JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.1505-1514
https://publishoa.com
ISSN: 1309-3452

Network Intrusion Attack Detection and Prevention using Various Soft Computing and Deep Learning Techniques

Maithili S. Deshmukh, Dr. A. S. Alvi
Department of Information Technology, Prof. Ram Meghe Institute of Technology & Research, Badnera

Received 2022 April 02**; Revised** 2022 May 20**; Accepted** 2022 June 18**.**

Abstract
In modern times, there has been a substantial improvement in network malwares and threats offences, which constructs excellent attention of network secrecy and protection appearances. Due to the elevation of the technology, network wormhole attacks are converting immensely complicated. Such that the current exposure systems are not satisfactory adequate to discuss this detection and prevention. Therefore, the execution of a unique and efficient network invasion detection system would be important to resolving the problems. In this research, we utilize hybrid deep learning nbased methods, especially, Convolutional Neural Networks (CNN) as well as Recurrent Neural Networks (RNN) to produce a unique apprehension system which can distinguish different network intrusions. Additionally, we estimate the representation of the proposed approach that demonstrates different evaluation techniques, and we impersonate an association among the consequences of our recommended explication to find the best model for the specific network intrusion detection scheme. The experiment analysis of system has done with numerous machines learning algorithm that executed in Weka environment on KDDCUP99 and NSLKDD dataset. Finally we conclude system provides better detection than machine learning algorithms on well-known network datasets.
Keywords: Network Intrusion, Host base Intrusion System, Deep Learning, Feature extraction, classification.

1. Introduction

Since technology has advanced, to exploit weaknesses and other disruptive behaviours, modern and sophisticated cyber-attacks would be introduced to break across networks. Communications infrastructure is among the critical systems with a dense variety of different cyber warfare styles, such as Denial of Service (DoS) as well as Distributed Denial-of-Service called (DDoS) threats, data Flooding attack, Ping of Death (PoD), hammered attack, Scan attacks, etc. Therefore, a great deal of effort is already put into implementing different effective types of methods to prevent these attacks and ensuring that the network is safe and protected while retaining high standards of protection efficiency to legal users in the system. Several main drawbacks of the Machine Learning (ML) based IDS frameworks is essential for the required training time to observe the large dataset of the network data stream. Nevertheless,

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.1505-1514
https://publishoa.com
ISSN: 1309-3452

with the use of some new methods and new emerging advanced technology, deep hybrid learning, delivers an operative learning contrivance. The training time is reduced with the asynchronously developed parallel method innovation, and the precision of the system's functionality is enhanced.

The limitations of current deep learning classifiers can be solved by Deep Learning communication networks such as CNN and RNN. Besides, Deep Learning Frameworks are known with their great accuracy and enhanced efficiency of Network Intrusion Schemes, the key aspect. Therefore, throughout this analysis, depending on its existence and efficacy, we selected CNN to identify artifacts that could be used to detect natural and malicious traffic processes. RNN, is from the other hand, was chosen because of its characteristics of recalling past incidents that could lead to critical success in classifying different types of attacks as a consequence. The aim of this research is always to study the efficacy of CNN and RNN implementation in NIDS. The research was conducted an observational experiment to analyze the effectiveness of the suggested framework in activity recognition.

2. Literature Survey

According to Dong Bo et. al. [1] based on deep learning approaches which are inspired by the complexities of the human cognitive ability from the higher characteristics to the concept of advanced concentration. The Deep Belief Network (DBN) includes study functions which map from inputs and outputs due to multilevel abstraction. The cooperative learning is not based on attributes created by humans. For-layer DBN utilizes the Restricted Boltzmann Machine (RBM) as an unsupervised algorithm. Advantages were also: Deep encoding is its prepared to change to evolving data contexts, ensuring the methodology performs comprehensive data analysis. It also detects device anomalies which include anomaly detection and traffic identification. The drawbacks are: requirement for better and quicker data assessment. The main purpose of [2] is to review and summarize the deep learning work on the monitoring of machine health. Deep learning implementations in computer healthcare monitoring were also extensively examined from of the following areas: Auto encoder or its variants, deep learning computers and their variants, as well as Deep Belief Network but also Deep Boltzmann Machines, CNN, and RNN. One potential benefit is: DL-based MHMS requires no considerable information systems and understand-how. Implementing deep learning techniques is not restricted to various types of computers. Major disadvantage are: DL-based MHMS efficiency depends to a large degree on the size and accuracy of the dataset. Indicates by use of a filled decoder, just one deep learning application, to construct an FDC model to retrieve and classify functions at the same time. The SDA model [3] Management support and distinct genetic for voltage sensor fault surveillance, and are robust against noise reduction. An SDA is a denunciation of auto-encoders which are layered layer after layer. This multidimensional architecture enables positive externalities to be learned from

