Unsupervised Feature Learning using a Novel Non-Symmetric Deep Auto encoder(NDAE) Model for NIDS Framework

Zohra Anzar Shaikh[1], Vidya Dhamdhere[2]
PG Student[1], Assistant Professor[2],
Department of Computer Engineering,
G.H. Raisoni College of Engineering and Management, Pune[1,2]

ABSTRACT: Network intrusion identification frameworks play a pivotal role in guarding computer networks. As of late, one of the fundamental concentrations within Network Intrusion Detection System(NIDS) inquire about the usage and application of Machine Learning(ML) Techniques. This paper proposed to enable NIDS network traffic for novel deep learning model. The novel approach proposes non-symmetric deep auto encoder(NDAE) for unsupervised feature learning. Moreover, it proposes novel deep learning classification display built utilizing stacked NDAEs. Our proposed classifier has been executed in KDD Cup '99 and NSL-KDD data-sets. The KDD Cup 99 and NSL-KDD dataset particularly are performance evaluated network intrusion detection datasets. The contribution work is to implement intrusion prevention system (IPS) contains IDS functionality but more sophisticated systems which are capable of taking immediate action in order to prevent or reduce the malicious behavior.

Keywords:- Deep Learning, Machine learning, Intrusion Detection, Auto Encoder, KDD, Network Security, Novel Approach.

1 INTRODUCTION

In today's world one of the superior defy in network security is the provision of a robust and effective Network Intrusion Detection System(NIDS). as opposed to anomaly detection techniques, now a days also majority of solutions is still manage using less capable Signature Based Techniques, The existing techniques is prevalent issues that prompts to inadequate and incorrect recognition of assaults. There are three constraints, for example, volume of system information, inside and out observing and granularity required to amend viability and precision and in conclusion the quantity of unmistakable conventions and disagreement of information crossing. The vital of NIDS ascertainment is the use of shallow learning and AI systems. The rudimentary deep learning probe will be vanquished that is upper layer-wise element realizing which is improved or if nothing else banter the performance of Shallow Learning Techniques. It competent of presenting a deeper scansion of accelerated discernment and network data of any anomalies. Right now, this paper conclude a novel deep learning model to empower NIDS activity within present days systems.

2 RELATED WORK

In paper [1] purposes the center of interest in methods of Deep Learning, Which is intuitive having structure profundity of human brain characteristics learn from lower level to higher level ideas. the Deep Belief Network (DBN) assists with taking in capacities which are mapping from contribution to the yield, in view of deliberation from different levels. The methodology of learning doesn't depend on human-created highlights. DBN has an unaided learning calculation and Restricted Boltzmann Machine(RBM) in every layer Advantages: Its Deep Coding has respectability to regulate to trading settings concerning information. It Detects variations and abnormalities in the system which contain traffic distinguishing proof and irregularity discovery. Disadvantages are: Need for quickened and viable information evaluation.

The paper [2] is to view and survey the Deep Learning methods working on Machine Health Monitoring System(MHMS). Application of deep learning in the MHMS are see from following tendency: Auto encoder (AE), Restricted Boltzmann Machine(RBM) and its variations having Deep Boltzmann Machines (DBM) and Deep Belief Network (DBN), Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN). Preferences are: Deep Learning based MHMS don't required expansive human work and achieve information. Deep learning models are not limited to explicit sort of machines. Inconveniences: The tricks of deep learning put together MHMS more hinge with respect to the measure and quality of data sets.

The Paper[3] having Stacked denoising Auto-encoder(SdA), and algorithm of deep learning to conquer FDC model for synchronous classification and feature extraction. In paper[3] SdA model can discriminate invariant and global worldwide highlights is vigorous against measurement noise in the sensor signals for deficiency observing. denoising auto encoders is stacked layer by layer which contain SdA. Multilayered Architecture has ability of learning complex input information from global features, for example, high resolution images and

time-arrangement informational collections. Points of interest are: In real applications SdA model is beneficial. SdA model contain flaw related highlights without preprocessing from sensor signals. Demerits are: Required to interrogate a trained SdA which essentially impact the classification results to find the procedure parameters.

