Journal of Science and Technology
ISSN: 2362-1443 (Print) 3059-9490 (Online)
2024, Vol. 4, No. 1, 21-25

Deep Learning Model for Security of IoT Network

Sukhachandra Thakur
Nepal College of Information and Technology,
Pokhara University
thakur2042@gmail.com

Shashidhar Ram Joshi
Institute of Engineering,
Tribhuvan University
srjoshi@ioe.edu.np

Article History:

Received: 24 February 2024
Revised: 16 April 2024
Accepted: 26 May 2024

Keywords— ***IoT, Deep Learning, MLP, 1DCNN, KNN, Naïve Bayes, KDDCUP99, NSLKDD, UNSW-NB15***

***Abstract*— The Internet of Things (IoT) is the emerging and rapidly rising network of physical objects that are provided IP addresses for network connectivity and having the ability of transferring data between objects and other Internet-based devices and systems. There are billions of IoT devices connected and there is a high cyber security and data privacy risk. Computers and mobile devices have many software and security solutions to secure and defend from attacks, but a similar type of security solution is missing to secure IoT networks. In this paper, the One-Dimensional Convolution Neural Network (1DCNN) is proposed to measure efficiency using UNSW-NB15 dataset which is the latest and covered modern attacks data in comparison to NSL KDD and KDDCUP99. For comparison study of performance, we have compared attic Machine learning models with KNN and Naïve Bayes. In each experiment, the model ran up to 200 epochs and with 0.001 learning rate. Deep learning models have outperformed in comparison to the attic machine learning model.**

I. INTRODUCTION

The IoT device is a physical device which is not a standard device like Computer, Laptop and Smartphones but they have capability of transmitting data over the internet. IoT devices communicate with the environment using various sensors and actuators and pass those data to the internet for further processing, monitoring and controlling purposes.

Each IoT device connect with network using unique identifier i.e. Ip address devices include Smart watch, CCTV, Smart electrical meter , Smart home appliance etc. Approximately there are over 20 billion IoT devices connected with the internet and these devices are increasing in a massive way day by day [11]. One research found that in most 10 popular IoT gadgets there are more than 250 vulnerabilities including open telnet port, old firmware, absolute Linux version etc. [5].

In 2016, Oct, there were more than 100K IoT devices, mainly CCTV cameras captured and used for distributed denial of service (DDoS) attacks in Dynamic DNS server. This attack was named as Merai Botnet attack [20]. Because of that attack, the web services like Netflix, Amazon, GitHub, CNN, Twitter etc were down for a couple of hours. As per Gartner analysts 25% of cyberattacks will be done from IoT devices by 2020[4]. Because of limited resources i.e Memory and CPU, complex and efficient security mechanisms can’t be applied.

There are a lot of signature-based IDS implemented in the past days in the IoT network, but they did not perform well and had 1. High False Alarm 2. Not able to detect zero-day attack 3. Not able to deal with advanced types of attack. To handle issues in signature-based IDS for IoT network, researchers invented a solution i.e. use of machine learning algorithm specially use of neural network or deep learning.

In this research, we have used security model at IoT gateway layer also called FoG layer where near real time IoT traffic can be monitored because of comparatively very less traffic in compared to cloud layer IoT security model .To classify attack and normal traffic we have used various deep learning models such as Multi-layer Perceptron (MLP), 1-Dimensional Convolutional Neural Network (1DCNN) Model with single convolution layer , 1DCNN model with two convolution layers and various form of Recurrent Neural Network which are Simple Recurrent Neural Network (RNN) Model, Long short Term Memory Model (LSTM) , Gated Recurrent Unit Model (GRU).

II. RELATED WORK

Several studies revolving around IoT security have attempted to design IDS systems tailored specifically for the IoT ecosystem. In [1], the authors proposed multi-stage, Naïve Bayes Multinomial in stage 1 for unique word classification in dataset and Random Forest classifier at stage 2 for other single-valued quantitative attributes used to classify IoT and non-IoT traffic with the accuracy 99%. They used three major attributes Set of domain names, a set of remote port numbers and a set of cipher suites.

