ISSN 2347 - 3983

Volume 7, No. 8 August 2019

International Journal of Emerging Trends in Engineering Research

Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter12782019.pdf

https://doi.org/10.30534/ijeter/2019/12782019

Intrusion Detection System Using Recurrent Neural Networks and Attention Mechanism

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

Protecting sensitive information over the internet requires specialized algorithms that can detect subtle abnormalities in the data. These abnormalities are harder to detect as most of the data contains redundant information that does not contribute to classification of normal and abnormal classes. Finding the most optimal features is still a difficult task. To optimize the model to focus on the most relevant features and to make the model train considerably faster, we are proposing an attention based recurrent neural network for intrusion detection in large scale networks. The model is optimized to leverage the power of GRU network while focusing on the most important features for classification. The latter is achieved by the attention vector that is introduced to the network. The efficiency of the network is determined using several measures such as accuracy, precision, recall and F1-score. These measures are compared among state-of-the-art classification algorithms to determine how our proposed model performs on par with the current approaches.

Key words : Attention, Feature Selection, Intrusion detection, Neural networks.

1. INTRODUCTION

Intrusion detection system (IDS) is a service that is quintessential to the overall security of the network. It monitors the system events in order to detect the unauthorized access and mitigate resource misuse [1]. An IDS system captures the traffic in real-time to detect these abnormalities in the data. These logs are achieved to audit the reliability of the network. Since the inception of internet and internet based services the need for robust IDS has been ever growing. Today network security management has become the most sought out skill in all the disciplines. There has been a drastic change in the approaches for Intrusion detection. Intrusion detection systems are broadly divided into two approaches, supervised and unsupervised. Initial supervised models used data mining techniques to evaluate the behavior of intrusions and normal activity [2]. With the wide adoption of machine learning algorithms have paved way for new algorithms intrusion detection that can learn from data using various statistical analysis techniques. Widely adopted techniques include Support Vector Machines (SVM) [3], Decision trees and Random Forests [4], and neural networks [5].

Although these models were able to find the network intrusions effectively. Most of these algorithms often require manual feature extraction by a domain expert. Even though the neural network architectures can learn to find the important features, the redundancy in the data decreases the efficiency of the network. This also makes the architecture to focus on the set of features that are not optimal to generalize well for unseen data. Recent literature is mostly focused on using recurrent neural networks for tackling the intrusion detection problem [6]. Various variants of recurrent neural networks such as LSTM and GRU networks made the model to train faster and achieve higher results. Although they can produce better results, they still suffer from the disadvantages specified above.

In this paper, we are proposing a GRU based recurrent neural network that leverages attention mechanism. As the GRU network can train relatively faster and give better results, the attention mechanism helps the network to focus on the optimal set of features. This helps the network to generalize well for unseen traffic from the network. The proposed network is evaluated using two benchmark datasets namely NSL-KDD [7] and CICIDS [8]. These datasets are widely used in experimenting intrusion detection models across major literature. So experimenting using these datasets can give an accurate baseline for the system. Performance of these models are measured using Accuracy, Precision, Recall and F1-Score. All these performance measures are compared against various machine learning models.

2. LITERATURE REVIEW

Farnazz et al. [9] proposed intrusion detection system that can identify the actions of the users on the internet that can notify the abnormal activities of the users over internet. They tested their accuracy at different levels to identify the flaws in network. Random forest classifier was used for identifying the flaws in the network and they are saved for future and overfitting problem canovercome. KDD Dataset was used. The results concluded that the model works efficiently with low false alarm rate when compared to other models. Ambusaidi et al. [10] proposed a mutual information method. This method will select the optimal features from the data in analytical process. classification is performed on that data that is derivedanalytically.Linearity and nonlinearity are the two basic problems that we face in data these two problems can be overcome by using this approach. This approach was performed on three datasets KDD Cup 99, NSL-KDD and Kyoto datasets and the feature selection method was used that contributed the major features responsible for classification.

Niyaz et al. [11] proposed an architecture that can identify the unknown and unpredictable and unforeseen features that are responsible for causing flaws in the system.NSL-KDD dataset was used. Aself-thought deep learning model was used for flexible identification of the unknown or unforeseen data that is responsible for the flaws in the system. Kevric et al. [12] introduced a tree based algorithm for intrusion detection.KDDCUP 99 dataset was used. This is one of the benchmark datasets that is for intrusion detection. In this model 41 features are used for describing hidden patterns present in the data. This model is used for identifying the incoming data traffic into the network and the hidden relationship among them is identified.