This paper[4] contain a Novel Deep Learning based Recurrent Neural Networks(RNNs) model for short messages of programmed security review, It characterize short messages(Secure and Non-Secure). Right now short messages which is separated by word2vec and catches word request data of each sentence and mapped with highlight vector. Specifically words which contain same significance is used for comparable situation in vector space which arranged by RNNs. Merits are: The RNNs model prevail for a normal 92.7% precision which is unrivaled by the SVM. Exploiting benefit of ensemble frameworks to merge other classification algorithms to support the general execution. Drawbacks are: Apply for just short messages not for long messages.

In paper[5] a deep convolutional neural system in a cloud stage is utilized for plate confinement as Signature-based feature technique, division and character recognition. Extracting significant highlights makes the LPRS in a provoking circumstance to successfully perceive the license plate , for example, I) blocked numerous plates in the picture with traffic, ii) brightness towards plate direction, iii) plate information, iv) wear and tear distortion v) caught pictures in awful climate like cloudy pictures. Points of interest : The inclinations accuracy of proposed algorithm such as LP rather than other traditional LPRS. Impediments is: miss-detection or unrecognized pictures .

The paper [6] proposed approach of a deep learning (DL) for anomaly detection and deep belief network using a Restricted Boltzmann Machine(RBM) are achieve. To execute unsupervised feature reduction this strategy utilizes a one-concealed layer RBM. The outcome loads of this RBM is trade by another RBM. The prepared loads is initiate by fine tuning layer with multi-class soft-max contain a Logistic Regression (LR) classifier . Points of interest : Achieves 97.9% accuracy. It draft a subordinate false negative rate of 2.47%. Drawbacks : Necessity to meliorate the strategy to expand the element decrease procedure to improve the data set in the deep learning system.

The paper[7] contain Deep Learning based which access for increment adaptable and effective NIDS. A soft max regression and inadequate auto-encoder based NIDS was achieve. merits are: STL accomplished have rate over 98% accuracy for all the types of classification. Demerits are: require to give a continuous NIDS to authentic systems utilizing deep learning techniques.

The paper[8] contain mufti-core CPU's as well as GPU's to finish the performance of Deep Neural Network(DNN) based IDS to hold the huge network data. The equal processing capability of the neural system to construct the DNN to employable glance through the system traffic with quicker execution. Points of interest are: DNN based IDS is legitimate and viable in interruption recognition to detect predefined attacks classes for preparing with required number of tests. The multicore CPU's was accelerated than the sequential preparing component. Drawback are: Requirement to meliorate the recognition exactness of DNN based IDS.

The paper[9] having a system for identifying network attacks in large scale utilizing Replicator Neural Networks(RNNs) for building abnormality identification models. Our entrance is unaided and without labeled information. It follow network anomalies accurately without assuming that the preparation information is totally free of attacks. Focal points are: It is proposed technique have ability to investigate effectively all apparent SYN Port scans injected DDoS assaults. Drawbacks are: prerequisite to meliorate proposed approach by utilizing stacked auto encoder deep learning techniques.

The Paper[10] based on flow based nature of SDN, this paper contain a flow-based anomaly recognition system framework utilizing deep learning. For flow-based anomaly detection SDN contain deep learning approach. Points of interest are: It finds an ideal hyper-parameter for DNN and concurrences the identification rate false alarm rate. The model contain the accuracy and execution with 75.75% which is sensible utilizing six essential system highlights. Drawbacks are: It won't work on real SDN condition.

