ISSN 2278-3091
Volume 8, No.4, July – August 2019
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse21842019.pdf
https://doi.org/10.30534/ijatcse/2019/21842019

Effective Intrusion Detection using Deeper Recurrent Neural Networks

Praveen Kumar Kollu[1], R. Satya Prasad[2]
[1]Research Scholar, Department of CSE, Acharya Nagarjuna University, Guntur, India, praveenman@gmail.com
[2]Professor, Department of CSE, Acharya Nagarjuna University, Guntur, India, profrsp@gmail.com

ABSTRACT

Computer networks are susceptible to a variety of security threats. With the ever growing number of devices and people that are connecting to the network, it has become an utmost priority to defend these networks at a large scale. The constant changing nature of the networks and the increase in the type of attacks have made the traditional approaches to intrusion detection obsolete. In this paper, we are proposing a deeper recurrent neural network based approach for intrusion detection in large scale networks. The proposed model uses independent neurons in each layer to construct a deeper recurrent neural network. It helps in faster training and classification time as well as adaptability and scalability to dynamic environments. To evaluate our proposed model, CICIDS 2017 dataset was used to implement and compare against popular deep learning based approaches in network intrusion detection. The experiments have shown promising results that our proposed model can produce improved results over existing approaches.

Key words: Computer security, Information security, Intrusion detection, Recurrent neural networks

1. INTRODUCTION

With the advancements and democratization of computer networks in the past decade, network architectures and topologies have vastly diversified. The non-linear growth of networks like wireless sensor networks and ad-hoc networks have made self-sustaining networks that may or may not be connected to the outside world. This is especially true for IOT networks [1]. Due to the increase in number and the types of attacks, security has become important than ever [2]. Network Intrusion Detection System (NIDS) detects the abnormalities in the network before any malicious activity is happened. It constantly monitors all the traffic in the networks for any unusual, fluctuating and abnormal traffic.

Most of the traditional approaches for NIDS are based on using Machine Learning models [8]. They were a significant improvement over hardware based approaches. Most simple of the machine learning models was to use K-Nearest Neighbor [3]. Other Machine learning approaches include Naïve Bayes [4-5], SVM [6-7] and Random Forest [6]. Although these were able to effectively detect intrusions, there were several shortcomings. One of the main disadvantages of Machine learning approaches is that the features have to be manually engineered. This will be a problem in applications like intrusion detection because humans cannot perceive the subtle differences in the network data. Higher false alarm rates are often caused by these approaches.

To overcome these, most of the research work in intrusion detection is being based on deep learning techniques one way or the other. The automatic feature extraction of the deep learning methods has shown that it can effectively detect the changes in the highly dynamic nature of the network. Recurrent neural networks (RNN) are found to be better at producing accurate results with lower false alarm rates. These approaches are also found to have several shortcomings. RNN architectures such as LSTM and GRU cannot scale well with the data. It is especially a drawback for intrusion detection because the data from the network is growing rapidly along with the size of the network. They also have vanishing gradient problem due to the use of saturated activation functions such as tanh and sigmoid. Inspired from the IndRNN architecture [9], we are proposing a RNN architecture that can have deeper layers without the vanishing gradient problem. It has been tailored for the dynamic and scalable nature of the network data. The independent nature of the neurons can facilitate for larger sequence data processing The results have shown that our proposed model trains faster than the other deep learning approaches even with the same type of parameters.

2. RELATED WORK

Shone et al. [11] proposed a network intrusion detection system by using Non-Symmetric Deep Auto-Encoder approach. Two NDAE are stacked upon each other while random forest is used as final classifier. Tang et al. [12] proposed a IDS for software defined networks(SDN). The IDS model was deployed at the SDN controller. It contains neurons with multiple hidden layers. [13] proposed normal recurrent neural networks with Forward approach responsible for output values. The proposed model gave best results at 80 hidden nodes and learning rate of 0.1. Kasongo et al. [14]

proposed FFDNN based approach from Intrusion detection. FFDNN model has 60 nodes that are spread over three hidden layers. The proposed model worked best at learning rate 0.05 with high accuracy. [15] proposed auto encoder based intrusion detection system. Initially they used principle component analysis(PCA) to reduce the dimensionality and keep only the significant features then the output is given to the auto encoder with support vector machine(SVM) as classifier.