In [4], the authors proposed an IDS model which classifies DDoS attack traffic and normal traffic in an IoT network using various machine learning algorithms. Mentioned that there are two types of features in IoT traffic 1. Stateless features 2. Stateful features. Stateless features are length of packet, Packet interval, communication protocol and state full features are bandwidth, Ip of destination address. More than 90% of attack packets were found under 100 bytes, whereas normal packet size was between 100 and 1,200 bytes.

The CNN based model to enhance security in IoT Security Risk Assessment (SRA) .The main advantage of this network is the ability to learn complex features from a large amount of unlabeled data which is very crucial in IoT SRA [6]. CNN model is composed of different stages where two most important phases are "Extraction" and "Classification". The first phase, "Extraction", is responsible for learning and extracting features automatically from raw data which has two basic layers 1. Convolution layer 2. Max pool layer the proposed model was able to detect DDoS attack with highest accuracy rate in compare to classifying other types of attack i.e. Malware, R2L, Probe and normal traffic.

In [8], researchers implemented deep learning algorithm for Network Intrusion Detection System, used CNN and the combination of convolution and sequential modeling such as RNN, LSTM and GRU. To analyze model performance KDDCup 99 and NSL-KDD data set used. He achieved 99% accuracy using single convolution layer in 1DCNN model whereas adding extra convolution with 2 layers achieved 99.8% and with 3 layer 80.1% respectively.

III. Methodology

A. System Model

In this proposed model, The 1DCNN model to classify normal and benign traffic using UNSW-NB15 intrusion detection data set which includes IoT and non IoT traffic as shown as fig. 1. The data set has a total of 49 features where there are some categorical features which are removed. So after preprocessing we applied 42 input features to the 1DCNN model. The first layer is convolution layer where we have used 32 convolution filters , 5 kernel size and 1 time steps .Use of only one filter in first convolution layer of 1DCNN would make convolution layer to learn only 1 features which would not enough so we used 32 filters to get 32 different features in first convolution layers. Relu has been used as an activation function in 1dCNN and for optimizers we have used Adam . In the second layer i.e Max pool which does down sampling of convolution layer output. In the third layer, Flattening operation is done using the Flattened Layer which is used to make input data ready to load in an artificial neural network. The fourth layer is dropout which is added to provide regularization in CNN model as it is capable of resulting better generalization and has less chance to over-fit in the training process. The flattened output has been applied to Fully Connected Layer and final is the output layer where sigmoid activation function is used to get classified output. The output layer provides result 0 or 1 where 0 is normal traffic and 1 is attack traffic.

B. Dataset and Preprocessing

We have used the UNSW-NB 15 data set available at Kaggle [19] which is used for attack classification in IoT networks by many researchers. It includes 49 features of the dataset and the number of records in the training set is 175,341 records and the testing set is 82,332 records from the different types attack and normal as shown in Tab. 1.And, data set and its description is shown in Tab 2.

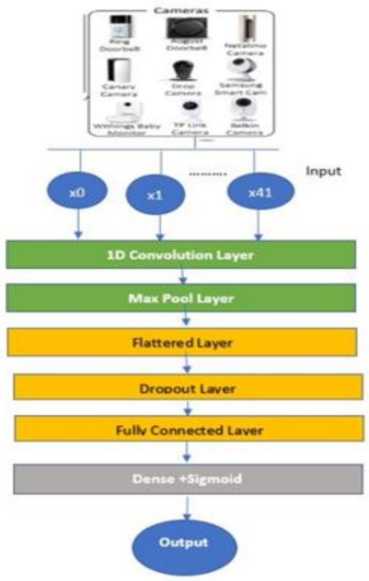
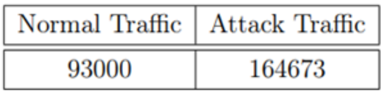


Fig. 1. 1D CNN model

Typically, real-world data is incomplete, inconsistent and inaccurate contains errors or outliers. Data preprocessing is integral steps to maintain high quality of data so that ML can give best result.

