979-8-3503-1860-9/24/$31.00 ©2024 IEEE

Shema Alosaimi and Saad M. Almutairi [10] presented an Ensemble Machine Learning approach for IDS in IoT network. The captured signals are transferred to other devices if it was an IDS in ML classifier for create accurate decision. Two techniques are used to operate the data, the primary one was used to minimize the data size and second one was used to resolve the class imbalance problem.

Muhammad Zeeshan et al. [11] suggested a Protocol-Based Deep Intrusion Detection (PB-DID) for IDS in IoT. The UNSW-NB15 and BoT-IoT datasets are used to recognize the common features among two networks and the union of dataset are fell in flow and Transmission Control Protocol (TCP). The suggested model enhances the identification accuracy and reduce the complexity. However, it required huge data for training.

Mohanad Sarhan et al. [12] introduced an Ensemble ML technique for IDS in IoT networks. The six various techniques such as Decision Tree (DT), Naive Bayes (NB), Logistic Regression (LR), Recurrent Neural Network (RNN), Deep Feed Forward (DFF), and Convolutional Neural Network (CNN). The feature extraction such as Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA) and Auto-Encoder (AE) are used for extracting features. However, it not considered the determination, heterogeneous traffic and range of communication networks.

III. Proposed Method

The BiLSTM is proposed for IDS in IoT network which includes two various datasets such as BoT-IoT and ToN-IoT. The normalization is used for preprocessing and Gain Ratio is used for feature selection and BiLSTM is used for classification. Figure 1 shows the process of proposed method.

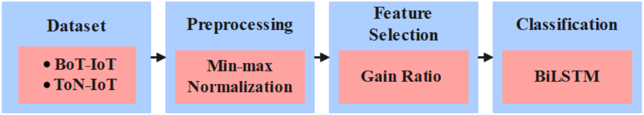

Fig. 1. Process of proposed method

A. Dataset

The BoT-IoT [13] and ToN-IoT [14] datasets are used for IDS in IoT networks. In BoT-IoT dataset, the normal and botnet traffic integrated and it comprises 72 million records in which 364562 records are utilized for dataset training and it minimized 243043 records for testing. The ToN-IoT dataset comprised 461043 annotations and 300000 for normal and 162043 for attack annotations.

B. Preprocessing

The min-max normalization is used for normalize the input data in the range of [0, 1] which is measured by using eq. (1),

$$\tilde{z}_{fj} = \frac{z_{fj}-min(z_f)}{max(z_f)-min(z_f)} \qquad (1)$$

Where, z_{fj} is a normalized value, $min(z_f)$ and $max(z_f)$ are minimum and maximum value.

C. Feature Selection

The Gain Ratio (GR) is used for feature selection that has benefits over Information Gain (IG) if it is not biased to feature. The IG results in overfitting and it select non-optimal features for prediction. The GR includes various number of branches while ranking the features. The GR value high if the information is distributed uniformly and less when every data comes to one branch. GR modifies IG by considering Intrinsic Information (II) of split in consideration. It described as data needed for define the branch into instances. The GR is a feature $f \in \{\mathcal{F}1, \mathcal{F}2, \mathcal{F}3, \dots, \mathcal{F}N\}$ is measured in eq. (2),

$$GR(f) = \frac{IG_f}{II_f} \qquad (2)$$

To estimate the GR for every feature f, the IG required to estimated in eq. (3),

$$IG(f) = (I_b) - (I_a) \qquad (3)$$

Where, I_b and I_a are a data feature before and after split, data feature before split is measured as eq. (4),

$$I_b(t) = Entropy(t) = -\sum_{i=1}^{k} p(v_i) log_2 p(v_i) \qquad (4)$$

Where, t is a target variable, p is a possibility and v is a value which measured the data for every feature through eq. (5),

$$I_a(t,f) = -\sum_{i=1}^{k} p(v) Entropy(v) \qquad (5)$$

Where, v is a probable value of feature f, the II is measured as eq. (6),

$$II(t,f) = -\sum_{i=1}^{k} \frac{|v_i|}{|v|} \times log_2 \frac{|v_i|}{|v|} \qquad (6)$$

The GR value falls among [0,1] because of the normalization, additionally, the GR value of feature f is 1, it denotes the data from f entirely predicts the t, if GR value is 0, it denotes there is no connection among f and t.