Aljawarneh et al. [13]proposed a model that can effectively gather the sensitive information from the data.Meta heuristic data is taken in this model and assessments are performed on the meta data that are used for identifying the possibility of the intrusion in the data. NSL KDD which is one of the famous dataset for intrusion detection was used. This is a hybrid model that can also identify the degrees up to which the intrusion has occurred in the network.

3. METHODOLOGY

The proposed methodology diagram is given in Figure 1. The proposed methodology contains taking either of the two intrusion detection benchmark datasets and pre-processing it. The pre-processed data is then given to the proposed model for training and once the model is trained the weights are saved for future use. The test data is then evaluated on the trained model for intrusion classification

3.1 Datasets

Two benchmark datasets were used to evaluate the proposed model namely NSL-KDD [7] and CICIDS [8]. These two datasets have been used in intrusion detection research since past decade. Details about these datasets are given below.

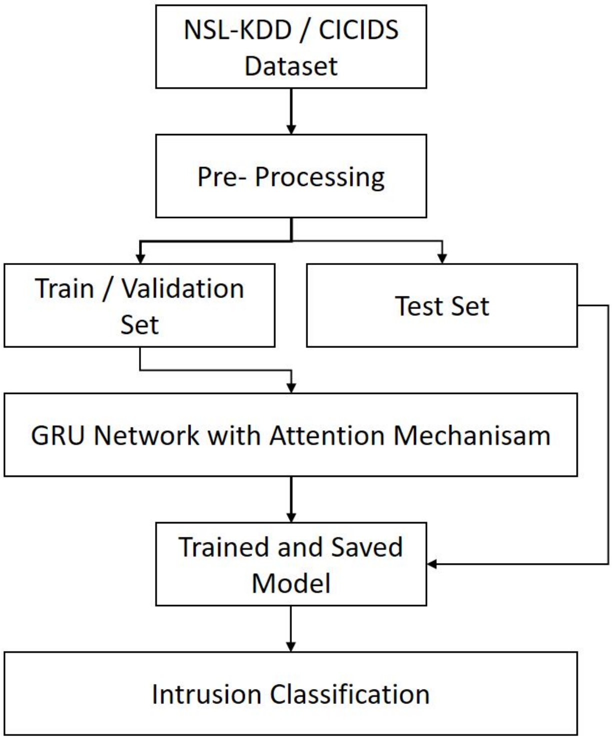


Figure 1: Proposed Methodology Diagram

3.1.1 NSL-KDD Dataset

NSL-KDD is an improved version of the famous KDD cup 99 dataset. Initial dataset had a lot of redundant information in both training and testing set. It was captured and prepared by DARPA in 1998. NSL-KDD dataset eliminated the redundant data from the original KDD cup 99 data and partitioned the records in the dataset into different difficult levels. The NSL-KDD dataset contains around 41 features and the class labels are labelled either normal or the specific type of attack like DOS probing, user-to-root, and root-to-local. The training set of the dataset has class labels of around 23 classes. One being the normal and rest of them are attacks on the network.

3.1.2 CICIDS Dataset

CICIDS 2017 is a relatively new dataset that is being used as benchmark for intrusion detection systems. One of the reasons for choosing this dataset is that it contains data for most recent and common attacks than any other dataset. Total of 5 days' data is captured for this dataset. Although it contains data from 5 days, Monday contains normal traffic without any attacks. Each day is given as a csv file with the corresponding labels. For training the model all the csv files are merged to a single file. Some of the attacks in the dataset include DOS, DDOS, Brute Force, XSS and Botnet etc.

3.2 Pre-Processing

NSL-KDD contains attributes with non-numerical data which has to be converted into numerical data using label encoder and one-hot encoder. CICIDS dataset does not contain any non-numerical attributes. Both the datasets are then processed to fill the null values with the mean of that particular attribute. Finally, both datasets are normalized using the mean and standard deviation. NSL-KDD dataset is used for binary classification where the class labels are converted into normal and attack irrespective of the attack type, whereas CICIDS is used for multi label classification for each specific attack type. This is done to evaluate the models performance on both binary and multi label classification.