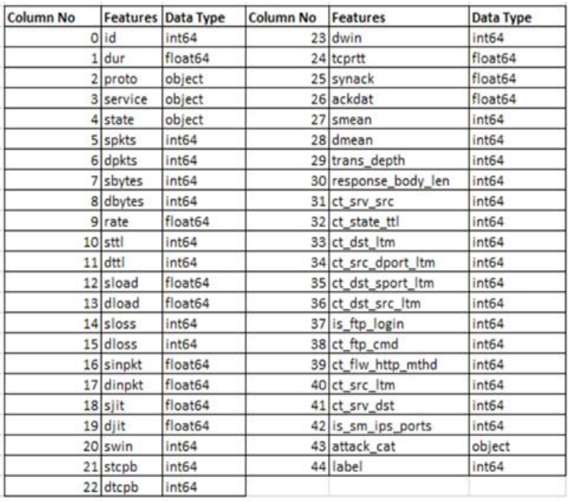
TABLE I. .Distribution of Normal and Attack traffic



Normal Traffic	Attack Traffic
93000	164673

In the given data set, there are 42 features where 6 features are in string value which need to convert into numerical value because the input value of CNN should be a numeric matrix. We converted some non-numeric features, such as 'protocol_type', 'service' and 'flag' features, into numeric form.

TABLE II. Dataset Description



Column No	Features	Data Type	Column No	Features	Data Type
0	id	int64	23	dwin	int64
1	dur	float64	24	tcprtt	float64
2	proto	object	25	synack	float64
3	service	object	26	ackdat	float64
4	state	object	27	smean	int64
5	spkts	int64	28	dmean	int64
6	dpkts	int64	29	trans_depth	int64
7	sbytes	int64	30	response_body_len	int64
8	dbytes	int64	31	ct_srv_src	int64
9	rate	float64	32	ct_state_ttl	int64
10	sttl	int64	33	ct_dst_ltm	int64
11	dttl	int64	34	ct_src_dport_ltm	int64
12	sload	float64	35	ct_dst_sport_ltm	int64
13	dload	float64	36	ct_dst_src_ltm	int64
14	sloss	int64	37	is_ftp_login	int64
15	dloss	int64	38	ct_ftp_cmd	int64
16	sinpkt	float64	39	ct_flw_http_mthd	int64
17	dinpkt	float64	40	ct_src_ltm	int64
18	sjit	float64	41	ct_srv_dst	int64
19	djit	float64	42	is_sm_ips_ports	int64
20	swin	int64	43	attack_cat	object
21	stcpb	int64	44	label	int64
22	dtcpb	int64			

In UNSW-NB 15 data set, there are some missing values and we handled by replacing with mean value of that columns. Also, to ensure no any feature dominates other important features based on their value we scale all features in similar range and to scale we used Minmax scaler function.

C. Implementation

All the deep learning models are evaluated using UNSW NB15 data set. We combined both train and test data available in dataset and randomly splitted 80/20 train test data split . The available train and test data set caused over fitting. In our proposed 1D CNN model, we have used 32 convolution filter of kernel size 5 and 1 stride .In output layer we have used activation function as sigmoid. We used Adam optimizer with 0.001 learning rate .Also tested with different optimizer like

SGD, AD-Grad and RMS pro but performance did not improve .Also tested with different learning rate 0.1,0.01 and 0.001 and found 0.001 gives best performance .

We used a 'Relu' activation function in the convolution layer. Use of one filter in convolution layer allow the 1D-CNN model to learn one single feature from the dataset which is not sufficient as we have 41 features so we used 32 filters which make convolution layer to extract 32 different features from input dataset.

