Deep Learning Models for Cyber Security in IoT Networks

Monika Roopak
School of Engineering
Newcastle University, UK

Prof. Gui Yun Tian
School of Engineering
Newcastle University, UK

Prof. Jonathon Chambers
School of Engineering
Newcastle University

Abstract: **In this paper we propose deep learning models for the cyber security in IoT (Internet of Things) networks. IoT network is as a promising technology which connects the living and non-living things around the world. The implementation of IoT is growing fast but the cyber security is still a loophole, so it is susceptible to many cyber-attack and for the success of any network it most important that the network is completely secure, otherwise people could be reluctant to use this technology. DDoS (Distributed Denial of Service) attack has affected many IoT networks in recent past that has resulted in huge losses. We have proposed deep learning models and evaluated those using latest CICIDS2017 datasets for DDoS attack detection which has provided highest accuracy as 97.16% also proposed models are compared with machine learning algorithms. This paper also identifies open research challenges for usage of deep learning algorithm for IoT cyber security.**

Keywords: IoT, DDoS, Deep Learning, CNN, LSTM, RNN CICIDS2017

I. INTRODUCTION

Internet of Thing is latest emerging promising technology which connects everything around the world through internet. IoT technology guarantee to improve and help our personal, professional life and society[1]. IoT consist of network of smart objects around the world through internet without any human interference, which is great but it is susceptible to cyber attacks like any other network. Intrusion Detection System (IDS) is an effective technique for the detection of cyberattacks in any network. Most of the latest IDS are based on machine learning algorithm for the training and detecting cyber-attack on the network. Fog computing is improved extension of centralized Cloud computing in which the distributed fog nodes are closer to the IoT network objects and it resolves the scalability bottlenecks, high bandwidth consumption QoS (Quality of Service) abasement and limitation of high latency in the cloud computing. Fog-to-node computing is ideal for the practical implementation and success of the IoT networks. Fig. 1 illustrates the architecture of fog-to-node model with distributed parallel computation providing intelligence to the distributed fogs by providing computation, control and storage of IDS closer to IoT network objects. IDS detects cyber-attack efficiently and quickly at fog nodes in comparison with cloud. IoT network consist of connections among different kinds of smart objects ranging from super computers to tiny devices which may have very low computation power, so securing such kind of network is challenging and hence cyber security is a big

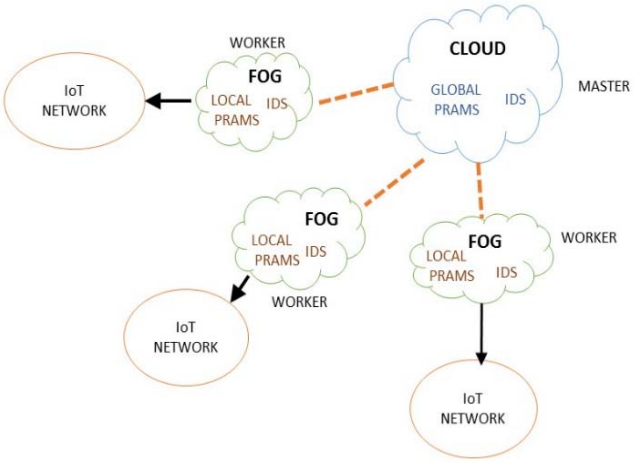

Fig. 1. Fog-to-Node architecture for IoT network

loophole in the implementation of IoT network [2]. DDoS is one of the major cyber-attack that affected IoT network in recent past and has resulted in massive loses [3]. In DDoS attack hacker uses number of hosts to overwhelm target server which result in complete crash down of system and hence the legitimate users are not able to access the service on the server system. According to information provided in [4] the denial of service attack will reach 17 million by 2020.