Abusitta et al. [16] proposed cooperative intrusion detection system it is particularly useful for detection under incomplete information. Stacked denoising auto encoders is used. Finally, logistic regression is used for binary classification. Niyaz et al. [17] proposed two stage self-taught learning system for intrusion detection system. It contains a sparse auto encoder connected to normal neural network with softmax regression. NSL KDD cup dataset was used to evaluate their approach

3. METHODOLOGY

3.1 RNN

Recurrent neural networks (RNN) are widely used networks for applications such as language translation and speech recognition etc. due to their ability to process sequence of information. They consist of memory cell which can remember information from the previous computation to compute the next output. Although they can hold the information across time steps, it is often limited to a small number of steps. Carrying of information as a memory to the next computation can be described as

$$I_t = \sigma(W x_t + U I_{t-1}) \quad (1)$$

Where I_t is the state at time t, σ is the activation function, W is the weight matrix, x_t is the input attribute, U is the state to state weight matrix and I_{t-1} is the state at time t-1.

To overcome this problems architectures like LSTM (Long Short Term Memory) and GRU (Gated Recurrent Unit) are developed. But these architectures like RNN suffer from vanishing gradient problem where the signals do not propagate through the layers. Due to this they cannot have a large number of layers as some information may not make it to the output layer. It is partly because the RNN contains saturated activation functions such as tanh. Using RNN and its variants for application like intrusion detection can be difficult due to the number of attributes at play also the volume of those attributes. In this paper, we are proposing a new model that uses recurrent layers with independent neurons and unsaturated activation function. Figure. 1 shows the proposed methodology diagram.

3.2 Deeper RNN

Each state of the RNN at a particular time is based on the significance of the particular attribute and the importance of the previous state as shown in (1). With the independent RNN the equation can be change to use the hadamard product to calculate the importance of the previous state. The state to state weights are taken as a vector. It can be described as

$$I_t = \sigma(W x_t + U \Theta I_{t-1}) \quad (2)$$

Where I_t is the state at time t, σ is the activation function, W is the weight matrix, $x_{(t)}$ is the input attribute, U is the state to state weight matrix, Θ is hadamard product and $I_{(t-1)}$ is the state at time t-1.

Each neuron in the recurrent layer has access to the information of its own previous state at a previous time step whereas neurons in the normal RNN contains information of all the other neurons in the layer. Having independent neurons is analogous to the feed forward networks where the data points are independent for each other. The recurrent layer is shown in the Figure. 2.

Each recurrent unit consists of weights and Batch normalization is performed before giving to the RNN cell. The ReLU activation is then performed along with the Batch normalization to avoid overfitting. The network can contain arbitrary number of these units based on the application.

These layers can be effectively stacked on top of each other to build deeper networks. Unsaturated activation functions such as ReLU and its variants like LeakyReLU and ELU can also be used between the layers. In this paper we have used a stack of recurrent layers with ReLU activation function.

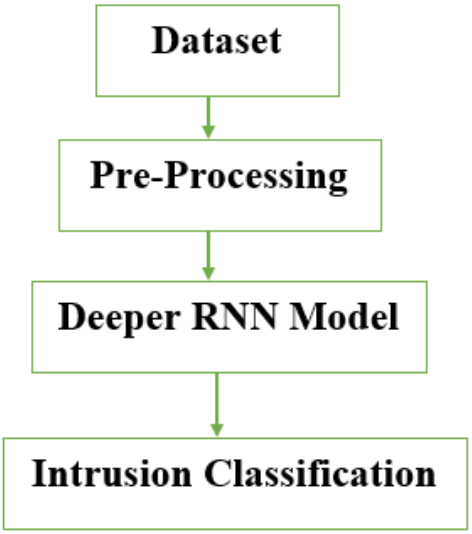

Figure 1: Methodology Diagram

. 3.3 Dataset

CICIDS2017 dataset [10] was used to implement and evaluate our proposed model. This particular dataset is taken because it contains all the necessary criteria for reliable network intrusion detection. It covers a variety of network attacks which are generally not found in other benchmark datasets. They have provided a corresponding CSV files of the entire dataset for use in Machine Learning / Deep Learning applications. The dataset contains data captured over 5 days starting from Monday and ending on Friday. Although it is captured for 5 days each day contains a particular type of attacks except for Monday. Monday captured data contains normal traffic flow. The attacks include DoS, DDoS, Brute Force, XSS, SQL Injection, Infiltration, Port scan and Botnet. The days and their corresponding labels are given in Table I. The dataset contains 78 attributes and a corresponding labels divided by each day.