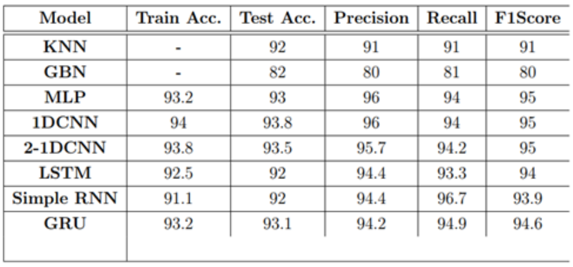
Each filter in convolution layer have their own weight with the defined kernel size, based on length of input matrix. Max pooling layer is used to down sampled output from convolution , a max-filter in polling operation extract the maximum value of the region to where the filter is mapped. It help to minimize the spatial length of the model output, reduce number of features and computational complexity of CNN model. Dropout layer is provides regularization in CNN model by removing unnecessary neurons connections. It also helps model from over fitting. After feature being extracted we flattened the result to feed input to fully-connected layer for attack classification . For comparison of deep learning model with attic machine learning model we compared with KNN(N=15) and Naive Bayes model .Also evaluated data set using MLP , 1D single CNN , 1D two stacked CNN ,simple RNN , LSTM,GRU and CNN+LSTM combination models .We trained our model in 200 epochs also used increased number up to 1000 but but performance did not improve beyond 175 epochs .exponents.

D. Results

We have compared performance using various classic machine learning model as well as different deep learning models. Performance comparison of different models

The Tab. 3 shows the various performance matrix of different machine learning algorithms .We observed Deep learning model given better performance and yields 92-93.8% Accuracy 94-96% , Precision 91.25-96% Recall 93.3-96.7% and F1 Score 93.9-95% in compared to classical machine learning model whose Accuracy 82-92%,Precision 80-91%, Recall 81-91% and F1 Score 80-91%. Among different deep learning model we found 1DCNN with single convolution layer resulted better Accuracy i.e 93.8% in compared to other deep learning models whose accuracy 93.1% , 93%,92% and 92% for GRU , MLP LSTM and Simple RNN respectively .Also we noted that adding of more convolution layer did not improve accuracy of model.

TABLE III. OVERALL PERFORMANCE MATRICES OF DIFFERENT MACHINE LEARNING MODELS.



Model	Train Acc.	Test Acc.	Precision	Recall	F1Score
KNN	-	92	91	91	91
GBN	-	82	80	81	80
MLP	93.2	93	96	94	95
1DCNN	94	93.8	96	94	95
2-1DCNN	93.8	93.5	95.7	94.2	95
LSTM	92.5	92	94.4	93.3	94
Simple RNN	91.1	92	94.4	96.7	93.9
GRU	93.2	93.1	94.2	94.9	94.6

The fig 2 shows the overall performance parameter kept is single graph so that we can visualize which model is best in which performance metrics.

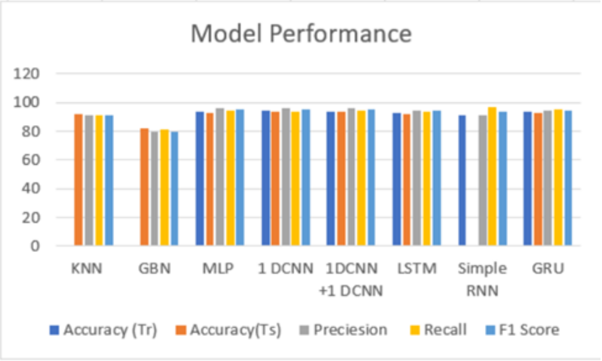


Fig. 2. Overall Performance chart

Loss curve is one of the mostly used graph to monitor a neural network during training. It gives us a snapshot of the training process and the direction in which the network learns. In the proposed model, after 177 epochs, loss function minimized to approx. 0.045 as shown in fig 3. Also we can observed there is smooth curve for training and testing loss which indicate that model have perfectly fit with train and test data.

