IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 5, May 2023

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.423

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205230 |

Intrusion Detection System Using Machine Learning Algorithm

Pachhade Rasika Chandrashekhar, Prof. Joshi S. G.

Department of Computer Engineering, Vishwabharti Academy's College of Engineering, Sarola Baddi, Ahmednagar, India

ABSTRACT- Information redundancy and inapplicability have generated a long-term problem in network traffic classification. One of the key topics of NIDS research in recent years has been the application of machine learning and shallow learning approaches. This research provides a new deep learning model for NIDS operation in current networks. The model demonstrates a combination of deep and shallow learning, allowing it to accurately analyse a wide spectrum of network traffic. For unsupervised feature learning, the unique approach provides a non-symmetric deep autoencoder (NDAE). Furthermore, it presents a unique deep learning classification display built with stacked NDAEs. Our suggested classifier was implemented in GPU-enabled TensorFlow and evaluated using the KDD Cup '99 and NSL-KDD datasets. The performance of network intrusion detection analysis datasets, specifically KDD Cup 99 and NSL-KDD, was tested. The contribution task is to develop intrusion prevention systems (IPS), which are more advanced systems capable of taking rapid action to prevent or reduce hostile behaviour.

KEYWORDS- Deep learning, Anomaly detection, Auto-encoders, KDD, Network security

I. INTRODUCTION

The deployment of a comprehensive and effective Network Intrusion Detection System (NIDS) is one of the primary difficulties in network security. Despite substantial advancements in NIDS technology, the bulk of solutions continue to rely on less capable signature-based techniques rather than anomaly detection techniques. The present challenges are that existing methodologies result in ineffective and inaccurate attack detection. The volume of network data, in-depth monitoring and granularity required to increase efficacy and accuracy, and lastly the number of different protocols and diversity of data travelling are the three key limits. The implementation of machine learning and shallow learning approaches has been the primary focus of NIDS research. Deep learning research has shown that its superior layer-wise feature learning can outperform or at least match shallow learning strategies. It is capable of aiding deeper network data analysis and faster detection of any irregularities. We present a novel deep learning model in this paper to enable NIDS functioning within current networks.

II. RELATED WORK

The paper [1] focuses on deep learning methods which are inspired by the structure depth of human brain learn from lower level characteristic to higher levels concept. It is because of abstraction from multiple levels, the Deep Belief Network (DBN) helps to learn functions which are mapping from input to the output. The process of learning does not dependent on human-crafted features. DBN uses an unsupervised learning algorithm, a Restricted Boltzmann Machine (RBM) for each layer. Advantages are: Deep coding is its ability to adapt to changing contexts concerning data that ensures the technique conducts exhaustive data analysis. Detects abnormalities in the system that includes anomaly detection, traffic identification. Disadvantages are: Demand for faster and efficient data assessment.

The main purpose of [2] paper is to review and summarize the work of deep learning on machine health monitoring. The applications of deep learning in machine health monitoring systems are reviewed mainly from the following aspects: Autoencoder (AE) and its variants, Restricted Boltzmann Machines and its variants including Deep Belief Network (DBN) and Deep Boltzmann Machines (DBM), Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN).Advantages are:DL-based MHMS do not require extensive human labor and expert knowledge.The applications of deep learning models are not restricted to specific kinds of machines. Disadvantages are: The performance of DL-based MHMS heavily depends on the scale and quality of datasets.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205230 |

Proposes the use of a stacked denoisingautoencoder (SdA), which is a deep learning algorithm, to establish an FDC model for simultaneous feature extraction and classification. The SdA model [3] can identify global and invariant features in the sensor signals for fault monitoring and is robust against measurement noise. An SdA is consisting of denoisingautoencoders that are stacked layer by layer. This multilayered architecture is capable of learning global features from complex input data, such as multivariate time-series datasets and high-resolution images.Advantages are:SdA model is useful in real applications.The SdA model proposes effectively learn normal and fault-related features from sensor signals without preprocessing.Disadvantages are:Need to investigate a trained SdA to identify the process parameters that most significantly impact the classification results.

Proposes a novel deep learning-based recurrent neural networks (RNNs)model [4] for automatic security audit of short messages from prisons, which can classify short messages(secure and non-insecure).In this paper, the feature of short messages is extracted by word2vec which captures word order information, and each sentence is mapped to a feature vector. In particular, words with similar meaning are mapped to a similar position in the vector space, and then classified by RNNs.Advantages are:The RNNs model achieves an average 92.7% accuracy which is higher than SVM.Taking advantage of ensemble frameworks for integrating different feature extraction and classification algorithms to boost the overall performance.Disadvantages are:It is apply on only short messages not large-scale messages.

