IoT Network Intrusion Detection Using Long Short-Term Memory Recurrent Neural Network

Yee Mon Thant
Cyber Security Research Lab
University of Computer Studies, Yangon (UCSY)
Yangon, Myanmar
yeemonthant@ucsy.edu.mm

Mie Mie Su Thwin
Cyber Security Research Lab
University of Computer Studies, Yangon (UCSY)
Yangon, Myanmar
drmiemiesuthwin@ucsy.edu.mm

Chaw Su Htwe
Cyber Security Research Lab
University of Computer Studies, Yangon (UCSY)
Yangon, Myanmar
chawsuhtwe@ucsy.edu.mm

***Abstract*— The Internet of Things (IoT) is the latest technologies for everyday physical devices which the digitally connected to the internet or each other. With the rapidly development of IoT platform, IoT have been encountered many malicious activities. IoT security is one of the most critical issues in developing and implementing of IoT platform. Intrusion Detection System (IDS) play an important role for security solution in IoT network. In our paper, we proposed an effective IDS model for intrusion detection in IoT network by using Long Short-Term Memory Recurrent Neural Network (LSTM RNN). Performance of proposed model to identify correctly the normal and attack has been evaluated on the benchmark intrusion dataset, UNSW-NB15 dataset which applied in the most of IoT intrusion detection research. Moreover, in this survey, we studied the performance of the proposed LSTM RNN IDS model by contrasting the Recurrent Neural Network (simple RNN) algorithm. The experimental results show the efficiency of proposed model with accuracy, precision, recall and F1 score. Our proposed LSTM RNN model outperformed than the simple RNN model to build a highly effective intrusion detection model with accuracy over 99% for IoT network attacks.**

Keywords—Internet of Things, IoT, Intrusion Detection System, Long Short-Term Memory, Recurrent Neural Network

I. INTRODUCTION

The Internet of Things (IoT) security and privacy is the significant issues in modern day cyber security. The number of devices connected to the internet is expected to reach 50 billion worldwide at the end of 2030. Global IoT connections will reach 38.6 billion by 2025[1]. The Internet of Things (IoT) technologies play an important role to enhanced smart environment, such as smart home, smart city, smart healthcare, smart education, smart vehicles, smart industrial/ manufacturing and agriculture etc. The nature of IoT, large scale data, various platform and large number of devices are connected. The Internet of Things (IoT) systems are inclined to different cyber-attacks since of IoT gadgets have less computational capabilities, less power and constrained resources. In recent year, machine learning and deep learning are used as an advanced detection technique for network attack detection survey. The most commonly used algorithm for IoT security system of machine learning are Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and deep learning methods are Convolution Neural Network (CNN), Recurrent Neural Network (RNN) and restricted Boltzmann machine etc. [2].

Deep learning applications contain image recognition, computer vision, speech recognition, pattern recognition and self-driving cars etc. IoT solutions can be optimized that the deep learning algorithms with strong computation capabilities are more suit for complex and heterogeneous datasets compared to conventional machine learning strategies.

The main contributions of our research are

- To propose an effective IDS model for attacks detection in the Internet of Things (IoT) network by applying LSTM RNN.
- To analyze the deep learning application for the IoT network security by the performances evaluating on the intrusion dataset with simpleRNN and LSTM RNN algorithms.
- To discuss the development of deep learning based intrusion detection approach for binary classifications with high accuracy.

This paper is organized as follows. We will firstly introduce the related work on intrusion detection in the internet of things networks and conventional network based on machine learning, deep learning approaches, in Section II. And then, we present the proposed methodology for the intrusion detection in IoT networks utilizing the Recurrent Neural Networks models in Section III. In Section IV describes the intrusion dataset, experimental environment, data preprocessing, evaluation metrics and the analysis of experiment results obtained. Finally, the paper is concluded in Section V.

II. RELATED WORK

Machine learning and Deep learning approach are used in many related research work for the internet of things (IoT) malicious network detection techniques. In [3] , authors have proposed an intrusion detection mechanism based on Recurrent Neural Networks (RNN). Both binary and multi class classification are included in network intrusion detection, the benchmark NSL-KDD intrusion dataset are used in their proposed model performances evaluation. In performance evaluation, other machine learning methods such us Artificial Neural Network, Decision Tree, Random Forest, Support Vector Machine, are used to compare their proposed RNN-IDS model. Moreover, they have been proposed RNN model with high accuracy and low false positive rate to recognize the type of malicious.