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.1505-1514
https://publishoa.com
ISSN: 1309-3452

complex input data, such as multivariate time series datasets and high definition images. The advantages are: The SDA model is useful in real-life applications. The SDA model suggests learning from different sensors without having to effectively pre-process the normal and fault-related features. Disadvantages are: really have to investigate a trained SDA to determine the processing parameters that so many impact classified outcomes.

The novel deep learning-based recurrent neural networks (RNNs) theory specifies short prison texts in [4] for automatic vulnerability analysis, which can categories short (safe and non-insecure) messages. In this research word2vec collects the function of simple comments capturing word customer orders for each phrase and is converted to an input image. In special, words to similar structure are assigned to a similar position throughout the subspace, but instead classified by RNNs. The potential benefits are: The RNNs model achieves an average 92.7% precision which is higher than SVM. To boost overall performance, the use of ensemble frameworks to integrate various extraction features and classification algorithms. The disadvantages are: messages that are not large-scale but only simple codes are applicable. Signature-based features methodology as a profound convolution neural network [5] explains plate detection, personality detection, as well as cloud platform differentiation. Obtaining important features allows the LPRS to adequately recognize the license plate in a difficult situation including such I busy traffic with numerous plates in image (ii) plate alignment towards brightness, (iii) new details on the plate, (iv) distortion of excessive wear and (v) distortion of the captured images as image data during bad weather. Advantages are: The proposed algorithm's supremacy in recognizing LP accuracy rather than other traditional LPRS. Downsides are: there are some unknown or falsely-detected images.

ArwaAldweeshet. al. [6] the main simulations evaluating deep learning for malware detection were evaluated and compared, as well as the current sample was based on historical ones. It provided an interesting fine-grained categorization that considered different modelling aspects, namely data input, recognition, implementation, and strategies for assessment. This thus offered an in-depth analysis of the relevant scientific investigations in auditory learning style IDS. Throughout the context of deep processing, that auto-encoder is among the most important models for extracting features from the high dimensional data. Fahimeh et al. [7] implemented a comprehensive auto-encoder method to improve the detection mechanism. Their significant risk factor Auto-Encoder-based intrusion prevention system comprises of four auto-encoders using the auto-encoder output to the current layer as that of the auto-encoder result of the interactions. The proposed method achieved accuracy of detection 94.71% of total KDDCUP99 testing set of 10%. The performance of the scheme proposed by Zhang et. al. [8] is evaluated by experiments performed on the UNSW-NB dataset. IDS consist basically of a discovery existing engine on the DAE function as well as an MLP based classifier. Nathan Shone et. al. [9] System proposed the Non symmetric Deep Auto-Encoder which is evaluated by supervised

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.1505-1514
https://publishoa.com
ISSN: 1309-3452

Random forest classification algorithm on KDDCup98 and NSLKDD. This system fails in redundant feature selected by feature selection method. Multilayer classification with DBN has used for IDS by [10]. This research has shown that DBN should really learn a better training algorithm and execute on acknowledgement of intrusion activities. This system is not able to detect real time attacks in network environment.

New flexible, deep-learning intrusion detection software, this research effectively tackles the problem of processing high quantities of security-related data for network security tasks in [11]. It utilizes Apache Spark as just a big data analysis platform to manage a large range of network traffic data. Framework also suggests a hybrid scheme which incorporates the benefits of deep network and computer vision solutions. Initially, system that ensures transceiver network is used to extract latent features, followed by several classification-based detection techniques such as supporting vector machines, random forests, discriminate analysis and naive bays that are being used to find vulnerabilities in massive network traffic effectively and efficiently, and also some UNB ISCX 2012 repository.