3.4 Preprocessing

In the pre-processing all the five consecutive dates data files are merged in to a single file which is easy to process further. The rows containing null values are replaced by the mean value of that particular attribute. As the dataset does not contain any categorical values all the attributes are normalized before given to the Deeper RNN model.

Table 1: Days and Labels in the Dataset

Days	Labels
Monday	Benign(Normal)
Tuesday	BForce, SFTP and SSH
Wednesday	DoS and Hearbleed Attacks slowloris, Slowhttptest, Hulk and GoldenEye
Thursday	Web and Infiltration Attacks Web BForce, XSS and Sql Inject. Infiltration Dropbox Download and Cool disk
Friday	DDoS LOIT, Botnet ARES, PortScans

4. EXPERIMENTS

The following metrics and measures are used for evaluating the model.

1. True Positive (TP) – Intrusion data correctly classified
2. False Positive (FP) – Normal data incorrectly classified
3. True Negative (TN) – Normal data correctly classified
4. False Negative (FN) – Intrusion data incorrectly classified

The measures for evaluating and comparing the proposed model are:

Accuracy is the percentage of correct classifications

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

Precision is the ratio of correct classifications to the incorrect classifications

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

Recall measures the ratio of correct classification by missed entries

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

F-Score is the harmonic mean of precision and recall.

$$F - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (6)$$

5. RESULTS

TensorFlow framework was used to implement the model in the cloud using the Google Colaboratory environment. Adam [18] optimizer was used for back propagation as it provides faster inference. As the predicted values are categorical, categorical cross entropy is used as the loss function. The model was trained for 5 epochs with a 20% validation and 5% testing split using random sampling. The proposed model was compared against vanillaRNN which is a normal RNN architecture, LSTM network and GRU network. For better comparison each of the architectures were optimized to produce the best results they can offer.

Initially all the other models other than the proposed models were overfitting. But the proposed model was able to circumvent this because of the Batch Normalization in the network. Dropout was added to all the other model to minimize overfitting. Our proposed model has deeper layers than the rest of the models and minimizes loss faster. Figure 3 shows the Accuracy of the different models. It can be seen that having independent neurons and the ReLU can help attain higher accuracy and faster inference. Figure 4 shows the loss minimization of the models. The normal vanillaRNN has a slow and steady decrease whereas the other models have sudden decrease.

Having faster training and loss minimization will give an advantage on finding intrusions in the network as the data streams are always changing and rapid. Other important measures in intrusion detection are Precision, Recall and F1-Score. The results of these measures for test data are given in the Table 2.

The proposed Deeper RNN model has outperformed all the other model in every measure. As the network is increasing the Deeper RNN model can keep steady measures. It can handle large data sequences than the other models that we have tested. Especially using ReLU has clearly helped the model as the gradient is propagated through the layers which can be seen in Figure 3, the initial accuracy was higher than rest of the models.