3.3 GRU with Attention Mechanism

Gated Recurrent Unit (GRU) is a variant of recurrent neural networks. GRU is relatively faster than other variants such as Vanilla RNN and LSTM networks. GRU is very similar to LSTM in functionality but the main difference is in the gates in the cell. Each LSTM contains three gates input, forget and output respectively where as GRU contains only the reset and update gate. Each of these gates are controlled by activation functions. Only the Addition and Hadamard product operations are performed inside each cell which makes them more computationally efficient. The architecture of a single GRU cell is given in Figure 2. From the Figure 2 it can be seen that the current GRU cell takes input from previous cell $h_{t-1}$to determine the current cell content $h_{t.}$ $x_t$is the feature that is being considered by the unit at time step t. Finally, the update gate operation is performed by $z_{t.}$

Attention layer is incorporated in to the network after the GRU layer. Attention layer is a vector which is multiplied to the output from the GRU layer. Each value in the vector is a probability for each cell in the GRU layer. If the cell is important it contains high probability value by the softmax function. The full architecture is given in Figure 3. As the attention vector is fully derivable. It can learn to focus on the most significant features for the GRU layer. This can help the network to achieve high performance and can also generalize well. The last layer is a normal Dense layer with a single neuron and sigmoid function for binary classification and Dense layer with multiple neurons and softmax function for multi-label classification.

4. EXPERIMENTS

The proposed model was implemented using the keras framework with tensor flow backend. Adam optimizer with an initial learning rate of 0.001 was used to optimize the training. Categorical cross entropy and binary cross entropy were used as loss functions for multi-label classification and binary classification respectively. The CICIDS dataset achieved inference in around 15 epochs and NSL-KDD dataset in 10 epochs. The following evaluation metrics were used to measure the performance of the model.

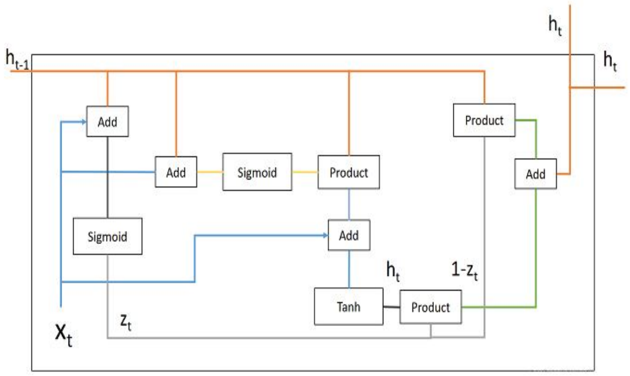


Figure 2: GRU Cell

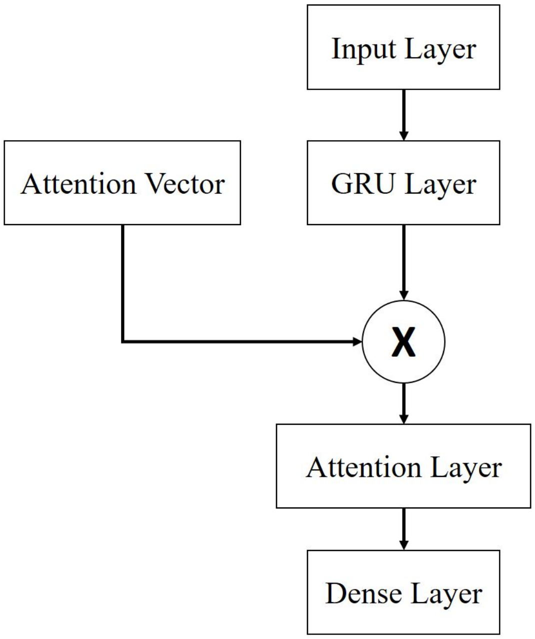


Figure 3: Proposed Network Architecture

Accuracy is the percentage of correct classifications

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

Precision is the ratio of correct classifications to the incorrect classifications

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

Recall measures the ratio of correct classification by missed entries

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

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

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \qquad (4)$$

4. RESULTS

The results were given in tables 1 and tables 2. It can be seen that the proposed model was able to achieve higher scores across all the evaluation metrics. The attention mechanism in the proposed model has helped optimizing the loss of the network. Consistent results were obtained from both datasets. The accuracies of training and testing for NSL-KDD are given in Figure 4 and for CICIDS in Figure 6. Similarly, the loss is given in Figure 5 and Figure 7 respectively. The model optimized the loss to a great extent in just a couple of epochs and has shown consistency in minimizing. Based on the results it can be seen that the proposed model is able to work on Intrusion Detection System irrespective of the dataset and also be able to utilize for binary as well as multi-label classification.