DilaraGumusbaset. al. [12] describes widespread impression of machine learning approaches for cyber security of IDS, with particular emphasis on current deep learning (DL) approaches. The analysis analyses recent approaches with regard to their mechanisms for intrusion detection, success outcomes and drawbacks, as well as how system use large historical databases to guarantee a best detection accuracy and evaluation. However, a thorough analysis of the information security benchmark network log has obtainable .Jun Gao et. al. [13] focused on investigate of deep learning-based omni-intrusion detection (IDS) networks for supervisory control and data acquisition (SCADA) capable of detecting both temporarily uncorrelated and correlated assaults. With regard to the IDSs developed in this paper, a feed forward neural network (FNN) may detect temporarily uncorrelated attacks at a rate of 99.96% but correlated attacks at a rate of 58.2%. In comparison, long-term memory (LSTM) detects 99.56 % of associated attacks while 99.30 % of uncorrelated attacks. The combination of LSTM and FNN with an ensemble method further increases the efficiency of IDS with F1 by 99.68 % irrespective of the time differences between the data packets.

Wenjuan Wang et. al. [14] intentions to use profound defined to automatically mine indispensable feature demonstrations as well as efficiently achieve high recognition enactment. For unsupervised extraction of features an efficient stacked contractive auto encoder (SCAE) method is provided. When using the SCAE process, from raw network traffic, improved and reliable lower dimensional structures can be automatically taught. A new cloud based IDS is developed using the SCAE classification algorithm and supporting vector machine (SVM). Amine Ferrag Mohammed et. al. [15] developed a dynamic energy storage energy network focused on deep learning and blockchain, entitled DeepCoin. The DeepCoin system uses two methods, a public blockchain scheme and a profound supervised learning scheme. The blockchain-based

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.1505-1514
https://publishoa.com
ISSN: 1309-3452

framework consists of five phases: initialization phase, compromise phase, block construct phase and approval build phase, then changing of view process. It implements a modern, scalable peer-to - peer resource network on the practical Byzantine fault algorithm tolerance which achieves high performance. The proposed machine learning framework is an IDS, which uses neural networks in the blockchain-based power grid to automatically detect attacks or financial fraud.

3. Methods

The deep learning is an evolving trend in the sub section of machine learning techniques. It is sub-field of machine intelligence in a simulated neural network. To be repaired, we can arrange extensive evaluations of items that utilised an in-depth learning approach in the application field. Millions of data points are located in the process. The properties of deep learning are received from the data. If extensive capacities of data are usable, the device output can be reduced. Machine learning is a well-suited learning means of achieving better accuracy in expressions of results. System denotes three main kinds of knowledge: supervised, mid-supervised, as well as unsupervised.

Current techniques for identifying attacks on cloud infrastructures or the VMs inhabitant inside them don't adequately address cloud specific issues. Despite the huge efforts employed in past studies in regards to the behaviour of certain kinds of malware in the Internet, so far little has been done to handle malware present in cloud. Specifically, the investigations in expected to modify the performance of traditional Intrusion Detection Systems (IDS). Intrusion detection or prevention systems are as yet basic to by and large data security achievement. Throughout the years, IDS/IPS has over and over been proclaimed dead just to be restored each time. The objective of IDS is to distinguish cyber-attack by examining the signature of data packets as they cross the network. By analysing these packets we get real time alarms when awful things occur. IDS are a passive, detect and alarm just device. IPS includes a active protection technique for adjusting to the risk and hindering the traffic so the planned individual host is never come to.

Here, with attention to the deep neural network, intrusion measurement is carried in. Intrusion is a concept that may compromise the integrity of the internal quality management system. And the method of detecting intrusion is another identification of intrusion. The methodology of vulnerability scanning is divided into two approaches, i.e. detection of irregularities or detection of exploitation. These two approaches are briefly listed below:

3.1 Misuse Detection: It is also known as the signature-based detection. Here, the behaviour of user is compared with existing pattern. But the problem with this technique is that only the known attack. Intrusion detection based on signature is not suitable for real time applications.