Deep learning is broader subfield of machine learning which is actually a larger deep neural network and can be employed for supervised, unsupervised and semi supervised learning. The concept of deep learning method was first proposed in [5] based on deep belief network and it has been proven to be highly effective in the fields such as image processing, natural language processing and self-driving car etc. One of limitation of deep learning methods is longer training time it requires as larger the training data, greater will be the training time but deep learning methods need huge data for training for performing well.

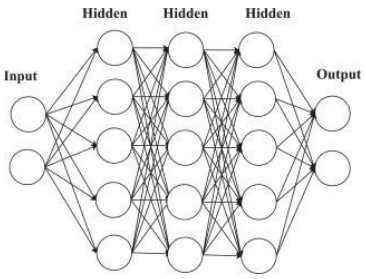

Fig. 2. Architecture of deep learning model

978-1-7281-0554-3/19/$31.00©2019 IEEE

The basic architecture of deep learning model is shown in Fig. 2 it consist of one input layer followed by number of hidden layers which further fed input to the output layer. CNN (Convolutional Neural Network) is a deep learning model which has been used extensively in the field of computer image processing[6, 7]and language processing[8]. Raw image is fed to CNN model directly without any pre-processing, it than evaluate features by convolution operations [9]. RNN(Recurrent neural network) is another type of deep learning model that has provided promising progress in the fields such as NLP(natural language processing) [10] and text processing [11]. LSTM (Long Short Term Memory) network is evolution of RNN network which is cable of learning patterns in ling sequences, this can be used to classify data as attack and normal. One of the advantage of LSTM is that it can be applied directly on raw data without applying any feature selection method. This paper compare performance multilayer perceptron, CNN, LSTM and hybrid CNN+LSTM model for the detection of cyberattack in the IoT network on a centralized system.

This paper aim to 1) propose deep learning method for the cyber security in IoT network , 2) proposed models are evaluated using CICIDS2017 dataset 3) compare the performance of deep learning algorithm with machine learning algorithms and 4) propose open research challenges for applying deep learning in cyber security field.

Remaining of the paper is organized as follows Section II illustrate literature review, Section III describes proposed deep learning models, Section IV contains experimentation environment and results and in Section V conclusion and future work are outlined.

II. Literature Review

Deep Learning methods have been employed in the fields such as image processing, speech recognition, healthcare etc. which as provided better performance in comparison with other machine learning methods. An deep learning based method for the detection of distributed attack in fog-to-thing computing is proposed in [12]. This work illustrates the drawback of cloud computing got IoT network as it is centralized processing which is not appropriate for large IoT network as it require the processing for cybersecurity at the edge of the network. Deep learning has been proven in the field of big data areas, so for IoT network a fog-to-node method is appropriate for the massive IoT network generating huge data. This work is conducted on NSL-KDD datasets by employing stacked auto encoder along with softmax as classifier and compared with a shallow learning model based on performance metrics such as accuracy, false alarm rate and detection rate. Author also demonstrated the fruitfulness of distributed parallel computing employed on fog to node model as improved accuracy and efficiency of attack detection.

In [13] author has proposed a self-taught deep learning based authoencoder in combination with SVM (Support Vector Machine) for intrusion detection in network. Deep learning is employed as feature selection method in unsupervised manner to reduce training and testing time and also improve the performance by increasing the accuracy of the SVM classifier. Author has compared the proposed method for both binary and multiclass classification along with comparison with other shallow machine learning algorithms such as J48, naïve Bayesian and random forest. Proposed method has provided better results in term of performance such as accuracy in comparison with other proposed methods.

An comparison of shallow and deep neural network has been proposed in [14]. Author has used KDDCup-'99' dataset to train and test proposed method with learning rate of 0.1 and compared the results obtained with other machine learning methods based on performance metrics such as accuracy, precision and recall. In their research author conclude that deep learning are promising technology for the cyber security field and in conducted work deep neural network model with 3 layer performed best in comparison with other models.