According to [4], the authors utilized Long Short Term Memory (LSTM) method and evaluated on CIDDS dataset, in the implementation of anomaly based network intrusion detection. This research achieved 0.85% of accuracy with using rmsprop optimizer and learning rate 0.1 in the construction of multiclass classifier. Especially, their anomaly based network intrusion detection model is better than conventional methods, support vector machine, naïve bayes and multilayer perceptron for multiclass classification.

In [5], authors have presented a technique of anomaly based intrusion detection using Convolutional Neural Networks (CNN) and two types of Recurrent Neural Networks (RNN), LSTM-RNN and GRU-RNN are utilized on the NSL-KDD dataset. They found that CNN have outperformed the RNNs techniques in term of accuracy above 97%.In this work have some limitation that they need to tackle computation resources and long training time for their proposed model. Researchers in [6] discussed a network intrusion detection method on Stack Bidirectional Long Short Term Memory (LSTM) model employed the KDD Cup-99 dataset. In this research, the proposed IDS model conduct with three hidden layers and 80 LSTM units to intent the classification model is the best classified in binary and multiclass classification. The experimental result showed Stack Bidirectional LSTM surpassed than the CNN, LSTM and Multilayer Perceptron on accuracy over 91%.

In [7], the authors had been explained the IoT botnet detection was constructed by applying Long Short Term Memory Recurrent Neural Network model. The proposed model is compared with K Nearest Neighbor, Logistic Regression and Support Vector Machine and their model assessment was used the UNSW-NB15 dataset for training and testing the model. The developed model attained over 99% of accuracy and zero false alarm rate. In this paper, the author does not present the detail networks construction, especially the learning rate and number of epochs.

In [8], the authors proposed an IDS model for the internet of things(IoT) to optimize the network structure for the number of hidden layers and number of neuron in different attacks type. Their proposed model manipulated for IoT intrusion detection by launching the Genetic Algorithm with Deep Belief Network. The NSL-KDD intrusion dataset are concerned in the model evaluation. Comparison with different models and different attack types, detection in Distributed Denial of Services (DoS) got highest accuracy than the other attack types.

Author [9] have presented an IoT malicious activities detection framework by applying four deep learning methods, MultiLayer Perceptron(MLP), Convolutional Neural Networks(CNN) , Deep Neural Networks (DNN) and Autoencoder. The performance of the models evaluates in term of accuracy, RMSE and F1score by applying two intrusion benchmark datasets (NSL-KDD99 and UNSW-NB15). In their experiment, they found that DNN model accomplish superior out of the all other techniques in accuracy over 99%.

In this paper, we construct deep learning approach based effective intrusion detection method for the Internet of Things (IoT) network by applying Long Short-Term Memory Recurrent Neural Networks (LSTM RNN) model. The proposed model, LSTM RNN is compared with simple RNN model. We will use the benchmark UNSW-NB15 dataset to evaluate the proposed model performances because most of the IoT intrusion detection research based on this intrusion dataset. In addition, IoT intrusion detection researchers encounter the lack of IoT intrusion dataset. Moreover, Recurrent Neural Networks (RNNs) architectures have been employed in the numerous network intrusion detection system because of it suite for detection known and unknown attacks.

III. METHODOLOGY

A. Deep Learning in IoT Security

The combination of IoT framework and application with deep learning innovation is the current turned for IoT security. The good thing about deep learning above conventional machine learning strategies is stronger that the huge amount of IoT data.

Deep Learning have a capability automatically take out high level features from the dataset, so it is suite for the internet of things (IoT) technology that generate heterogeneous massive amount of data [10].

Deep Learning is motivated by the working components of the human brain and neurons for taking care of signals. With the approach of the deep learning have a place to different models, it models can be divided into three categories, generative, discriminative and hybrid. Generative model moreover called unsupervised learning and Deep Belief Networks, Restricted Boltzmann Machines are a few examples of generative networks. Convolutional Neural Networks and Recurrent Neural Networks are example of discriminative strategies. Furthermore, Recurrent Neural networks can be utilized for discriminative since the output of model can be applied as labeled input sequences. Generative Adversarial Networks is one of the cases of hybrid strategies.

