International Organization of Research &Development (IORD)
ISSN: 2348-0831
Vol. **08** Issue **01** | **2020**

Review on Deep Learning Method for Intrusion Detection System using Recurrent Neural Networks

Dr. Yogesh Kumar Sharma
Associate Professor, Head/Research Coordinator
Department of Computer Science and Engineering
Shri. Jagdishprasad Jhabarmal Tibrewala University, Vidyanagari, Jhunjhunu, Rajasthan, India.

Mr. Harshal Ashokrao Karande
Ph.D. Research Scholar
Department of Computer Science and Engineering
Shri. Jagdishprasad Jhabarmal Tibrewala University, Vidyanagari, Jhunjhunu, Rajasthan, India.

ABSTRACT

Safety is a critical concern in the field of the information system, because of the enormous amount of internet traffic. The way individuals live, work and study is changing due to the increasing use of the Internet, which is causing more and more serious security threats in everyday life. This paper discusses how to model a deep learning-based intrusion detection system and proposes a deep learning approach to intrusion detection using recurrent neural networks (RNN-IDS). here analyze the model's performance in binary classification and multiclass classification, and the number of neurons and different learning rate impacts on the model's output. They equate it with J48's, artificial neural network, random forest, help vector machine, and other machine learning approaches proposed on the benchmark data set by past researchers. The RNN-IDS model increases intrusion detection performance and provides a new method of examining intrusion detection

Keywords—Intrusion detection, Recurrent Neural Network, Deep Learning Introduction

1. INTRODUCTION

Intrusion detection system is software for detecting intrusion or malicious activities occurring in a particular network. Computer system goes on a high risk, when it is connected to a network. With the increasingly comprehensive integration of the Internet and social life, the Internet is transforming how people learn and function, but it also exposes us to ever more serious threats to health. How to recognize various attacks on the network, particularly attacks not seen before, is a key issue that needs to be resolved urgently.

Intrusion Detection Systems (IDS) focus on identifying possible incidents or threats, logging information, attempting to stop intrusion or malicious activities, and report it to the management station. In addition, it record information related to observed actions, notify security administrators of significantly observed actions and generate reports.

With the increasing integration of the Internet and society, the Internet is changing people's way of living, studying and working, but the various security threats we face are becoming ever more serious. Detection of intrusion is typically equivalent to a classification problem, such as a binary or multi-classification problem, i.e. determining whether network traffic behavior is normal or anomalous, or a classification problem of five categories; I.e. determining whether it is natural or any of the other four types of attack: Denial of Service (DOS), User to Root (U2R), Probe (Test) and Root to Local (R2L). In short, the main motivation for detecting intrusion is to improve the accuracy of classifiers in effectively recognizing disruptive behaviour.

Machine learning methodologies have been commonly used to detect different types of attacks, and an approach to machine learning may help the network administrator take appropriate measures to avoid intrusion. Deep learners, on the other hand, have the ability to derive better representations from the data and create much better models. As a result, technology for intrusion detection experienced rapid evolution after dropping into a relatively slow era. Due to growing computing resources, recurrent neural networks that have existed for decades but have only recently begun to be widely recognized, such as convolution neural networks have recently created significant development in the field of deep learning. Since deep learning has the ability to

extract better representations from the data to create even better models, and motivated by recurring neural networks, we have proposed a deep learning approach for an intrusion detection system using recurring neural networks (RNN-IDS). The main contributions of this paper are summarized as follows.

Authors present the detection system design and implementation, based on recurrent neural networks. In addition, we are analysing the model's success in binary classification and multi-classification, and the number of neurons and different learning rate impacts on the accuracy.

In contrast, on the benchmark NSL-KDD dataset, they analyse the performance of the naive bayesian, random forest, multi-layer perceptron, support vector machine, and other methods of machine learning.
They equate RNN-IDS output with other machine learning methods in both binary and multiclass classification.