3.2 Anomaly Detection: It is relating to the normal behaviour of user. If any action should pretend to be different with respect to the normal behaviour of user or system. The anomaly detection consists of two different categories: Threshold- based and profile-based. In the threshold-based

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.1505-1514
https://publishoa.com
ISSN: 1309-3452

detection, it counts the number of occurrences of any event with some time interval.In this process, intrusion detection system can check the behavior of network on the basis of binary classification and on the multiclass classification, means that identify whether it is normal or anomalous. The multiclass classification is categorized in following different category. They are as follows: It is either normal or from other category i.e. Denial-of-Service, User to Root, Root to Local and Probing

4. System Design

The proposed work shows how to identify different characteristics of attacks caused by an attacker on the targeted victim computer, and the system automatically evolves against the protection measures taken. Alongside the security mechanisms, the IDS continually introduce new functionality. The proposed solutions like firewalls or antivirus, which are classic methods of detection using threat fingerprint using rules and policies available in repositories. To detect the runtime problems to identify and remove such situation can be focused in this work. Hence it is a smarter approach to use both static and behavioural approaches to identify and avoid IDS.

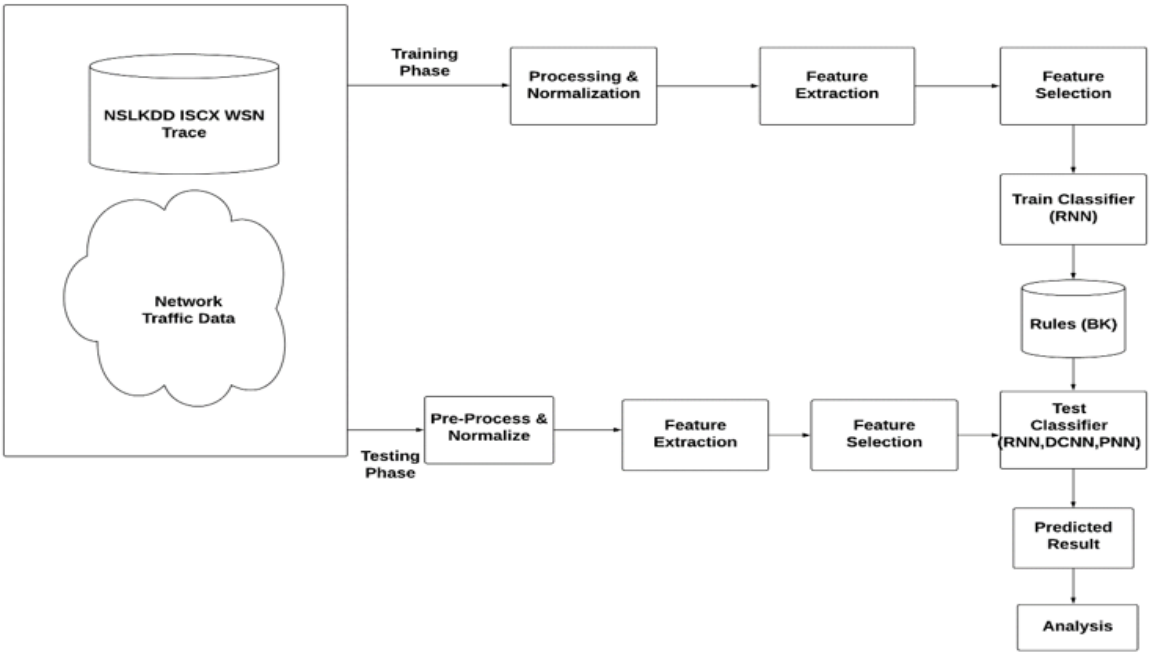


Figure 1 : Proposed system design

Figure 1 demonstrates the device architecture for IDS dependent detection of runtime anomalies. The first section is the training process that stores them into the rule repository, learning the rules according to historical knowledge of the method. During the detection stage, this repository is used. The optimal solution is used to create dynamic rules in the learning strategy or strategies such as similarity measures are used to check hybrid deep learning methods.

The proposed system works with deep learning approach. Program first collects

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.1505-1514
https://publishoa.com
ISSN: 1309-3452

examination data from various online and offline outlets. Once the data is collected through the program it applies pre-process and feature extraction. After this rules are created and stored into local database directory called as Background rules (BK Rules). Background rules are given as an input to the deep learning approach for the classification of sub attack. In this work, the hybrid deep learning algorithm was applied to pre-processing distilled data to create a learning model, and the entire NSLKDD dataset was used for testing. In the end, the accuracy, detection rate, and false alarm rate were determined to assess the detection efficiency of the model.