B. Intrusion Detection System

When scanning the network activities to distinguish the malicious and normal behavior, intrusion detection system (IDS) is used as a security framework. There are two fundamental sorts of IDS, signature based (misuse detection) and anomaly based detection. Signature based approach is utilized to distinguish known assaults by utilizing pattern matching or rule-based strategies. Anomaly based approach is applied for detecting both known and unknown assaults by considering their behavior. Artificial neural network[11], machine learning[12]and deep learning[13] are the advanced intrusion detection methods. In these methods, deep learning based intrusion detection technique is better than the other detection techniques.

Technically, IDS is used for classification assignment, to distinguish whether any activity behavior is “normal” or “malicious”. Two type of classification work: binary classification and multiclass classification. In binary classification, the system distinguishes as it was one of two outputs (attack or normal). In multiclass classification distinguishes the sorts of assault categorically. In our paper, we utilized anomaly based detection for binary classification in IoT network intrusion detection.

C. Recurrent Neural Network(RNN)

One of the discriminative (supervised) deep learning methods is Recurrent Neural Network (RNN). RNN is the appendix of traditional Feed Forward Neural Network (FFNN) and it can handle the sequential data correctly. Distinctions with feed forward neural networks, RNN have repetitive connections making them powerful for modeling sequences. RNN is incorporated with several layers with feedback connections and is able to propagate previous information forward to the present time. When we train the RNNs, must be address the gradients vanishing and exploding problem because of it is the main issues of RNN algorithm[14]. In addition, RNNs can be deploying for the internet of Things security aspect. Detection system are needed to apply to detect several potential network malicious activities that the generation of the large volume of sequential data from various sources of IoT devices. Fig. 1 illustrates the typical structure of recurrent neural network. In Figure, x(t) is the input at time t, y(t) is the output at time t and current hidden layer is denoted by h(t) and h(t-1) is previous hidden layer. The hidden layers of RNNs has feed-back loop.

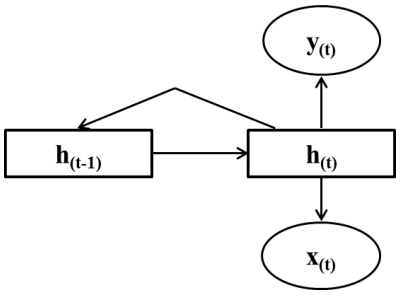

Fig. 1. Typical Structure of RNN

D. Long Short Term Memory

Long Short Term Memory (LSTM) networks are derived from the Recurrent Neural Networks (RNNs), which can easily remember pervious cell state. The prediction of Recurrent Neural Networks (RNNs) may be miss critical information from the starting of the pervious time step. LSTM was introduced in [15]to address the issue of RNN that cannot support long-term dependency problem. LSTM was designed to handle the vanishing gradient problem of RNN by using a gating mechanism. Three gates are contained in a LSTM memory cell; they are forget gate, input gate and output gate denoted with $\mathbf{f_t}$, $\mathbf{i_t}$ and $\mathbf{o_t}$. Forget gate determine what data should be admit as the current input from the previous output by passing through the sigmoid function which output between 0 and 1. If the value is 0, do not pass any information and if the value is 1, can be pass all information. Input gate chooses what data is included to the cell state and it compose of two function sigmoid and tanh function. Sigmoid function chooses what values got to be upgraded and tanh function creates the values of new vector to add to the cell state. Output gate elect by utilizing the sigmoid and tanh function what data is created from the current cell state as the output [16]. LSTM is enhanced to predict the time series given time slacks of unknown length. Fig.2 shows a Long Short-Term Memory(LSTM) cell [17]. The following equations are used to compute the value of three gates, cell state and output.

$$i_t=\sigma(W_{xi}x_t+W_{hi}h_{t-1}+W_{ci}c_{t-1}+b_i) \quad (1)$$

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right) \quad (2)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (3)$$

$$c_t = f_t c_{t-1} + i_c \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (4)$$

$$h_t = o_t \tanh(c_t) \quad (5)$$

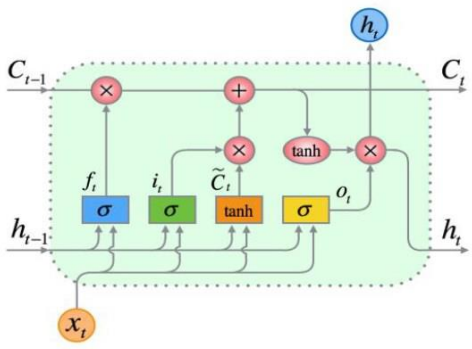

Fig.2. LSTM Cell

IV. Experiments

In our experiment, we used the Keras Python Deep Learning API [18]and Google TensorFlow library [19] on to simulate proposed RNNs models. We constructed LSTM model an input layer with 5 neurons, 4 LSTM memory units in a hidden layer and a single neuron for output layer that will produce only one output form attack or normal traffic. In our work, we used **binary crossentropy** as loss function that we need the attack or normal classification task. Although, we have been tried the various parameters especially the number of epochs, different optimizers and batch size for each experiment. The experiment result shown only better performances result with batch size 32. Fig.3 presents the overview of our detection process.