Table 1: Results of NSL KDD Dataset

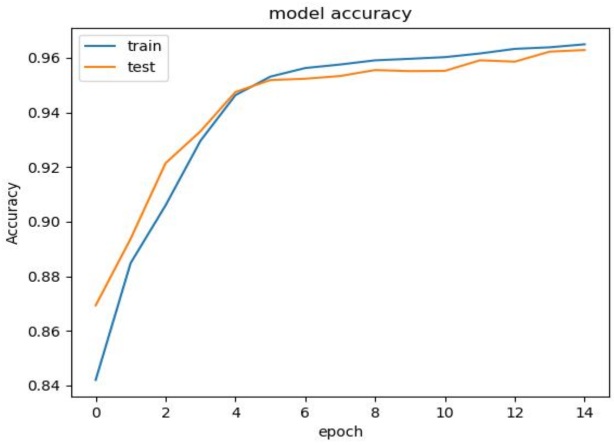
	Accuracy	Precision	Recall	F1-Score
simpleRNN	0.94	0.94	0.95	0.94
LSTM	0.97	0.98	0.97	0.97
GRU	0.97	0.98	0.98	0.98
Proposed	0.99	0.98	0.99	0.98

Table 2*:* Results of Evaluation Measures for CICIDS dataset

	Accuracy	Precision	Recall	F1-Score
simpleRNN	0.92	0.97	0.96	0.97
LSTM	0.95	0.98	0.95	0.96
GRU	0.98	0.98	0.96	0.97
Proposed	0.99	0.99	0.98	0.98

5. CONCLUSION

In this paper, a new architecture for intrusion detection is proposed. The architecture contains the attention mechanism that can focus on the most optimal features for detecting subtle changes in the network traffic. The model was implemented on two most widely used benchmark datasets and the results have shown that our proposed model performs significantly better in terms of accuracy, Precision, Recall and F1-score. Although we used only two datasets the model can further be used for any dataset and it can extract the significant features for that particular data. This can significantly increase the efficiency of the detections in the network. The GRU is computationally efficient and trains relatively faster. This research can help in more robust intrusion detection systems that can tackle the day to day changes happening in the network attacks.