2. REVIEW OF LITURATURE

Deep learning, a branch of machine learning, has become increasingly popular in recent years and has been applied for intrusion detection; studies have shown that deep learning overcomes traditional methods altogether.

According to A. Javaid et.Al (2016), a deep learning approach based on a deep neural network for flow based anomaly detection, and the experimental results show that deep learning can be applied for anomaly detection in software defined networks.

According to Harshal Karande, Shyam Gupta (2015), "Ontology based intrusion detection system for web application security", model derived from was tested for various vulnerabilities and attacks listed out in OW ASP. Also successfully showed that the model performance for correctness, response time, and completeness are all at par and beyond with the state of the art. Time efficiency and reduction in space requirement are added benefits.

According to T. A. Tang et.Al(2016), a deep learning based approach using self-taught learning (STL) on the benchmark NSL-KDD dataset in a network intrusion detection system. When comparing its performance with those observed in previous studies, the method is shown to be more effective. However, this category of references focuses on the feature reduction ability of the deep learning. It mainly uses deep learning methods for pre-training, and it performs classification through the traditional supervision model. It is not common to apply the deep learning method to perform classification directly, and there is a lack of study of the performance in multiclass classification.

Dr.Yogesh Kumar Sharma and Rokade Monika D (2019) proposed a Deep Learning base intrusion detection system for synthetic as well as real time network environment. Various dataset have been used to evaluate the proposed experimental analysis. The partial implementation of system shows the better results than existing systems. CIDDS-001, KKDCUP99, NSLKDD, ISCX network dataset used for evaluate the system with different algorithms. The proposed system describe a common drawback that affects machine learning, much, is that the unbalanced category distribution drawback as a results of disproportionate categories. This study was taken off to style associate degree economical anomaly based intrusion detection system from the unbalanced network intrusion dataset and to check other ways of treating original unbalanced category distributions. The experiments make sure that RNN along side down-sampling and sophistication balancer cause effective and comparable ends up in terms of accuracy.

Dr. Yogesh Kumar Sharma and S Pradeep (2019), focuses on the integration of deep learning approach with real time feature descriptor from camera of flying FWN aircraft so that the definite target can be identified accurately. Proposed approach is presented to be effective in military as well as civil defense to recognize the activities of suspicious persons or even to locate the flying objects released by opponent country. In this approach, the integration of TensorFlow, Keras and PyTorch with feature descriptors in the camera of Flying Deep Learning Integrated Drone (FDLID)is proposed and found that the strategic and tactical decisions can be made using this methodology. Proposed approach of deep learning the evaluation of feature points can be done with the training data and for further appropriate actions. Deep neural networks are widely used in other engineering applications including optimization, soft computing, and biological computations and even in pure sciences. This approach is presented in the flying IoT based networks as it is an effectual and high performance paradigm and proven to be powerful in different domains.

According to Manoj s. Koli proposed An Advanced method for detection of botnet traffic using Internal Intrusion Detection, confirmed that on the advanced method for detection to improve the security by identifying and tracking the attacker using machine learning, ranking and Voronoi clustering is proposed the paper ensure reducing the size of data set and high detection accuracy. A data set called ISOT has been used keeping in mind the processing delay in the large scale network UDP and TCP are examined to recognize achieve instruction growth in network traffic is taken care of machine learning modules act like deep neural network various botnet techniques are provided DNA based method is developed by the system help. The paper also uses characteristics of the network flow to detect the botnet intrusion despite packet payload content, which helps in encryption of packet.

Deep learning approaches have blossomed exponentially with the continuous development of big data and computing power, and have been widely used in various fields. A deep learning approach for intrusion detection using recurrent neural networks (RNN-IDS) is posed in this paper following this line of thinking.