A. Intrusion Dataset

UNSW-NB15 dataset was designed in 2015 at the Cyber Security Lab of Australian Centre. It has been included the nine attack types (Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode and Worms).

It has composed of 49 features and diversity of normal and attack activities including with labeled data over two million records. It was supported the splitting training and testing set, the number of records in training set is 175,341 and testing set is 82,332 records with the diversity of attack types and normal traffic, described in Table I. These two dataset file can be download from [20].The UNSW-NB15 dataset have been chosen for our research work purpose is it contain diversity of modern attack patterns, IoT attacks and contains only two classes attack and normal. Moreover, the UNSW-NB15 dataset consist of sequential data, which is very appropriate for training our proposed recurrent neural network models.

1) *Attack types in UNSW dataset* [21].

a) *Fuzzers: In the field of cyber security, fuzzing attempt to suspend a program or network by obtaining the weakness of target system beside randomly generated the data.*

b) *Analysis: Analysis is the process of investigating the target network information by scanning and them attacker try to attack the target system.*

c) *Backdoors: Backdoor is a malware that can apply to obtain prohibited access to a computer or network.*

d) *Dos: Dos attacks try to harmed or shut down the system or network by flooding packet to the target system which cannot access the legitimate users.*

e) *Exploits: Attacker recognizes the defects of an operation system or chunk of data and used the information by abusing the vulnerabilities.*

f) *Generic: The generic attack also called birthday attack. It is to hash a large number of distinct messages prior to the same hash appear once more consideration to collision resistance.*

g) *Reconnaissance: Reconnaissance is a type of general information collecting attacks. Attacker try to attack the system known the vulnerabilities of the system by doing traffic scanning, packet sniffing etc.*

h) *Shellcode: Shellcode attack is a segment of code that can utilized as payload to explore the vulnerabilities of the system.*

i) *Worms: Worms malware can reproduce itself in arrange to distribute the one computer to others. It can frequently copy itself without human instruction when they once enter the target system.*

B. *Preprocessing*

In our research work, we used a comprehensive benchmark dataset, UNSW-NB15 dataset. This dataset contains 49 original features [21]. In [22], author proposed five attributes (service, sbytes, sttl, smean, ct_dst_sport_ltm) for improve the model accuracy and reduce the training time. Hence, we selected these feature set to improve our model performance. Python language does not support for multiple data type. So we need to convert non numerical (categorical data) to numerical. We used LabelEncoder for convert non numerical data to numerical data. And them, normalizer is used to scale the input data into a unique format. Data reshaping is the last work for data preprocessing state. This phase adjust features to ensure the data representation is suitable for our applied algorithm.

Table II shows five rows of the dataset features sample after preprocessing. The first five columns represent the features proposed according to [22]. In column six (attack_cat) shows the attack types including the normal samples, and the last column, label represents the value 0 for normal and 1 for attack samples. The data distribution of intrusion dataset for both training and testing are show in Fig.4 and Fig.5.