5. Algorithm

The proposed system categorized into two different parts, one is machine learning classifiers, and another is deep learning classifiers. Weka toll has used for validation of machine learning classification algorithms while the below algorithm has used for deep learning classification..

Algorithm 1 : Deep Learning based training algorithm

Input : Training data TrainDB, Testing data TsetDB

Output :Train Module TM

Train(TrainDB)

Step 1 :Normlized ← *Normalization (* $\sum_{k=0}^{n} TrainDB\ [k].select\ if\ descrete$*)*

Step 2 : Extract co-relational features from Normalized set

Step 3 :i ← *0*

epoch ← *100*

while (I != epoch)

Module[] ←*TrainClassifier(* $\sum_{k=0}^{n} Normlized\ [k]$*)*

i=i+1

end loop

Step 4 : Return Module[]

Test(TestDB, Module[])

Step 1 : for each (read instance from TestDB)

Step 2 :AttributesSet[] ←*instance.splite*

Step 3: calculate weight of respective instance using below equation $weight = Testclassifier\ (AttributesSet[index] = \sum_{i=1}^{n} Module[i])$

Step 4: if (weight > Threshold)

Flag =1

Return attack

Step 5: if(flag==0)

Return normal

6. Results

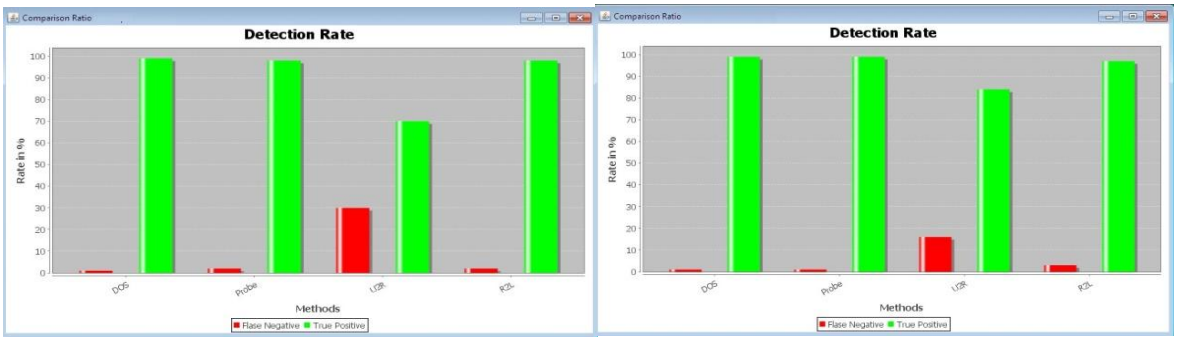
The proposed implementation has done in open source environment with Weka environments. The investigations conducted this work completely are reviewed in Table 1. From the result generated from these practices, one can understand that have essentially exceeded in terms of advanced algorithm F1 score, correctness and efficiency with the lowest number of epochs needed to reach its optimal execution.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.1505-1514
https://publishoa.com
ISSN: 1309-3452

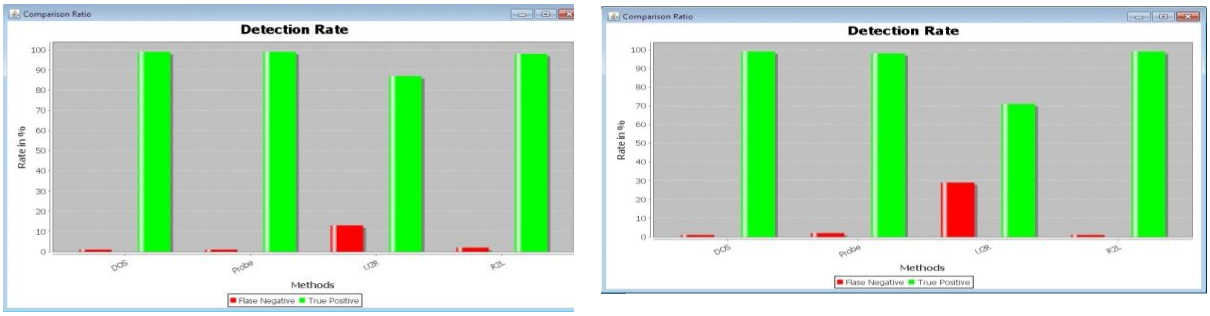
Classifier	DOS		Probe		U2R		R2L	
	TPR	FNR	TPR	FNR	TPR	FNR	TPR	FNR
SVM	99%	1%	98%	2%	70%	30%	98%	2%
Naïve Bayes	99%	1%	99%	2%	84%	16%	97%	3%
J48	99%	1%	99%	1%	87%	13%	98%	2%
Random Forest	99%	1%	98%	2%	71%	29%	99%	1%
ANN	99%	1%	98%	2%	76%	24%	97%	3%
Adaboost	99%	1%	98%	2%	77%	23%	99%	1%