3. RELATED WORK

Recurrent neural networks involve input units, output units and hidden units and the most important work is performed by the hidden unit. The RNN model essentially has a one-way stream of information from the input units to the hidden units and the reconstruction of the one-way flow of information from the previous temporal concealment unit to the new timing hiding unit is shown in Fig. They will consider hidden units as the entire network's storage, which is reminiscent of end-to-end information. As we unfold the RNN we may find it reflects the profound thinking. For supervised classification learning an approach to RNNs can be used.

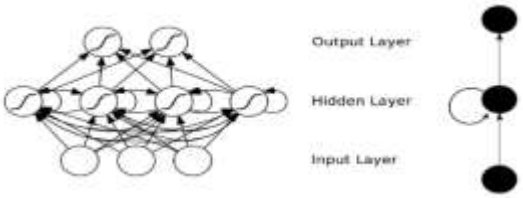

Fig. 1 Recurrent Neural Network

3.1 Dataset Description

The NSL-KDD dataset generated in 2009 is widely used in intrusion detection experiments. all the researchers use the NSL-KDD as the benchmark dataset, which not only effectively solves the inherent redundant records problems of the KDD Cup 1999 dataset but also makes the number of records reasonable in the training set and testing set, in such a way that the classifier does not favor more frequent records. The dataset covers the KDDTrainC dataset as the training set and KDDTestC and KDDTest datasets as the testing set, which has different normal records and four different types of attack records. . There are 41 features and 1 class label for each record and the features includes basic features, content features, and traffic features. Some specific attack types that disappear in the training set are added to the testing set to provide a more realistic theoretical basis for intrusion detection.

3.2 Data Preprocessing

Numericalization: NSL-KDD dataset contains 38 numeric features and 3 non-numeric features. Recurrent neural network model requires a numeric matrix as an input, so we need to convert these non-numeric features into numeric form. The non-numeric features are ‘protocol type ’, ‘service’ and ‘flag’ features. For example, the feature ‘protocol type ’has three types of values, ‘udp’, ‘tcp ’, and ‘icmp ’, and its numeric values are encoded as binary vectors (1,0,0), (0,1,0) and (0,0,1). Similarly, the feature ‘service ’has 70 different types of values, and the feature ‘flag ’has 11 different types of values. Continuing in this way, 41-dimensional features map into 122-dimensional features after transformation.

Normalization: For some features, the difference between the maximum and minimum values has a very large scope. Such features are ‘duration[0,58329] ’, ‘src bytes [0,1.3X109] ’and ‘dst bytes [0,1.3X109] ’. Apply the logarithmic scaling method for scaling to these features to obtain the ranges of ‘duration[0,4.77]’, ‘src bytes [0,9.11] ’and ‘dst bytes [0,9.11]’. Then map every feature to the [0, 1] range linearly using equation (1), where Max is the maximum value and Min is a minimum value for each feature.

$$x_i = \frac{x_i - Min}{Max - Min}$$

3.3 Methodology

It is apparent that the RNN-IDS model training consists of two parts-Forward Propagation and Back Propagation. Forward Propagation is responsible for measuring the out-placed values, and Back Propagation is responsible for passing the residuals that have been accumulated to update the weights that are not fundamentally different from the normal neural network creation.

3.4 Training

Forward Propagation: After initializing the model, forward propagation is the step to be taken to test its performance, i.e. how well the neural network is acting. Forward propagation in neural networks is to predict the output values and equate them to the actual / real value to get the error or loss. To measure the loss the actual value and the expected values are used. As the measurement flow goes in the normal forward direction, i.e. from the input through the neural network to the output, the forward propagation is therefore named.

Weights Update: Weight updating is done using back propagation which means error spreading backwards. In this first, calculate the partial loss derivatives regarding weight matrices and bias vectors, then propagate backwards to update the weights.

Testing: Testing is the final step in assessing model performance. The performance of the model is predicted using KDDTest+ and KDDTest-21 sets that are in the NSL-KDD dataset. Confusion matrix can be used to determine model accuracy which is the most significant performance metric used to test RNN model performance.