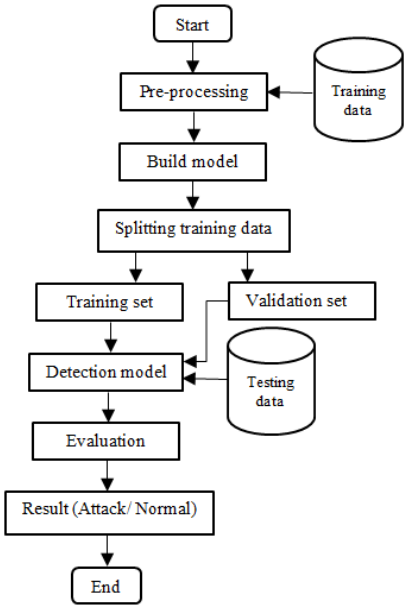

Fig.3. Overview of Intrusion Detection Process

C. *Experimental Environment*

The experiment was done by using Dell laptop Window 10 Pro 64 bit, Processor was Intel(R) Core (TM) i7-7500U CPU @ 2.70GHz 2.90 GHz. The memory for laptop was 8.00 GB and RADEON graphics card was used for experiment, does not apply Graphics Processing Unit (GPU). Pandas and Numpy frameworks are used for data analogy, cleaning and feature engineering. For data visualization used Matplotlib and Seaborn framework and Scikit-learn framework is used for data analysis. Python programming language with Keras deep learning libraries and Google Tensorflow frameworks have been widely used in research work for proficient implementation of RNNs models. Furthermore, our research is aim to solve the binary classification of malicious attack or normal, so we used loss function which is defined in Keras as binary_crossentropy.

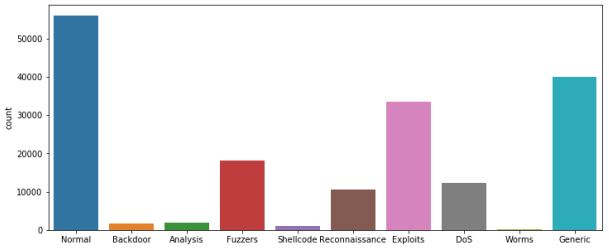

Fig. 4. Data distribution on Training set

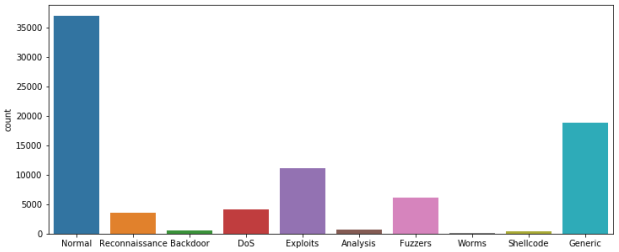

Fig. 5. Data distribution on Testing set

D. *Evaluation Metrics*

In our experiment, the performance of the proposed LSTM RNN model and simple RNN model are evaluated by applying the following metrics.

Accuracy: calculates the percentage of correctly classify to the number of total predictions.

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Precision: measures the ratio of correctly predicted the malicious packets to actual packets.

$$\text{Precision}=\frac{TP}{TP+FP} \quad (7)$$

Recall: measures what proportion of malicious packets were identified correctly.

$$\text{Recall}=\frac{TP}{TP+FN} \quad (8)$$

F1 score: calculates the harmonic mean of precision and recall, and gives a single weighted metric to assess the generally classification execution.

$$\text{F1 score}=\frac{2*(\text{precision} * \text{recall})}{\text{precision} + \text{recall}} \quad (9)$$

where, true positive denoted with TP, true negative denoted with TN, FP means false positive and FN means false negative.

Table III shows the used parameters for LSTM and simple RNN intrusion detection models. In our experiment, used the sigmoid as the activation function and also utilized the Adam optimizer.

E. *Experimental Result*

After the implementation of deep learning models, simple RNN and LSTM, we have been repeated experiments these model to obtain a satisfactory result. In our work, we examined the different learning rate with batch size 32 and various epochs to analyze the each model performance.

TABLE I. ATTACK DISTRIBUTION IN DATASET

	Training set	Testing set
Attack	119341	45332
Normal	56000	37000
Total	175341	82332

TABLE II. SAMPLE DATASET FEATURES AFTER PREPROCESSING

service	sbytes	sttl	smean	ct_dst_sport_ltm	attack_cat	label
smtp	37372	31	46	719	Normal	0
http	68199	254	1137	1	Fuzzers	1
-	200	254	100	4	Analysis	1
-	200	254	100	1	Backdoors	1
-	152	254	76	1	DoS	1
-	2202	254	1101	1	Exploits	1
dns	114	254	57	16	Generic	1
http	918	254	92	1	Reconnaissance	1
-	236	254	118	1	Shellcode	1
dns	114	254	57	16	Worms	1

Firstly, we choose the default learning rate (0.001), batch size 32, “adam” optimizer and “sigmoid” activation function on epoch (200, 500, 1000) for each experiments. All of these experiment, LSTM got above 93% on accuracy and 99% recall value with epoch 1000, show in Table IV. In Table V presents the performance result that we used the learning rate (0.01) in batch size 32 with same optimizer and activation function on various epochs. In this case, RNN obtain the accuracy only 83.68% when LSTM accuracy is being above 98% in epoch 1000.