An deep learning model based on Bidirectional Long Short Term Memory based Recurrent Neural Network(BLSTM-RNN) for the detection of botnet is proposed and compared with LSTM which is a RNN model in [15]. Author has generated dataset for this work for including four attack vectors as used by mirai botnet. They have tested and validated their proposed method on four attack vector as mirai, udp, dns and ack. Proposed method has shown to be performing well for mirai, udp and dns attack vectors with accuracies as 99%, 98% and 98% respectively but does not perform well for ack attack vector comparatively for which they suggest it could improve with more training data.

An illustrated literature survey and brief tutorial on machine learning and deep learning methods for the cyber security is given in [16]. They have discussed various problems existing in datasets available for IDS training and testing and also the challenges in employing machine learning and deep learning for cybersecurity. Author has raised the problem of training both the methods as the network data update very fast and this lead to the retraining of the models so author has suggested the lifelong training as future work.

III. Deep Learning Models

In this paper we have implemented four different classification deep learning models as MLP (Multilayer Perceptron), 1d-CNN, LSTM, CNN+LSTM. In later section the deep learning performance is also compared with machine learning algorithms. For all the models the last layer is dense with sigmoid activation function as our data has two classes as normal and attack.

A. MLP Deep Learning Model

Fig. 3 shows the flow chart of MLP model we implemented for this work. Input shape for MLP model is 2d data, out dataset is in form of matrix so we don't need to change the shape of our dataset for this model. The proposed model consist of first input layer followed by three dense layers. Output from each layer become input to next layer. One dropout layer is added to save system from heating. Output from the dropout layer is fed to fully connected layer which than provide input for the dense layer with sigmoid function.

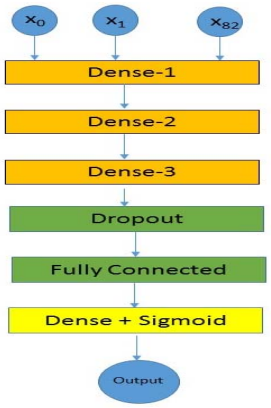

Fig. 3. MLP Deep Leraning Model Architecture

B. CNN Deep Learning Model

Fig. 4 shows the architecture of CNN models, in any CCN model there are three types of main layer as convolutional layer which, pooling layer and dense layer. 1d-CNN accepts input shape of data in the 3d form as (batch, steps, channels) so we converted out data to 3d shape accordingly. Dataset used total has 83 attributes including last label attribute so we converted data using reshape function as {data.shape(0), data.shpe(1), 1} and fed input shape as {82,1} and used relu as activation function. Max pooling layer is added discard features with low score and keep only features with highest score. Last layer is dense with sigmoid activation function.

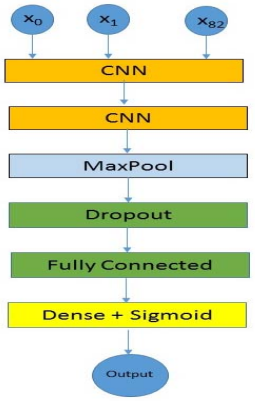

Fig. 4. CNN Deep Leraning Model Architecture

C. LSTM Deep Learning Model

LSTM is a type of RNNs in this nodes are connected to other nodes in same layer improve learning by removing and remembering specific information. The flow graph of LSTM model is presented in Fig. 5, it consist of first LSTM layer with 128 kernel using adam activation function followed by a dropout layer with rate 0.5. Output from dropout layer is connected to a fully connected layer which provides input to a dense layer with sigmoid function to classify attack and normal data.

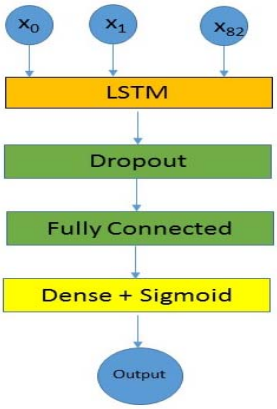

Fig. 5. LSTM Deep Leraning Model Architecture

D. CNN+LSTM Deep Learning model

A hybrid CNN with LSTM model is implemented, Fig. 6 illustrates architecture of this proposed model. This model has first 1-dCNN layer with relu activation function, which is followed by a LSTM layer with adam activation function. Rest of the parameters are same as used in CNN and LSTM models.