Signature-based features technique as a deep convolutional neural network [5] in a cloud platform is proposed for plate localization, character detection and segmentation. Extracting significant features makes the LPRS to adequately recognize the license plate in a challenging situation such as i) congested traffic with multiple plates in the image ii) plate orientation towards brightness, iii) extra information on the plate, iv) distortion due to wear and tear and v) distortion about captured images in bad weather like as hazy images. Advantages are: The superiority of the proposed algorithm in the accuracy of recognizing LP rather than other traditional LPRS. Disadvantages are: There are some unrecognized or miss-detection images.

In [6] paper, a deep learning approach for anomaly detection using a Restricted Boltzmann Machine (RBM) and a deep belief network are implemented. This method uses a one-hidden layer RBM to perform unsupervised feature reduction.The resultant weights from this RBM are passed to another RBM producing a deep belief network. The pre-trained weights are passed into a fine tuning layer consisting of a Logistic Regression (LR) classifier with multi-class soft-max.Advantages are:Achieves 97.9% accuracy.It produces a low false negative rate of 2.47%.Disadvantages are:Need to improve the method to maximize the feature reduction process in the deep learning network and to improve the dataset.

The paper [7] proposes a deep learning based approach for developing an efficient and flexible NIDS. A sparse autoencoder and soft-max regression based NIDS was implemented. Uses Self-taught Learning (STL), a deep learning based technique, on NSL-KDD - a benchmark dataset for network intrusion. Advantages are: STL achieved a classification accuracy rate more than 98% for all types of classification. Disadvantages are: Need to implement a real-time NIDS for actual networks using deep learning technique.

In [8] paper choose multi-core CPU's as well as GPU's to evaluate the performance of the DNN based IDS to handle huge network data.The parallel computing capabilities of the neural network make the Deep Neural Network (DNN) to effectively look through the network traffic with an accelerated performance.Advantages are:The DNN based IDS is reliable and efficient in intrusion detection for identifying the specific attack classes with required number of samples for training.The multicore CPU's was faster than the serial training mechanism.Disadvantages are:Need to improve the detection accuracies of DNN based IDS.

In [9] paper, proposes a mechanism for detecting large scale network-wide attacks using Replicator Neural Networks (RNNs) for creating anomaly detection models.Our approach is unsupervised and requires no labeled data. It also accurately detects network-wide anomalies without presuming that the training data is completely free of attacks.Advantages are:The proposed methodology is able to successfully discover all prominent DDoS attacks and *SYN Port* scans injected. Proposed methodology is resilient against learning in the presence of attacks, something that related work lacks.Disadvantages are:Need to improve proposed methodology by using stacked autoencoder deep learning techniques.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205230 |

Based on the flow-based nature of SDN, we propose a flow-based anomaly detection system using deep learning. In [10] paper, apply a deep learning approach for flow-based anomaly detection in an SDN environment.Advantages are:It finds an optimal hyper-parameter for DNN and confirms the detection rate and false alarm rate.The model gets the performance with accuracy of 75.75% which is quite reasonable from just using six basic network features.Disadvantages are:It will not work on real SDN environment.

III. OPEN ISSUES

'The current network traffic data, which is frequently large in size, poses a significant challenge to IDSs. Because of the computing difficulties in dealing with large data, "big data" slows down the entire detection process and may result in inferior classification accuracy.Machine learning technologies are commonly utilised in intrusion detection systems. However, most classic machine learning solutions are based on shallow learning and are incapable of properly addressing the massive intrusion data classification problem that arises in the face of a real network application environment. Shallow learning is also incompatible with intelligent analysis and the predefined needs of high-dimensional learning with massive data.

Disadvantages:
Computer systems and the internet have become key components of the critical system. IDSs face a significant issue with today's network traffic data, which is frequently massive in size. Because of the computing difficulties in dealing with large data, "big data" slows down the entire detection process and may result in inferior classification accuracy. Classifying a large amount of data frequently results in numerous mathematical issues, which leads to increased computer complexity.

IV. SYSTEM OVERVIEW

The authors present a unique deep learning approach to enable NIDS functioning within current networks in this research. The proposed model is a hybrid of deep and shallow learning that can appropriately analyse a wide range of network traffic. We combine the power of stacking our proposed Non-symmetric Deep Auto-Encoder (NDAE) (deep learning) with the accuracy and speed of Random Forest (RF) (shallow learning). This study introduces our NDAE, a non-symmetrical multiple hidden layer auto-encoder. NDAE is a hierarchical unsupervised feature extractor that scales effectively to handle high-dimensional inputs. It learns non-trivial features in a manner similar to that of a normal auto-encoder. Stacking the NDAEs provides a layer-wise unsupervised representation learning approach, allowing our model to learn the complicated correlations between various features. It may also refine the model by prioritising the most descriptive characteristics because it includes feature extraction capabilities.