On the other hand, the performance result of our implemented IDS model via LSTM networks is better than the other methods on the accuracy, precision, recall and f1 score above 99% in epoch 1000, learning rate 0.1, presented in Table VI. We observed that these two model superior on changing the learning rate than using the default learning rate on increasing the epoch numbers. In the performance evaluation of two models, LSTM is always excellently superior to simple RNN expect in learning rate 0.1, batch size 32 with epoch 200.

TABLE III. DEEP LEARNING MODEL PARAMETERS

Algorithm	Batch Size	Optimizer	Activation Function	Epochs
LSTM	32	adam	sigmoid	200,500,1000
RNN	32	adam	sigmoid	200,500,1000

TABLE IV. PERFORMANCE METRICS FOR BATCH SIZE 32 WITH DEFAULT LEARNING RATE (0.001)

Algorithm	Epoch	Accuracy	Precision	Recall	F1score
RNN	200	81.99 %	78.58 %	92.48 %	84.97 %
LSTM	200	84.37 %	78.36 %	99.16 %	87.54 %
RNN	500	82.61 %	79.34 %	92.49 %	85.41 %
LSTM	500	82.88 %	78.17 %	95.59 %	86.01 %
RNN	1000	83.78 %	79.63 %	94.77 %	86.54 %
LSTM	1000	93.37 %	89.80 %	99.23%	94.28 %

TABLE V. PERFORMANCE METRICS FOR BATCH SIZE 32 WITH LEARNING RATE (0.01)

Algorithm	Epoch	Accuracy	Precision	Recall	F1score
RNN	200	82.13 %	78.77%	92.43%	85.06%
LSTM	200	87.65 %	84.27%	95.38%	89.48%
RNN	500	81.56 %	77.98%	92.66%	84.69%
LSTM	500	97.88 %	96.64%	99.61%	98.10%
RNN	1000	83.68 %	79.44%	94.92%	86.49%
LSTM	1000	98.12 %	97.01%	99.64%	98.31%

TABLE VI. PERFORMANCE METRICS FOR BATCH SIZE 32 WITH LEARNING RATE (0.1)

Algorithm	Epoch	Accuracy	Precision	Recall	F1score
RNN	200	88.49%	85.29%	95.57%	90.13%
LSTM	200	83.28%	76.76%	99.88%	86.8%
RNN	500	91.76 %	89.05%	96.95%	92.83%
LSTM	500	98.76%	97.84%	99.95%	98.89%
RNN	1000	82.43%	76.46%	98.36%	86.04%
LSTM	1000	99.32%	98.82%	99.94%	99.38%

The binary classification of the models performance was used accuracy and loss function. The experimental result of confusion matrix for LSTM and RNN model tested on 82332 samples with epoch 1000 and learning rate 0.1 are as shown in Fig.6 and Fig.7 respectively. In Fig.8 shows the LSTM model accuracy on the number of epochs for training and testing set. Plot figure for LSTM model loss as shown in Fig. 9.

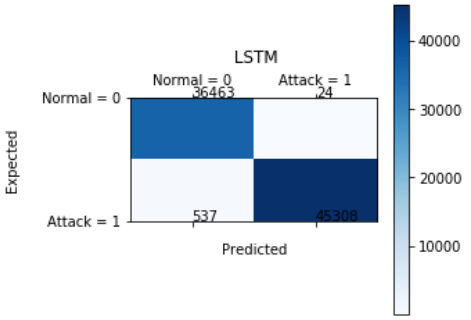

Fig. 6. Confusion matrix of attack or normal classification for LSTM model

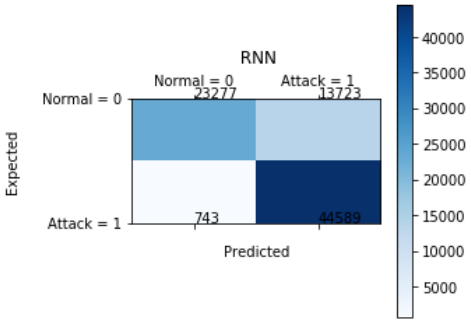

Fig. 7. Confusion matrix of attack or normal classification for RNN model

Fig. 10 and Fig.11 demonstrate the accuracy and loss for RNN model in epochs 1000 with batch size 32 respectively. When we compare with LSTM and RNN, RNN accuracy is very unstable; it is obviously fluctuate in testing set. RNN loss value is larger than LSTM loss; it is in terms to the behavior of model is poorly in each iterations of optimization. Our proposed LSTM IDS model is more effective detecting of the IoT intrusions when it deployed the Adam optimizer than other optimizers (RMSprop, Adagrad, Adamax, Nadam, and SGD).