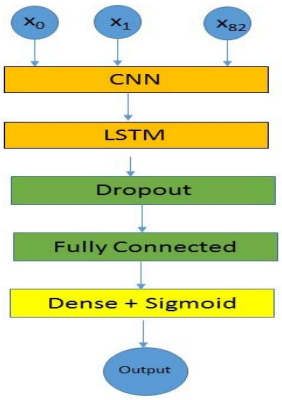

Fig. 6. Hybrid CNN+LSTM Deep Leraning Model Architecture

IV. Experimentation, Result and Discussion

In this section first we have illustrated the dataset used and environment of experiment. Then the metrics used are discussed for the measurement of performance of proposed models and later results are discussed.

A. Dataset and Environment

For conducting proposed work we have used latest DDoS attack CICIDS2017 dataset[17, 18]. Most of DDoS attack datasets have many limitations such as out of relevant data, redundancy which are unreliable. CICIDS2017 datasets contain up to date real work network resembling data. This dataset was collected for 5 consecutive days with many different cyber-attacks along with normal data. This dataset contains most recent up to date network data with and without attack which is very close to the real work network data. This dataset is unbalanced so we have balanced this dataset by duplicating method as it seriously affect the training of the deep learning method and hence the testing. This work is employed using Keras [19] on Tensorflow package for deep learning on 64-bit Intel Core-i7 CPU with 16 GB RAM in Windows 7 environment. Machine learning algorithm as implemented in MATLAB 2017a.

B. Performance Metrics

The performance of proposed deep learning models for the detection of DDoS attack is measured by standard matrices as Accuracy, Recall and Precision. The equation for the same is given below

$$\text{Accuracy: } \frac{TP+TN}{TP+\text{TN}+\text{FP}+FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{\text{TP}+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

where TP, TN, FP, and FN stand for true positives, true negatives, false positives, and false negatives, respectively.

C. Results

The deep learning models are implemented as discussed in section IV. All the models are evaluated on balanced CICIDS2017 dataset with learning rate of 0.1 and maximum number of epoch as 100 after this the models show no improvement in accuracy. Fig. 7 shows the comparison of accuracy value obtained from deep learning models employed along with SVM, bayes and Random forest machine learning algorithms. The parameters and attributes of deep learning models are discussed in section III. Accuracy obtained with 1d-CNN model is 95.14 %, with MLP is 86.34%, with LSTM is 96.24% and with CNN+LSTM is 97.16 %. As it is clear from the figure the highest accuracy we have obtained is with CNN+LSTN model while the lowest is with MLP layer.

Fig. 7 also illustrates the accuracy we obtained by employing machine learning methods on same dataset. The accuracy obtained with SVM is 95.5%, with bayes is 95.19% and with random forest is 94.64%. LSTM and CNN+LSTM perform better than machine learning algorithm while 1d-CNN is almost same but MLP accuracy is much lower by around 9.00% than the machine learning algorithm.

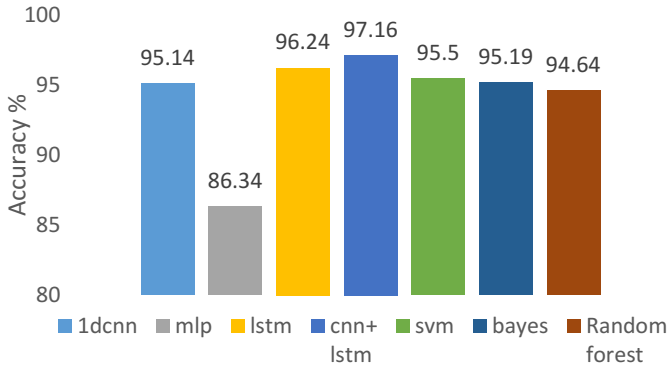

Figure 7 Comparison of accuracy of proposed deep models and machine learning methods