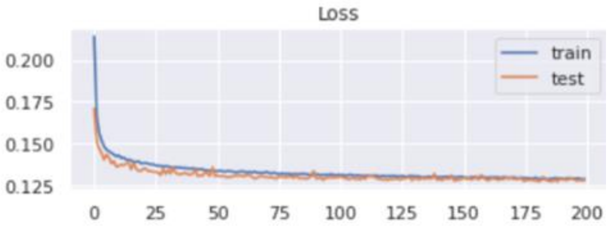


Fig. 3. 1D-CNN Model Epochs vs Loss Curve

The accuracy curve also call learning curve which is plot of model learning performance over training time or epochs .Monitoring learning curves of models during training can be used to identify problems with learning, that are as an over-fit and under-fit. It gives clear idea about whether data set is correctly split into training or testing set or not. The optimal accuracy of the proposed model is obtained at 177 epoch as shown in fig 4.

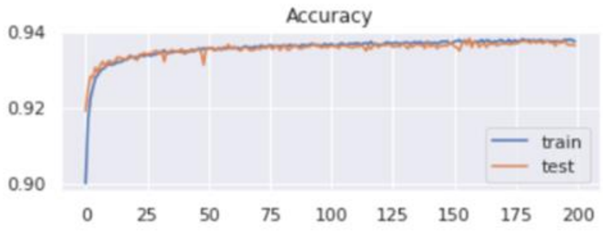
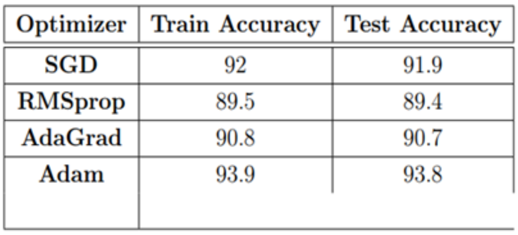


Fig. 4. 1DCNN Epochs vs Accuracy Curve

The process of minimizing loss function or update weight such that gives minimal loss , different optimization techniques used. The basic type of optimizer is gradient descent which is used to update the weights of neural networks such that the model gives minimum value of cost function. In deep learning there are some other advanced optimizers such as RMSprop , AdaGrad and Adam also being used based on requirement. In the experiment, the 1DCNN model is evaluated using different optimizations and found Adam outperformed i.e 93.8% Accuracy in compared to others optimizer which are SGD 91.9% Accuracy, AdaGrad 90.7% Accuracy and RMSprop 89.4% Accuracy as shown in Tab. 4. While varying optimizer , we made learning rate and no of epochs constant i.e 0.001 and 200 respectively.

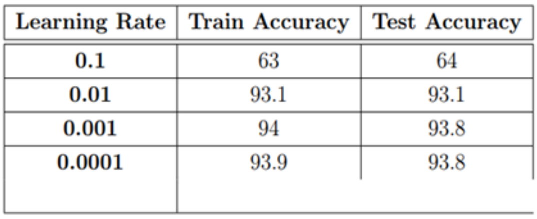
TABLE IV. TRAIN AND TEST ACCURACY OF DIFFERENT OPTIMIZERS



Optimizer	Train Accuracy	Test Accuracy
SGD	92	91.9
RMSprop	89.5	89.4
AdaGrad	90.8	90.7
Adam	93.9	93.8

The learning rate sometime also called step size is a hyperparameter which controls how much to change the 1DCNN model with respect to the estimated error each time when the model weights are updated. Very small learning rate may cause too much time to reach desired minimum loss or some time never reach to destination whereas when we choose it may cause overshooting from minimum point and caused high loss .Similar result we can see in below Tab. 5. where 0.1 learning caused very low accuracy i.e. 63% .We observed that when we chose 0.001 model performed well with accuracy 93.8% in terms of training and testing accuracy .While varying learning rate we made no of epochs 200 and Adam as optimizer constant.

TABLE V. TRAIN AND TEST ACCURACY OF DIFFERENT LEARNING RATE.