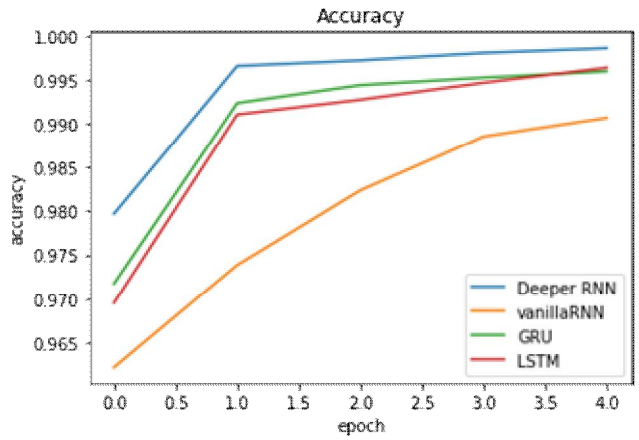

Figure 3: Accuracy of Proposed and Compared Models

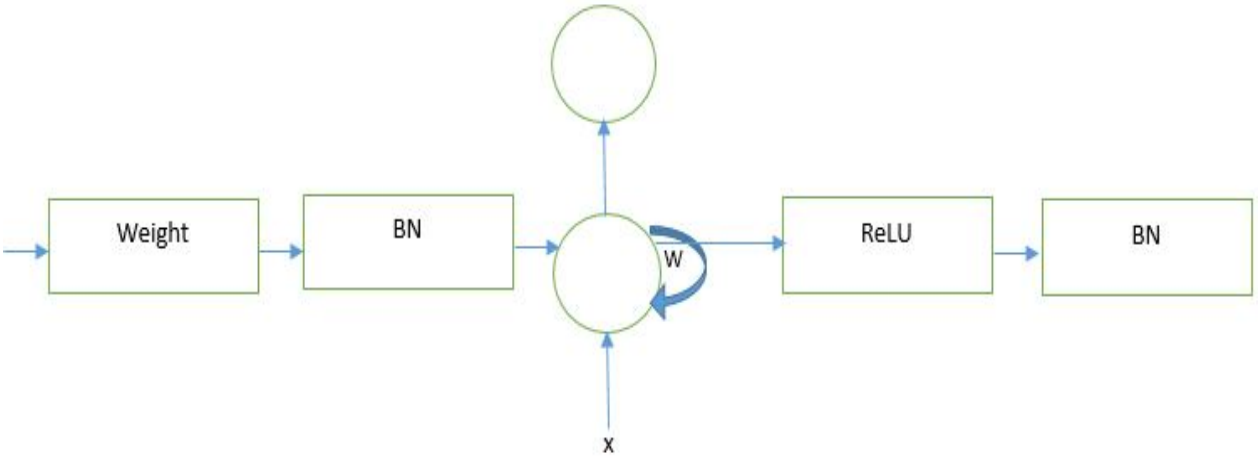

Figure 2: Single Recurrent Unit with Independent neurons

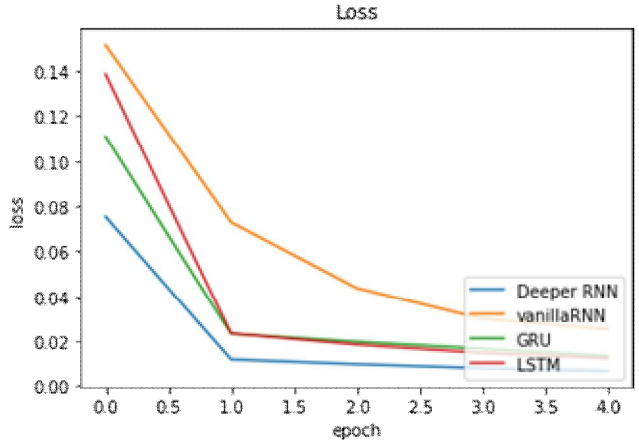

Figure 4: Loss of Proposed and Compared Models

Table 2 Accuracy, Precision, Recall and F1-Score

Model	Accuracy	Precision	Recall	F1-Score
Deeper RNN	**99.86**	**99.95**	**99.91**	**99.93**
vanillaRNN	99.07	99.36	99.68	99.52
GRU	99.60	99.85	99.74	99.79
LSTM	99.64	99.85	99.78	99.82

5. CONCLUSION

In this paper, we have proposed a new approach for intrusion detection system that can adapt to the dynamic nature of the network. The proposed model has deeper RNN layers which is suitable for large scale networks and facilitates information propagation. Gradient propagation to the deeper layers is achieved by the Batch normalization and the ReLU layers. The model trains faster because of the independence between the neurons in the layer. The proposed model can classify network data stream effectively than the other traditional state-of-the-art models for intrusion detection.

.