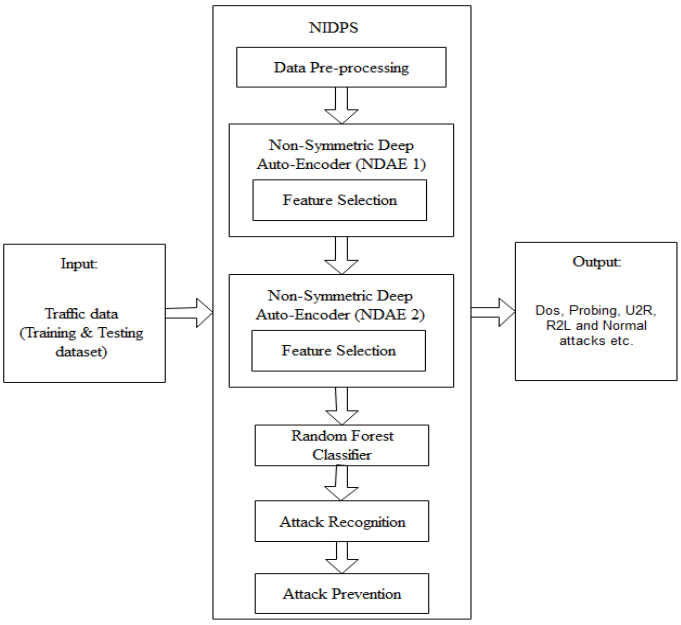

Fig. 1 Proposed System Architecture

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205230 |

V. CONCLUSION

In this study, we examined the issues that present NIDS approaches confront. As a result, we suggested our innovative NDAE approach for unsupervised feature learning. We then expanded on this by developing a new classification model based on stacked NDAEs and the RF classification algorithm. We also put in place an intrusion protection system. The results show that our approach provides high levels of accuracy, precision, and recall while requiring less training time. The proposed NIDS system improves accuracy by just 5%. As a result, precision must be improved further.

REFERENCES

1. 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.
2. R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring: A survey," Submitted to IEEE Trans. Neural Netw. Learn. Syst., 2016. [Online]. Available: http://arxiv.org/abs/1612.07640
3. H. Lee, Y. Kim, and C. O. Kim, "A deep learning model for robust wafer fault monitoring with sensor measurement noise," IEEE Trans. Semicond. Manuf., vol. 30, no. 1, pp. 23–31, Feb. 2017.
4. L. You, Y. Li, Y. Wang, J. Zhang, and Y. Yang, "A deep learning based RNNs model for automatic security audit of short messages," in Proc. 16th Int. Symp. Commun. Inf. Technol., Qingdao, China, Sep. 2016, pp. 225–229.
5. R. Polishetty, M. Roopaei, and P. Rad, "A next-generation secure cloud based deep learning license plate recognition for smart cities," in Proc. 15th IEEE Int. Conf.Mach. Learn. Appl., Anaheim, CA, USA, Dec. 2016, pp. 286–293.
6. K. Alrawashdeh and C. Purdy, "Toward an online anomaly intrusion detection system based on deep learning," in Proc. 15th IEEE Int. Conf. Mach. Learn. Appl., Anaheim, CA, USA, Dec. 2016, pp. 195–200.
7. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A deep learning approach for network intrusion detection system," in Proc. 9th EAI Int.Conf. Bio-Inspired Inf. Commun. Technol., 2016, pp. 21–26. [Online]. Available: http://dx.doi.org/10.4108/eai.3-12-2015.2262516
8. S. Potluri and C. Diedrich, "Accelerated deep neural networks for enhanced intrusion detection system," in *Proc. IEEE 21st Int. Conf. Emerg. Technol. Factory Autom.*, Berlin, Germany, Sep. 2016, pp. 1–8.
9. C. Garcia Cordero, S. Hauke, M. Muhlhauser, and M. Fischer, "Analyzing flow-based anomaly intrusion detection using replicator neural networks," in *Proc. 14th Annu. Conf. Privacy, Security. Trust*, Auckland, New Zeland, Dec. 2016, pp. 317–324.
10. 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.*,Oct. 2016, pp. 258–263.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.423

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

Scan to save the contact details