Table 1 : Confusion matrix of system using various machine learning algorithms

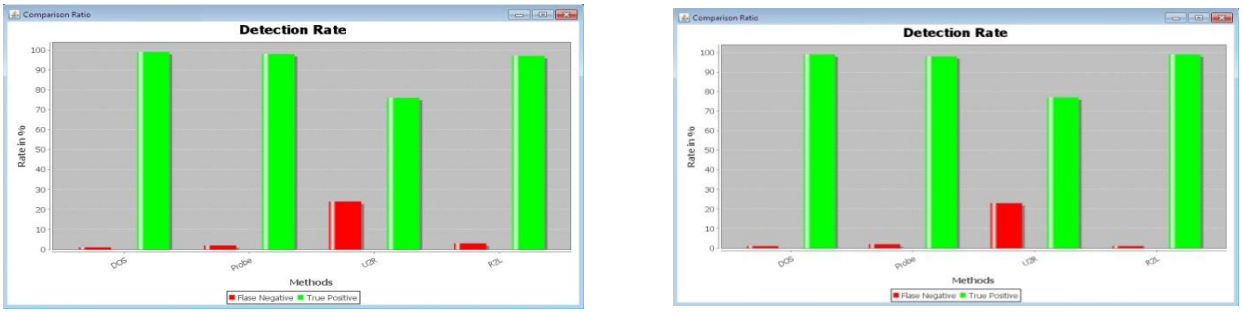
In another experiment we have tested NSLKDD benchmark dataset with various cross validation and also tuning some default hyper parameters. The below figure a to f shows detail explanation of system.



(a) **Accuracy with 20% test NSLKDD** (b) **Accuracy with10% test NSLKDD**



(c) Accuracy with 5 fold cross validation **(d) Accuracy with 10 fold cross validation**



(e) Accuracy with 15 fold cross validation **(f) Accuracy with 20 fold cross validation**

Figure 2 : Detection and classification accuracy with various machine learning algorithms using NSLKDD dataset.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p.1505-1514
https://publishoa.com
ISSN: 1309-3452

The above discussion demonstrates how to identify different characteristics of attacks caused by an attacker on the targeted victim computer, and the system automatically evolves against the protection measures taken. Alongside the security mechanisms, the IDS continually introduce new functionality. The proposed solutions like firewalls or antivirus, which are classic methods of detection using threat fingerprint using rules and policies available in repositories. To detect the runtime problems to identify and remove such situation can be focused in this work. Hence it is a smarter approach to use both static and behavioral approaches to identify and avoid IDS. Figure 2 demonstrates the device architecture for IDS dependent detection of runtime anomalies. The first section is the training process that stores them into the rule repository, learning the rules according to historical knowledge of the method. During the detection stage, this repository is used. The optimal solution is used to create dynamic rules in the learning strategy or strategies such as similarity measures are used to check hybrid deep learning methods.

7. Conclusion

This research basically demonstrates intrusion detection effective techniques in vulnerable network environment. We also proposed deep learning based techniques for network circulation anomaly detection. The proposed approach determines detect the malicious behaviours into the normal profiles based on risk score. Using proposed algorithms we able to extract some hidden features of network packet dataset. We also proposed new feature extraction algorithm to reduce noise with input training data and then classify using proposed deep learning algorithms.