D. Classification

The Recurrent Neural Network (RNN) is a type of Artificial Neural Network (ANN) that provides to manage few significant persistently. The LSTM has benefits in terms of long-term dependency which is considered to provide huge context-related data with high memory than RNN. The BiLSTM is an advanced version of LSTM which incorporates two layers of LSTM. The LSTM one unit is accomplishing input in forward side and another LSTM unit performs input in backward side. The input, output and forget gate are essential blocks of LSTM unit which is given in eq. (7)-(9),

$$i_t = \sigma(x_t \times U_i + H_{t-1} \times W_i) \qquad (7)$$

$$o_t = \sigma(x_t \times U_o + H_{t-1} \times W_o) \qquad (8)$$

$$f_t = \sigma\left(x_t \times U_f + H_{t-1} \times W_f\right) \qquad (9)$$

Where, x_t is a present input state at time t, H_{t-1} is a hidden state of past time, W_i is an input weight matrix, U_o is an output weight matrix, W_o is an output weight matrix related with hidden state, U_f is a weight matrix of forget gate, W_f is a forget gate weight matrix related with hidden state. The cell update of input, forget and output gates are given in eq. (10)-(12),

$$C_t = f_t \times C_{t-1} + i_t \times N_t \qquad (10)$$

$$H_t = o_t \times \tanh(C_t) \tag{11}$$

$$State\ updae = \begin{cases} forget\ everything\ if\ f_t = 0 \\ forget\ nothing\ if\ f_t = 1 \\ update\ bias\ otherwise \end{cases} \tag{12}$$

Where, N_t is a new data measured by eq. (13),

$$N_t = \tanh(x_t \times U_c + H_{t-1} \times W_c) \tag{13}$$

The LSTM output is normalized through SoftMax layer with SoftMax function which works through H_t as presented in eq. (14),

$$\sigma(H_t) = \frac{e^{H_t}}{\sum_{j=1}^{k} e^{H_j}} \tag{14}$$

However, in IDS applications, it needs current and previous involvements as dual inputs for prediction. In BiLSTM, dual devoted LSTM network are incorporated for work by forward and backward side. The sensitive property to context creates the BiLSTM for produce best solution for difficult issues.

IV. Experimental Result

The BiLSTM is simulated by python with system configuration of 16GB RAM, windows 10 operating system and i7 processor. The performance of BiLSTM is estimated by accuracy, f1score, recall and precision which is given in eq. (15)-(18),

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{15}$$

$$F1 - score = 2 \times \frac{Precision \times recall}{Precision + recall} \tag{16}$$

$$Recall = \frac{TP}{TP+FN} \tag{17}$$

$$Precision = \frac{TP}{TP+FP} \tag{18}$$

The TP, TN, FN and FP is a true positive, true negative, false negative and false positive.

A. Performance Analysis

The BiLSTM performance is estimated accuracy, f1-score, recall and precision for BoT-IoT and ToN-IoT dataset. Table 1 shows Bi-LSTM performance for BoT-IoT dataset and table 2 shows Bi-LSTM performance for ToN-IoT dataset.

TABLE I. Performance of BiLSTM for BoT-IoT dataset

Method	Accuracy (%)	F1score (%)	Recall (%)	Precision (%)
CNN	92.46	91.75	91.56	92.41
RNN	94.51	93.48	94.29	93.38
GRU	95.73	94.64	94.37	95.56
LSTM	97.67	96.51	96.43	97.45
BiLSTM	98.76	97.37	97.68	98.29

Table 1 shows the BiLSTM performance by accuracy, f1-score, recall and precision for BoT-IoT dataset. The Convolutional Neural Network (CNN), RNN, Gated Recurrent Unit (GRU), and LSTM performance are compared with BiLSTM. The BiLSTM attains accuracy 98.76%, f1score 97.37%, recall 97.68% and precision 98.29% when compared to existing techniques.

TABLE II. Performance of BiLSTM for ToN-IoT dataset

Method	Accuracy (%)	F1score (%)	Recall (%)	Precision (%)
CNN	90.46	89.54	89.37	90.26
RNN	92.59	91.63	91.48	92.37
GRU	93.71	92.86	92.61	93.58
LSTM	95.61	94.22	94.18	95.39
BiLSTM	96.84	95.81	95.75	96.61

Table 2 shows the BiLSTM performance by accuracy, f1-score, recall and precision for ToN-IoT dataset. The CNN, RNN, GRU, and LSTM performance are compared with BiLSTM. The BiLSTM attains accuracy 96.84%, f1score 95.81%, recall 95.75% and precision 96.61% when compared to existing techniques.

B. Comparative Analysis

The BiLSTM performance is compared with existing methods like GTBSS-HDNN [8], and DNN [9]. The accuracy, f1-score, recall and precision are used to estimate the model performance. Table 3 shows the comparative analysis for BoT-IoT and ToN-IoT dataset.