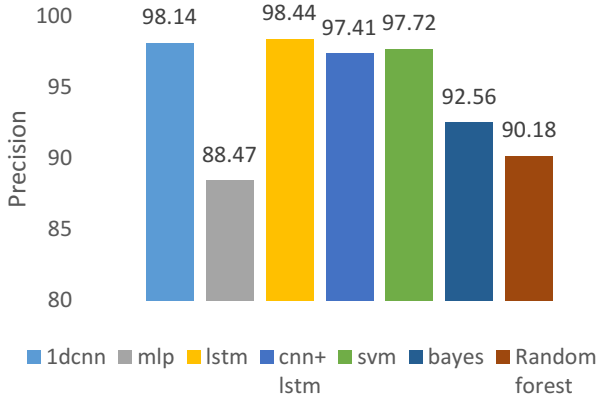

Figure 8 Comparison of Precision of proposed deep models and machine learning methods

Fig. 8 presents the comparison of precision obtained by deep learning model and machine learning methods. The parameter settings are same as those we employed to obtain accuracy value. The precision value with 1d-CNN model is 98.14%, with MLP is 88.47%, with LSTM is 98.44% and with CNN+LSTM model is 97.41%. The highest precision value we have obtained is with LSTM model which outperform the MLP model by around 10.00%. The LSTM model precision is more by 10.00% as compared to that of MLP. It is interesting to see that the precision of hybrid model is lower by 1.03% as compared with LSMT alone. Figure 8 illustrates comparison of deep learning models with machine learning algorithms. It can be seen from the figure that precision obtained with SVM is 97.72%, with bayes is 92.56% and with random forest is 90.18%. The precision obtained with MLP model is lower than machine learning algorithms, while other models precision is better than

machine learning algorithm. The precision of LSMT outperform the SVM by around 1.20%.

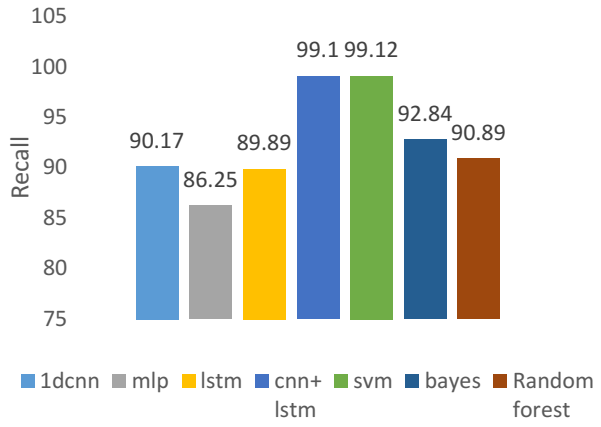

Figure 9 Comparison of Recall of proposed deep models and machine learning methods

The comparison of recall matric is summarized in Fig. 9, it can be seen that the value of recall with 1d-CNN is 90.17%, with MLP is 86.25%, with LSTM is 89.89% and with CNN+LSTM is 99.1%.The precision of CNN+LSMT is higher by at least around 9.20% as compared with other models but as accuracy and recall is higher. It is interesting to see that MLP performance is still lowest in Fig. 9. This can be concluded based on the results obtained that the CNN+LSTM is performing better than other deep learning models and machine learning algorithm for the detection of DDoS attack.