This approach effectively demonstrates the detection of malicious network behaviours from network traffic. The proposed IDS have used various classification algorithms and well-known network intrusion dataset. Detection and others are focused on action, scientifically referred to as 'detection focused on anomaly'. Utilizing CNN and RNN hybrid classification classifiers using the NSL-KDD dataset for training as well as testing, the present study introduced the subsequent methodology of anomaly-based intrusion prevention. The recommended classification model has been observed to have superseded some other strategies of precision, F1 ranking, recall and precision of above 97% amongst these evaluated deep learning models. The restricted computing power that caused a long time for training with each prototype was one of the primary constraints faced mostly during the study. We plan to implement faster systems that can train the algorithms in an even quicker period to address this problem. In future research, by incorporating different kinds of deep learning techniques with various levels and hypotheses testing, we will also continue studied the impact of various deep learning models and how these variables will have a significant effect on the effectiveness of the NIDS.

References

[1] Dong Bo, and Xue Wang. Comparison Deep Learning Method to Traditional Methods Using for Network Intrusion Detection, 8th IEEE International Conference on Communication Software and Networks 2016.

[2] Zhao, Rui, Deep Learning And Its Applications to Machine Health Monitoring: A Survey, arXiv preprint arXiv:1612.07640 (2016).
[3] Lee, Hoyeop, Youngju Kim, and Chang Ouk Kim. A Deep Learning Model For Robust Wafer Fault Monitoring With Sensor Measurement Noise, IEEE Transactions on Semiconductor Manufacturing 30.1 (2016): 23-31.
[4] You, Lina. A Deep Learning-Based RNNs Model for Automatic Security Audit of Short Messages, 16th International Symposium on Communications and Information Technologies (ISCIT), 2016.
[5] Polishetty, Rohith, Mehdi Roopaei, and Paul Rad. A Next-Generation Secure Cloud-Based Deep Learning License Plate Recognition for Smart Cities. 15th IEEE International Conference on Machine Learning and Applications (ICMLA). 2016.
[6] AldweeshArwa, AbdelouahidDerhab, and Ahmed Z. Emam. Deep Learning Approaches For Anomaly-Based Intrusion Detection Systems: A Survey, Taxonomy, And Open Issues. Knowledge-Based Systems 189 (2020): 105124.
[7] FarahnakianFahimeh, and JukkaHeikkonen. A Deep Auto-Encoder Based Approach for Intrusion Detection System. 20th International Conference on Advanced Communication Technology (ICACT), 2018.
[8] Zhang, Hongpo, et al. An Effective Deep Learning Based Scheme for Network Intrusion Detection. 24th International Conference on Pattern Recognition (ICPR).IEEE, 2018.
[9] Shone, Nathan, et al. A Deep Learning Approach To Network Intrusion Detection. IEEE transactions on emerging topics in computational intelligence 2.1 pp 41-50, 2018.
[10] Gao, Ni, et al. An Intrusion Detection Model Based On Deep Belief Networks. Second International Conference on Advanced Cloud and Big Data. IEEE, 2014.
[11] Mighan, SoosanNaderi, and Mohsen Kahani. A Novel Scalable Intrusion Detection System based on Deep Learning. International Journal of Information Security pp 1-17, 2020.
[12] Gümüşbaş, Dilara. A Comprehensive Survey of Databases and Deep Learning Methods for Cybersecurity and Intrusion Detection Systems. IEEE Systems Journal, 2020.
[13] Gao, Jun, Omni SCADA Intrusion Detection Using Deep Learning Algorithms. IEEE Internet of Things Journal, 2020.
[14] Wang, Wenjuan. Cloud Intrusion Detection Method Based on Stacked Contractive Auto-Encoder and Support Vector Machine. IEEE Transactions on Cloud Computing, 2020.
[15] Ferrag, Mohamed Amine, and LeandrosMaglaras. DeepCoin: A novel deep learning and Blockchain-based energy exchange framework for smart grids. IEEE Transactions on Engineering Management, 2019.
[16] Deshmukh MS, Alvi AS. Detection and Prevention of Malicious Activities in Vulnerable Network Security Using Deep Learning. InProceedings of the 2nd International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications 2022 (pp. 319-326). Springer, Singapore.