. **Figure 4:** Training and Testing Accuracy of NSL-KDD dataset

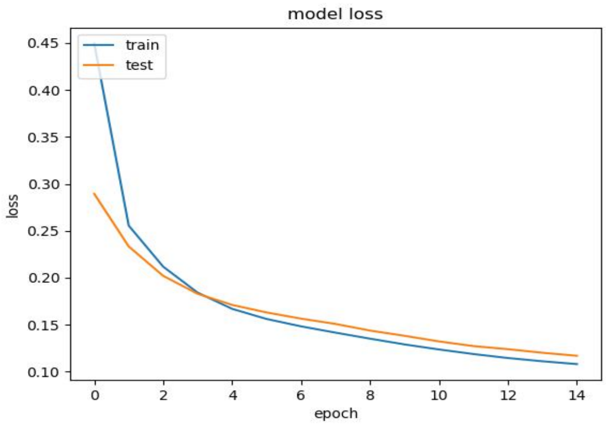


Figure 5: Training and Testing loss of NSL-KDD dataset

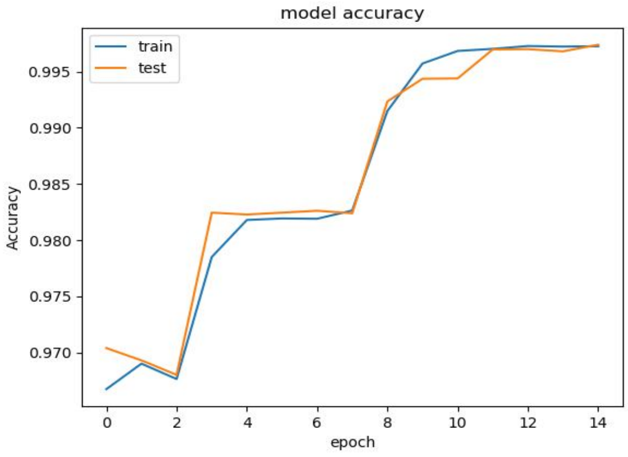


Figure 6: Training and Testing Accuracy of CICIDS dataset

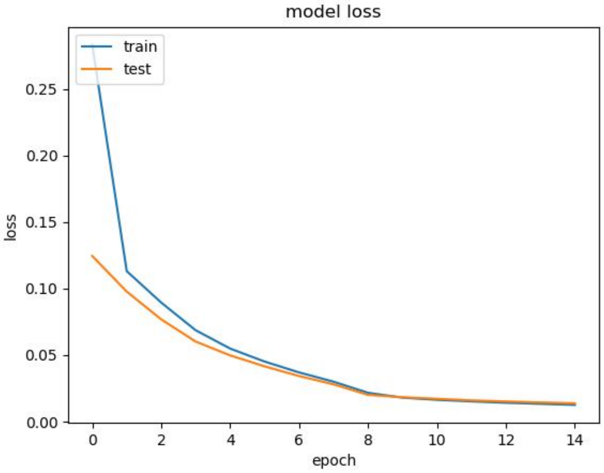


Figure 7: Training and Testing loss of CICIDS dataset

REFERENCES

1. Stallings W, Brown L (2008) **Computer security principals and practice**. Pearson Education, Upper Saddle River.
2. Wenke Lee, S. J. Stolfo and K. W. Mok, "**A data mining framework for building intrusion detection models**," Proceedings of the 1999 IEEE Symposium on Security and Privacy (Cat. No.99CB36344), Oakland, CA, USA, 1999, pp. 120-132.
3. S. Mukkamala , A. Sung ,**Detecting denial of service attacks using support vector machines**, in: Proceedings of the Twelfth IEEE International Conference on Fuzzy Systems, 2003.
4. 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. https://doi.org/10.1109/CSE-EUC.2017.118
5. G. Karatas and O. K. Sahingoz, "**Neural network based intrusion detection systems with different training functions,**" 2018 6th International Symposium on Digital Forensic and Security (ISDFS), Antalya, 2018, pp. 1-6. https://doi.org/10.1109/ISDFS.2018.8355327
6. 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. https://doi.org/10.1109/ACCESS.2017.2762418
7. M. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, "**A Detailed Analysis of the KDD CUP 99 Data Set,**" Submitted to Second IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA), 2009. https://doi.org/10.1109/CISDA.2009.5356528
8. 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. https://doi.org/10.5220/0006639801080116
9. Nabila Farnaaz, M.A. Jabbar, "**Random Forest Modeling for Network Intrusion Detection System**",Procedia Computer Science,Volume 89,2016, Pages 213-217.
10. M. A. Ambusaidi, X. He, P. Nanda and Z. Tan, "**Building an Intrusion Detection System Using a Filter-Based Feature Selection Algorithm**," in IEEE Transactions on Computers, vol. 65, no. 10, pp. 2986-2998, 1 Oct. 2016.
11. A. Javaid, Q. Niyaz, W. Sun, and M. Alam, ''**A deep learning approach fornetwork intrusion detection system**,'' presented at the 9th EAI Int. Conf.Bio-inspired Inf. Commun. Technol. (BIONETICS), New York, NY, USA,May 2016, pp. 21–26.
12. Kevric, J., Jukic, S. &Subasi, A."**An effective combining classifier approach using tree algorithms for network intrusion detection**" Neural Comput&Applic (2017) 28(Suppl 1): 1051. https://doi.org/10.1007/s00521-016-2418-1
13. Shadi Aljawarneh, Monther Aldwairi, Muneer Bani Yassein, "**Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model**",Journal of Computational Science, Volume 25, 2018, Pages 152-160. https://doi.org/10.1016/j.jocs.2017.03.006