V. Conclusion and Future Work

In this paper we proposed and implemented four different deep learning models and compared them with machine learning algorithms. We found that the hybrid CNN+LSTM model perform better than rest of the deep learning models and machine learning algorithms with accuracy of 97.16%. Also the MLP deep learning model is least performing deep learning model on the dataset used. We have found that except for MLP, the accuracy obtained by other three deep learning methods is more than 95.00% and are performing better than the machine learning algorithms. In future the implementation of IDS based on deep learning classification could be tested for fog to node architecture using distributed parallel processing as explained in the section I. Deep learning based method don't require feature selection to be done before the classification learning and testing but with large number of attributes in the datasets, the training time could be reduced by applying feature selection before training the model. Also for this work we balanced the dataset as it is highly unbalanced by duplicating the data, this could be improved in future by developing deep learning model which could work on unbalanced dataset.

References

[1] L. Coetzee and J. Eksteen, "The Internet of Things-promise for the future? An introduction," in *IST-Africa Conference Proceedings, 2011*, 2011, pp. 1-9.

[2] A. Chadd, "DDoS attacks: past, present and future," *Network Security,* vol. 2018, pp. 13-15, 2018.

[3] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "DDoS in the IoT: Mirai and other botnets," *Computer,* vol. 50, pp. 80-84, 2017.

[4] X. Yuan, C. Li, and X. Li, "DeepDefense: Identifying DDoS Attack via Deep Learning," in *2017 IEEE International Conference on Smart Computing (SMARTCOMP)*, 2017, pp. 1-8.

[5] G. E. Hinton, "Deep belief networks. Scholarpedia, 4 (5), 5947," *Available electronically at http://www. scholarpedia. org/article/Deep_belief_networks Hoppensteadt, FC,* pp. 129-35, 2009.

[6] A. Sharif Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: an astounding baseline for recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2014, pp. 806-813.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.

[8] B. Hu, Z. Lu, H. Li, and Q. Chen, "Convolutional neural network architectures for matching natural language sentences," in *Advances in neural information processing systems*, 2014, pp. 2042-2050.

[9] W. Hao, R. Bie, J. Guo, X. Meng, and S. Wang, "Optimized CNN Based Image Recognition Through Target Region Selection," *Optik-International Journal for Light and Electron Optics,* vol. 156, pp. 772-777, 2018.

[10] T. Hori, Z. Chen, H. Erdogan, J. R. Hershey, J. Le Roux, V. Mitra*, et al.*, "Multi-microphone speech recognition integrating beamforming, robust feature extraction, and advanced DNN/RNN backend," *Computer Speech & Language,* vol. 46, pp. 401-418, 2017.

[11] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE transactions on neural networks and learning systems,* vol. 28, pp. 2222-2232, 2017.

[12] A. Abeshu and N. Chilamkurti, "Deep learning: the frontier for distributed attack detection in Fog-to-Things computing," *IEEE Communications Magazine,* vol. 56, pp. 169-175, 2018.
[13] M. Al-Qatf, M. Alhabib, and K. Al-Sabahi, "Deep Learning Approach Combining Sparse Autoen-coder with SVM for Network Intrusion Detection," *IEEE Access,* 2018.
[14] V. K. Rahul, R. Vinayakumar, K. P. Soman, and P. Poornachandran, "Evaluating Shallow and Deep Neural Networks for Network Intrusion Detection Systems in Cyber Security," in *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2018, pp. 1-6.
[15] C. D. McDermott, F. Majdani, and A. V. Petrovski, "Botnet Detection in the Internet of Things using Deep Learning Approaches," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, pp. 1-8.
[16] Y. Xin, L. Kong, Z. Liu, Y. Chen, Y. Li, H. Zhu*, et al.*, "Machine Learning and Deep Learning Methods for Cybersecurity," *IEEE Access,* 2018.
[17] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization," in *ICISSP*, 2018, pp. 108-116.
[18] R. Vijayanand, D. Devaraj, and B. Kannapiran, "Intrusion detection system for wireless mesh network using multiple support vector machine classifiers with genetic-algorithm-based feature selection," *Computers & Security,* vol. 77, pp. 304-314, 2018.
[19] *Keras deep learning P.W.D. Charles Project Title* Available: https://github.com/charlespwd/project-title