REFERENCES

1. P.P. Ray, “A survey on Internet of Things architectures”, Journal of King Saud University – Computer and Information Sciences, Volume 30, Issue 3, 2018, pp 291-319. https://doi.org/10.1016/j.jksuci.2016.10.003
2. S. Latha and S. J. Prakash, "A survey on network attacks and Intrusion detection systems," 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, 2017, pp. 1-7. doi: 10.1109/ICACCS.2017.8014614
3. M. Govindarajan and R. Chandrasekaran, "Intrusion detection using k-Nearest Neighbor," 2009 First International Conference on Advanced Computing, Chennai, 2009, pp. 13-20. https://doi.org/10.1109/ICADVC.2009.5377998
4. Mukherjee, Saurabh & Sharma, Neelam. (2012). Intrusion Detection using Naive Bayes Classifier with Feature Reduction. Procedia Technology. 4. 119–128. 10.1016/j.protcy.2012.05.017.
5. F. Gumus, C. O. Sakar, Z. Erdem and O. Kursun, "Online Naive Bayes classification for network intrusion detection," 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining

(ASONAM 2014), Beijing, 2014, pp. 670-674. doi: 10.1109/ASONAM.2014.6921657

6. Farnaaz, Nabila & Akhil, Jabbar. (2016). Random Forest Modeling for Network Intrusion Detection System. Procedia Computer Science. 89.213-217.10.1016/j.procs.2016.06.047.
7. Y. Chang, W. Li and Z. Yang, "Network Intrusion Detection Based on Random Forest and Support Vector Machine," 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), Guangzhou, 2017, pp. 635-638. doi: 10.1109/CSE-EUC.2017.118
8. B. Dong and X. Wang, "Comparison deep learning method to traditional methods using for network intrusion detection," in Proc. 8th IEEE Int. Conf. Commun. Softw. Netw., Beijing, China, Jun. 2016, pp. 581–585.
9. S. Li, W. Li, C. Cook, C. Zhu and Y. Gao, "Independently Recurrent Neural Network (IndRNN): Building A Longer and Deeper RNN," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018, pp. 5457-5466. doi: 10.1109/CVPR.2018.00572
10. Iman Sharafaldin, Arash Habibi Lashkari, and Ali A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization", 4th International Conference on Information Systems Security and Privacy (ICISSP), Portugal, January 2018
11. N. Shone, T. N. Ngoc, V. D. Phai and Q. Shi, "A Deep Learning Approach to Network Intrusion Detection," in IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 2, no. 1, pp. 41-50, Feb.2018.doi:10.1109/TETCI.2017.2772792
12. T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi and M. Ghogho, "Deep learning approach for Network Intrusion Detection in Software Defined Networking," 2016 International Conference on Wireless Networks and Mobile Communications (WINCOM), Fez, 2016, pp. 258-263.doi:10.1109/WINCOM.2016.7777224
13. C. Yin, Y. Zhu, J. Fei and X. He, "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks," in IEEE Access, vol. 5, pp. 21954-21961, 2017.doi: 10.1109/ACCESS.2017.2762418
14. S. M. Kasongo and Y. Sun, "A Deep Learning Method With Filter Based Feature Engineering for Wireless Intrusion Detection System," in IEEE Access, vol. 7, pp. 38597-38607, 2019.
15. Dawoud A., Shahristani S., Raun C. (2019) Dimensionality Reduction for Network Anomalies Detection: A Deep Learning Approach. In: Barolli L., Takizawa M., Xhafa F., Enokido T. (eds) Web, Artificial Intelligence and Network Applications. WAINA 2019. Advances in Intelligent Systems and Computing, vol 927. Springer
16. Adel Abusitta, Martine Bellaiche, Michel Dagenais, Talal Halabi, "A deep learning approach for proactive multi-cloud cooperative intrusion detection system", Future Generation Computer Systems, Volume 98, 2019, Pages 308-318, ISSN 0167-739X.
17. Javaid, Ahmad & Niyaz, Quamar & Sun, Weqing & Alam, Mansoor. (2015). A Deep Learning Approach for Network Intrusion Detection System. EAI Endorsed Transactions on Security and Safety. 3. 10.4108/eai.3-12-2015.2262516.
18. Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.