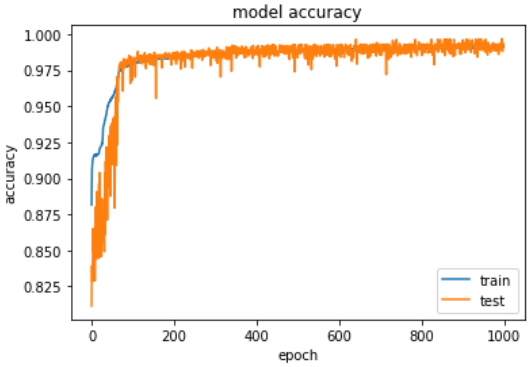

Fig. 8. LSTM model accuracy on epochs

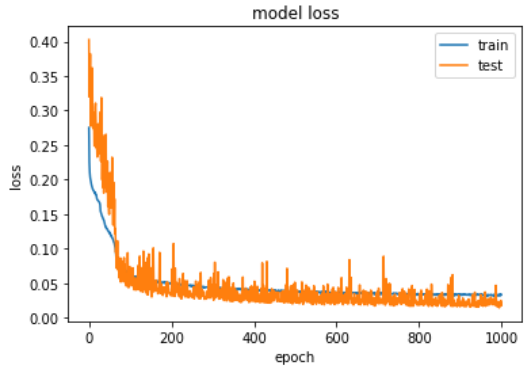

Fig. 9. LSTM model loss on epochs

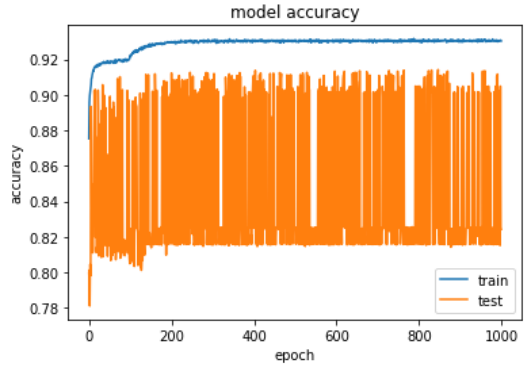

Fig. 10. RNN model accuracy on epochs

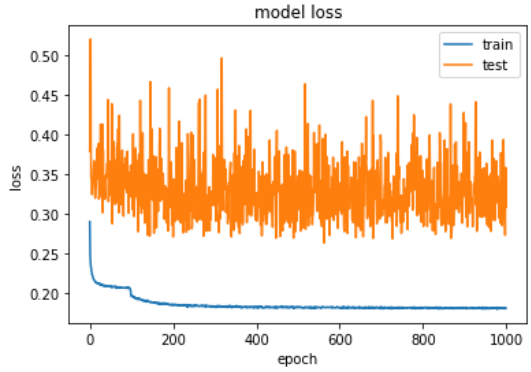

Fig. 11. RNN model loss on epochs

V. Conclusion And Future Work

In recent year, combining of deep learning with IoT network intrusion detection technique is a trend. In this paper, we proposed an IDS model by applying the Long Short-Term Memory Recurrent Neural Networks model (LSTM RNN) in the detecting of IoT intrusions. We have implemented our proposed model with best five attributes of UNSW NB-15 dataset to enhance the model accuracy and reduce the model training time. The experimental values are changed to find a best performances result. Moreover, we have been analyzed the model performance in binary classification and different learning rate, batch size 32 and iterations (epoch) impacts on the accuracy of the model. In the current study based on deep learning approach, LSTM RNN has be outperformed compare with simple RNN technique in accuracy over 99%. Consequently, we proved that LSTM RNN model is more effective than the conventional RNN model in the detection of IoT intrusions. In future work, Bidirectional Long Short-Term Memory Recurrent Neural Network (BLSTM RNN) will be used for network intrusions detection based on a realistic IoT intrusion dataset. In addition, we will employ the proposed model with other optimizer in both binary and multi-class classification.

References

[1] "Global IoT connections to reach 50 billion by 2030: study." https://enterpriseiotinsights.com/20190520/internet-of-things/global-iot-connections-reach-50-million-2030-study (accessed Nov. 04, 2020).