3 SYSTEM OVERVIEW

In this paper the model proposes a blend shallow learning and deep learning techniques, prepared to do accurately for examine a wide scope of network traffic. A novel deep learning model is used to empower Network Intrusion Detection System(NIDS) execution operation inside modern networks. More over, this paper consolidate the intensity of stacking our proposed Non-symmetric Deep Auto-Encoder (NDAE), exactness and speed of Random Forest (RF). This paper draft our NDAE, which is an auto-encoder highlighting non-balanced different shrouded layers. solo component separated that scale NDAE which can be utilized as a various leveled to suit high-dimensional information sources. Stacking the NDAE offer a layer wise unsupervised representation learning algorithm, which will allow a model to learn complex relationships between different features. It contain feature extraction capabilities so it can refine the model by organizing the most engaging features.

Fig. 1 shows the proposed system architecture of Network Intrusion Detection and Prevention System(NIDPS). The input traffic data is uses for NSL KDD data-set with 41 features. The training data-set contains data preprocessing which includes two steps: Data transformation and data normalization. After uses two NDAEs arranged in a stack, which uses for selecting number of features. After that apply the Random Forest Classifier for attack detection. Intrusion Intrusion Prevention Systems (IPS) contains IDS usefulness however increasingly complex frameworks which are equipped for making quick move so as to prevent or diminish the malicious conduct.

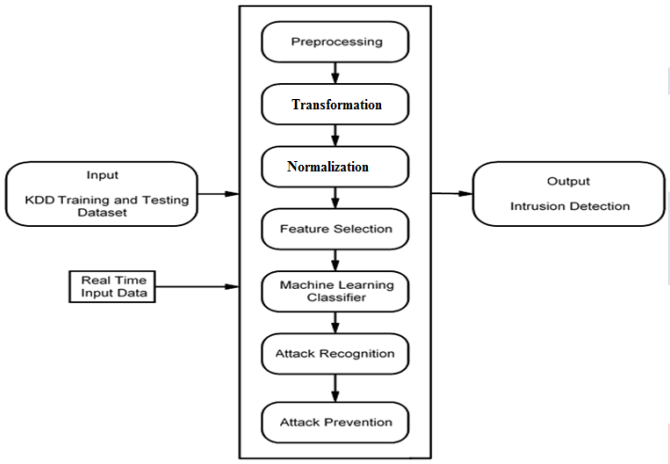

Fig. 1 Proposed System Architecture

Advantages are:

- deep learning method, meliorate accuracy of intrusion identification system(IDS).
- The network is consecutively examine for any intrusion or assault.
- The system can be modified by the need of explicit client and help external just as inward threads of the network and system.
- It effectively restrictions any harm to the system.
- It contribute easy to use interface which permit helpful security the board management system.
- Any changes to the documents and catalogs in system can be effectively trace and reported.

4 CONCLUSION

In this paper, we disputed the issues of existing NIDS procedures. Unsupervised feature learning proposed the Random Forest(RF) classification algorithm and novel NDAE approach. Additionally we accomplish the Intrusion prevention system. The inferences shows that this methodology offers superior level of precision, accuracy and recall together with diminished training time. In proposed NIDS framework is improved by 5% accuracy. Further work on real-time network traffic for improvement to assess and extend the capability of our model to handle zero-day attacks. Moreover, we will also looking forward to extend existing evaluations by utilizing real-world back bone network traffic to established the merits of extended models.

COMPARISON TABLE

Method	Accuracy	TP	FN	TN	FP
Deep Learning	93|68%	2334	166	2350	150
*CompTrav*_Graph+SVM (CG_SVM)	88.24%	2181	319	2231	269
*CompTrav*_Graph+ANN (CG_ANN)	87.88%	2190	310	2204	296
*CompTrav*_Graph+NB (CG_NB)	77.94%	1942	558	1955	545
*CompTrav*_Graph+DT (CG_DT)	87.42%	2185	315	2186	314