Learning Rate	Train Accuracy	Test Accuracy
0.1	63	64
0.01	93.1	93.1
0.001	94	93.8
0.0001	93.9	93.8

IV. CONCLUSION AND FUTURE WORK

Conventional machine learning techniques cannot effectively identify new intrusion and deep learning have the potential to extract superior representations to form way better models. Hence, in this research, we proposed different deep learning models to classify attack and normal traffic in IoT networks. We found that single 1D CNN model outperformed with accuracy 93.8% in compared to other deep learning models 2-1DCNN 93.5% ,MLP 93% ,LSTM 92%,Simple RNN 92% and GRU 93.1% .We also compared deep learning model with traditional machine learning such as KNN and Naive Bayes model where accuracy was not more than 92%.Also we found that adding extra convolution layers in CNN model did not improve the performance of CNN model. Also we have compared the 1DCNN model using different optimizers and different learning rates and identified that Adam optimizer performed best in comparison to other optimizers such as SGD, RMSprop and Adagard..

For future enhancement research around following things can be done – (1) Train and test model using real network by capturing live attack and normal traffic, (2) Use a combination of different deep sequential and convolution learning models such CNN+LSTM, CNN+LSTM+LSTM or CNN+CNN+LSTM etc. to test to see the performance, (3) Design a security model which can detect anomalous activity in a real time network.

REFERENCES

[1] Patel, S., Gupta, A., Nikhil, Kumari, S., Singh, M., Sharma, V. (2018) Network Traffic Classification Analysis Using Machine Learning Algorithms.. 201

[2] Arunan Sivanathan, Hassan Habibi Gharakheili, Franco Loi, Adam Radford, Chamith Wijenayake,Arun Vishwanath and Vijay Sivaraman Classifying IoT Devices in Smart Environments Using Network Traffic Characteristics. IEEE 2018 Transactions on Mobile Computin

[3] Anthi, E., Williams, L., Slowinska, M., Theodorakopoulos, G., Burnap, P.(2019) . A Supervised Intrusion Detection System for Smart Home IoT Devices. IEEE 2019 Internet of Things Journal, 1–1.

[4] Doshi, R., Apthorpe, N., Feamster, N. (2018). Machine Learning DDoS Detection for Consumer Internet of Things Devices. 2018 IEEE Security and Privacy Workshops (SPW

[5] Alrashdi, I., Alqazzaz, A., Aloufi, E., Alharthi, R., Zohdy, M., Ming, H. (2019). AD-IoT: Anomaly Detection of IoT Cyberattacks in Smart City Using Machine Learning. 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC).

[6] Wissam ABBAS, Zineb BAKRAOUY, Amine BAINA, Mostafa BELLAFKIH Classifying IoT security risks using Deep Learning algorithms IEEE 2018 6th International Conference on Wireless Networks and Mobile Communications (WINCOM)

[7] Monika Roopak,Prof. Gui Yun Tian,Prof. Jonathon Chambers Deep Learning Models for Cyber Security in IoT Networks 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)

[8] Vinayakumar, R., Soman, K. P., Poornachandran, P. (2017). Applying convolutional neural network for network intrusion detection. 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)

[9] Basant Subba, Santosh Biswas, and Sushanta Karmakar A neural network based system for intrusion detection and attack classification. In 2016 Twenty Second National Conference on Communication (NCC),pages 1–6. IEEE, 2016.

[10] Moustafa, N., Slay, J. (2015). UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). 2015 Military Communications and Information Systems Conference (MilCIS)

[11] Melody Moh , Robinson Raju. Raju Machine Learning Techniques for Security of Internet of Things (IoT) and Fog Computing Systems. IEEE

2018 International Conference on High Performance Computing Simulation

[12] Lopez-Martin, M., Carro, B., Sanchez-Esguevillas, A., Lloret, J. (2017). Network Traffic Classifier With Convolutional and Recurrent Neural Networks for Internet of Things. IEEE Access, 5, 18042–18050.

[13] Xiao, L., Wan, X., Lu, X., Zhang, Y., Wu, D. (2018). IoT Security Techniques Based on Machine Learning: How Do IoT Devices Use AI to Enhance Security. IEEE Signal Processing Magazine, 35(5), 41–49.

[14] J. Cañedo and A. Skjellum, "Using machine learning to secure IoT systems," 2016 14th Annual Conference on Privacy, Security and Trust (PST), Auckland, New Zealand, 2016, pp. 219-222, doi: 10.1109/PST.2016.7906930.