[2] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security," *IEEE Commun. Surv. Tutorials*, vol. 22, no. 3, pp. 1646–1685, 2020, doi: 10.1109/COMST.2020.2988293.

[3] 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, doi: 10.1109/ACCESS.2017.2762418.

[4] S. A. Althubiti, E. M. Jones, and K. Roy, "LSTM for Anomaly-Based Network Intrusion Detection," in *2018 28th International Telecommunication Networks and Applications Conference (ITNAC)*, Sydney, NSW, Nov. 2018, pp. 1–3, doi: 10.1109/ATNAC.2018.8615300.

[5] S. Al-Emadi, A. Al-Mohannadi, and F. Al-Senaid, "Using Deep Learning Techniques for Network Intrusion Detection," in *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, Doha, Qatar, Feb. 2020, pp. 171–176, doi: 10.1109/ICIoT48696.2020.9089524.

[6] Z. Ran, D. Zheng, Y. Lai, and L. Tian, "Applying Stack Bidirectional LSTM Model to Intrusion Detection," *Computers, Materials & Continua*, vol. 65, no. 1, pp. 309–320, 2020, doi: 10.32604/cmc.2020.010102.

[7] Jenny Costa and Shree Rayeshwar Institute of Engineering and Information Technology, "IoT-Botnet Detection using Long Short-Term Memory Recurrent Neural Network," *IJERT*, vol. V9, no. 08, p. IJERTV9IS080236, Aug. 2020, doi: 10.17577/IJERTV9IS080236.

[8] Y. Zhang, P. Li, and X. Wang, "Intrusion Detection for IoT Based on Improved Genetic Algorithm and Deep Belief Network," *IEEE Access*, vol. 7, pp. 31711–31722, 2019, doi: 10.1109/ACCESS.2019.2903723.

[9] A. Nagisetty and G. P. Gupta, "Framework for Detection of Malicious Activities in IoT Networks using Keras Deep Learning Library," in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, Mar. 2019, pp. 633–637, doi: 10.1109/ICCMC.2019.8819688.

[10] H. Li, K. Ota, and M. Dong, "Learning IoT in Edge: Deep Learning for the Internet of Things with Edge Computing," *IEEE Network*, vol. 32, no. 1, pp. 96–101, Jan. 2018, doi: 10.1109/MNET.2018.1700202.

[11] A. Shenfield, D. Day, and A. Ayesh, "Intelligent intrusion detection systems using artificial neural networks," *ICT Express*, vol. 4, no. 2, pp. 95–99, Jun. 2018, doi: 10.1016/j.icte.2018.04.003.

[12] A. Halimaa A. and K. Sundarakantham, "Machine Learning Based Intrusion Detection System," in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, Apr. 2019, pp. 916–920, doi: 10.1109/ICOEI.2019.8862784.

[13] L. Ma, Y. Chai, L. Cui, D. Ma, Y. Fu, and A. Xiao, "A Deep Learning-Based DDoS Detection Framework for Internet of Things," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, Dublin, Ireland, Jun. 2020, pp. 1–6, doi: 10.1109/ICC40277.2020.9148944.

[14] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training Recurrent Neural Networks," p. 13.

[15] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[16] M. Phi, "Illustrated Guide to LSTM's and GRU's: A step by step explanation," *Medium*, Jun. 28, 2020. https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21 (accessed Nov. 05, 2020).

[17] Z. Cui, R. Ke, and Y. Wang, "Deep Stacked Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-wide Traffic Speed Prediction," p. 12.

[18] "Keras: the Python deep learning API." https://keras.io/ (accessed Nov. 05, 2020).

[19] "TensorFlow," *TensorFlow*. https://www.tensorflow.org/ (accessed Nov. 05, 2020).

[20] "UNSW Canberra." https://www.unsw.adfa.edu.au/ (accessed Jun. 30, 2020).

[21] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *2015 Military Communications and Information Systems Conference (MilCIS)*, Canberra, Australia, Nov. 2015, pp. 1–6, doi: 10.1109/MilCIS.2015.7348942.

[22] T. Janarthanan and S. Zargari, "Feature selection in UNSW-NB15 and KDDCUP'99 datasets," in *2017 IEEE 26th International Symposium on Industrial Electronics (ISIE)*, Edinburgh, United Kingdom, Jun. 2017, pp. 1881–1886, doi: 10.1109/ISIE.2017.8001537.