Table: Comparison Table Between Different Models

GRAPH

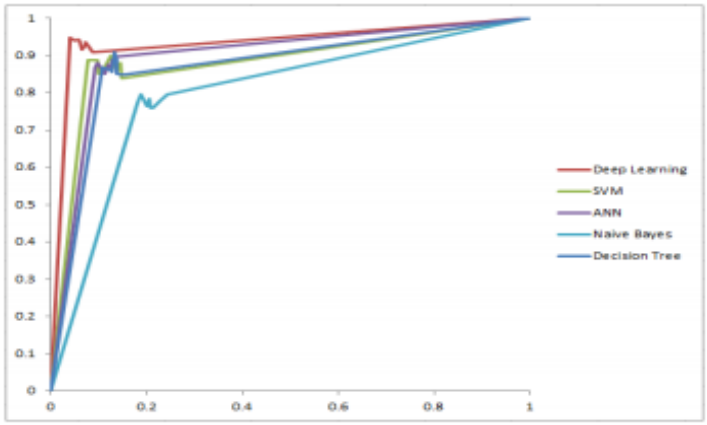

Fig 2: ROC Curves of Different Classification Model

5 REFERENCES

[1] B.Dong, et al comparison Deep Learning Methods to Traditional methods using Intrusion Detection System" IEEE International Conference, China, Jun. 2016.

[2] R.Zhao, R.Yan, et al "deep learning and its Application to machine health monitoring system: a Survey" IEEE Transaction Neural Network System, 2016.

[3] H.Lee, et al "A deep Learning Model for robust wafer fault monitoring with sensor mesurment noise" IEEE Transaction semiconductor manufacturing, Feb. 2017.

[4] Y. Li, et al "A deep learning based RNN model for automatic security audit for short messages" Communication Inf. Technology Qingdao, China, Sep. 2016.

[5] R. Pollishetty and et al "A next generation secure cloud based deep learning license plate recognition for smart cities" IEEE International Conference Machine Learning Application,USA, Dec.2016.

[6] K. Alrawashdeh and C. Purdy, "Toward an online anomaly intrusion detection system based on deep learning," IEEE International Conference Machine Learning Application Anaheim, CA. 2016.

[7] A. Javaid, et al "A deep learning approach for network intrusion detection system," International Conference, Bio-Inspired Technology, 2016.

[8] S. Potluri and C. Diedrich, "Accelerated Deep neural network for enhanced intrusion detection system," IEEE 21st International Conference Emergency Technology Factory, Germany, Sep. 2016.

[9] C. Garcia Cordero, S. Hauke, et al "Analyzing Flow Based Anomaly Intrusion Detection Using Replicator Neural Networks" Annual Conference Security Trust, New Zeland, Dec. 2016.

[10] T. A. Tang, L. Mhamdi, et al "Deep learning approach for network intrusion detection in software defined networking," International Conference Wireless Network Mobile Communication, Oct. 2016.

[11] Y. Aung and M. M. Min, "An analysis of random forest algorithm based network intrusion detection system," in *Proc. 18th IEEE/ACIS Int. Conf. Softw. Eng., Artif. Intell., Netw. Parallel/Distrib. Comput.*, Jun. 2017,

[12] M. Anbar, R. Abdullah, I. H. Hasbullah, Y. W. Chong, and O. E. Elejla, "Comparative performance analysis of classifification algorithms for intrusion detection system," in *Proc. 14th Annu. Conf.* Dec. 2016.

[13] Choudhury and A. Bhowal, "Comparative analysis of machine learning algorithms along with classifiers for network intrusion detection," in *Proc. Int. Conf. Smart Technol. Manage. Comput., Commun., Controls, Energy Mater.*, May 2015,

[14] Q. Niyaz, W. Sun, and A. Y. Javaid, A deep learning based DDOS detection system in software-defifined networking (SDN), Submitted to EAI Endorsed Transactions on Security and Safety, In Press, 2017,

[15] E. Hodo, X. J. A. Bellekens, A. Hamilton, C. Tachtatzis, and R. C. Atkinson, Shallow and deep networks intrusion detection system: A taxonomy and survey, Submitted to ACM Survey, 2017.