TABLE III. Comparative analysis

Method	Dataset	Accuracy (%)	F1score (%)	Recall (%)	Precision (%)
GTBSS-HDNN [8]	BoT-IoT	96.23	96.70	97.03	95.94
	ToN-IoT	95.3	95.67	95.23	96.54
DNN [9]	BoT-IoT	83.22	80.66	83.82	78.64
	ToN-IoT	69.53	61.96	69.53	56.84
Proposed BiLSTM	BoT-IoT	98.76	97.37	97.68	98.29
	ToN-IoT	96.84	95.81	95.75	96.61

C. Discussion

The GTBSS-HDNN [8] leads computational complexity that limits scalability, and it attains accuracy 96.23% and 95.3% for BoT-IoT and ToN-IoT datasets. The DNN [9] was exposed to manage malicious data that lead misclassification and it attains accuracy 83.22% and 69.53% for BoT-IoT and ToN-IoT datasets. The proposed Bi-LSTM enables the model to scale large dataset and enhance the accuracy. Gain Ratio is used for selecting optimal features that has benefits over information gain if it is not biased to feature. The BiLSTM achieves better accuracy of 98.76% and 96.84% for BoT-IoT and ToN-IoT dataset.

V. Conclusion

In this paper, the BiLSTM is proposed for anomaly-based IDS in IoT network. The BoT-IoT and ToN-IoT datasets are used for identifying IDS in IoT networks. The normalization is used as preprocessing and Gain Ratio is used for selecting optimal features that has benefits over information gain if it is not biased to feature. The BiLSTM is used for classification of IDS in network which enables the model to scale large dataset and enhance the accuracy. The performance of BiLSTM is estimated by accuracy, f1score, recall and precision. The BiLSTM achieves better accuracy of 98.76% and 96.84% for BoT-IoT and ToN-IoT dataset respectively which is better when compared to existing techniques like GTBSS-HDNN and DNN. In future, optimization algorithm is used for feature selection to enhance the performance.

References

[1] Li, J., Othman, M.S., Chen, H. and Yusuf, L.M., 2024. Optimizing IoT intrusion detection system: feature selection versus feature extraction in machine learning. Journal of Big Data, 11(1), pp.1-44.

[2] Verma, P., Dumka, A., Singh, R., Ashok, A., Gehlot, A., Malik, P.K., Gaba, G.S. and Hedabou, M., 2021. A novel intrusion detection approach using machine learning ensemble for IoT environments. Applied Sciences, 11(21), p.10268.

[3] Prabu, S., Ganesh Babu, R., Sengupta, J., Perez de Prado, R. and Parameshachari, B.D., 2020. A block bi-diagonalization-based precoding for indoor multiple-input-multiple-output-visible light communication system. *Energies, MDPI AG*, *13*(13), pp.1-16.

[4] Apostol, I., Preda, M., Nila, C. and Bica, I., 2021. IoT botnet anomaly detection using unsupervised deep learning. Electronics, 10(16), p.1876.

[5] Idrissi, I., Boukabous, M., Azizi, M., Moussaoui, O. and El Fadili, H., 2021. Toward a deep learning-based intrusion detection system for IoT against botnet attacks. IAES International Journal of Artificial Intelligence, 10(1), p.110.

[6] Rani, D., Gill, N.S., Gulia, P. and Chatterjee, J.M., 2022. An Ensemble-Based Multiclass Classifier for Intrusion Detection Using Internet of Things. Computational Intelligence and Neuroscience, 2022.

[7] Udaichi, K., Chinaveer Nagappan, R., Garcia-Torres, M., Bidare Divakarachari, P. and Bhukya, S.N., 2023. Large-scale system identification using self-adaptive penguin search algorithm. *IET Control Theory & Applications*, *17*(17), pp.2292-2303.

[8] Shobana, M., Shanmuganathan, C., Challa, N.P. and Ramya, S., 2022. An optimized hybrid deep neural network architecture for intrusion detection in real-time IoT networks. Transactions on Emerging Telecommunications Technologies, 33(12), p.e4609.

[9] Vishwakarma, M. and Kesswani, N., 2022. DIDS: A Deep Neural Network based real-time Intrusion detection system for IoT. Decision Analytics Journal, 5, p.100142.

[10] Deepshikha, K., Yelleni, S.H., Srijith, P.K. and Mohan, C.K., 2021. Monte carlo dropblock for modelling uncertainty in object detection. *arXiv preprint arXiv:2108.03614*.

[11] Zeeshan, M., Riaz, Q., Bilal, M.A., Shahzad, M.K., Jabeen, H., Haider, S.A. and Rahim, A., 2021. Protocol-based deep intrusion detection for dos and ddos attacks using unsw-nb15 and bot-iot data-sets. IEEE Access, 10, pp.2269-2283.

[12] Sarhan, M., Layeghy, S., Moustafa, N., Gallagher, M. and Portmann, M., 2022. Feature extraction for machine learning-based intrusion detection in IoT networks. Digital Communications and Networks.

[13] BoT-IoT dataset link: https://www.kaggle.com/datasets/vigneshvenkateswaran/bot-iot (Accessed on 27/02/2024).

[14] ToN-IoT dataset link: https://www.kaggle.com/datasets/amaniabourida/ton-iot (Accessed on 27/02/2024).