4. EVALUATION METRICS

In their model the most important intrusion detection efficiency indicator (Accuracy, AC) is used to calculate the RNN-IDS model performance. They implemented the detection rate and false positive rate, in addition to the accuracy.

The True Positive (TP) is equal to those properly rejected and represents the number of records of anomalies that are classified as anomalies.

The False Positive (FP) is the inverse of wrongly denied, which refers to the number of regular records marked as anomaly. The True Negative (TN) is equal to those accepted wrong, which represents the number of normal records that are defined as being regular.

The False Negative (FN) is equivalent to those incorrectly admitted, and it denotes the number of anomaly records that are identified as normal.

4.1 Accuracy

the percentage of the number of records classified correctly versus total the records shown as

$$AC = \frac{TP + TN}{TP + TN + FP + FN}$$

Table shows the definition of confusion matrix.

Predicted Class / Actual Class	anomaly	normal
anomaly	TP	FN
normal	FP	TN

Precision: It measures number of correct records penalized by number of incorrect records. Precision= TP/ (TP+FP)

Recall: It measures number of correct records as a number of missed entries. Recall= TP/ (TP+FN) 4. False Alarm Rate: Number of normal patterns classified as attack divided by total number normal patterns. FAR= FP/ (FP+TN)

5. CONCLUSION

Detection of intrusion plays an important role in security of the information. Using deep learning mechanism, not only improve system performance but also reduce some form of data dimension redundancies. It helps to achieve higher accuracy and detection rates with low false positives. Presenters can easily detect multi-class classification problems with respect to their category from the analysis of paper using different learning mechanisms.

6. ACKNOWLEDGMENT

I am glad to express my sentiments of gratitude to all who rendered their valuable guidance to us. I would like to express my appreciation and thank to Dr. Yogesh Kumar Sharma, Associate Professor, Head/Research Coordinator Department of Computer Science and Engineering, Shri Jagdishprashad Jhabarmal Tibrewala University, Vidyanagari, Jhunjhunu, Rajasthan.India for valuable guidance, thankful to my family and friends for their encouragement and support. I am sincere thank to experts and reviewers for their comments and suggestions.

7. REFERENCES

[1] Chuanlong yin, Yuefei zhu, Jinlong fei, and Xinzheng he, "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks" IEEE. Translations, VOLUME 5, 2017

[2] Harshal Karande, Shyam Gupta, "Ontology based intrusion detection system for web application security" International Conference On Communication Networks (lCCN), 2015

[3] A. Javaid, Q. Niyaz,W. Sun, and M. Alam, ``A deep learning approach for network 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.

[4] 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", in Proc. Int. Conf. Wireless Netw. Mobile Commun. (WINCOM) Oct. 2016, pp. 258_263

[5] Dr.Yogesh Kumar Sharma, Rokade Monika D. (2019), "Deep and Machine Learning Approaches for Anomaly-Based Intrusion Detection of Imbalanced Network Traffic", IOSR Journal of Engineering (IOSR JEN), ISSN (Online): 2250-3021, ISSN (Print): 2278-8719, pp.63-67

[6] Dr. Yogesh Kumar Sharma and S Pradeep (2019), "Deep Learning Based Real Time Object Recognition for Security in air defence", "International Conference on "Computing for Sustainable Global Development", ISSN 0973-7529; ISBN 978-93-80544-32-8, pp.64-67

[7] Manoj s. Koli, Manik K. Chavan, "An Advanced method for detection of botnet traffic using Interhnal Intrusion Detection", 2017 International Conference on (ICICCT), March 10-11, 2017, Sangli, India.

[8] M. Sheikhan, Z. Jadidi, and A. Farrokhi, ``Intrusion detection using reduced-size RNN based on feature grouping," Neural Comput. Appl.,vol. 21, no. 6, pp. 1185_1190, Sep. 2012.