International Journal of Computer Science & Communication (ISSN: 0973-7391)
Special Issue, Page 35-39, December 2020 ------ available at www.csjournals.com

A Survey on Intrusion Detection System Using Deep Learning Approach

Raksha R[1]
III Sem M.Tech
Dept of CS&E
SSIT, Tumkur

Dr. Sujatha S R[2]
Assoc. Professor
Dept of CS&E
SSIT, Tumkur

Dr. M Siddappa[3]
Professor and HOD
Dept of CS&E
SSIT, Tumkur

Abstract: The timely and automatic recognition and characterization of cyber threats at the source and destination levels, deep learning technologies are frequently utilized in construction of an Intrusion Detection System (IDS). However they are continually emerging due to malicious attacks and exist in very large amounts that require a scalable response, several problems arise. For further study by the data security community, numerous malware datasets are publicly accessible.

The thorough Analysis of the results on numerous publicly available databases of various machine learning algorithms has however, not been seen by a current report. The publicly available malware datasets are to be systematically updated and benchmarked Owing to the complex nature of malware, attack tactics are continually evolving. A deep neural network in this paper (DNN), Big Data based Hierarchical Deep Learning System(BDHDLS) and a form of Deep Learning(DL) model is being investigated for creation of scalable & efficient IDS in order for identify & recognize unexpected and unforeseen cyber attacks. The changing environment in network activity and rapid attack evolution makes it essential for evaluate various databases generated by simple and complex approaches all these years. The above kind for analysis makes it possible to find the right algorithm that can work reliably to identify potential cyber threats.

Keywords: Deep neural networks (DNN), Intrusion Detection Systems (IDS), NSL-KDD, NIDS, HIDS, BDHDLS, Deep Learning(DL), Neuron, Internet of Things(IoT).

I. INTRODUCTION

Applications and networks for information and communication technology (ICT) handle multiple sensitive user data exposed to multiple internal and external intrusion attacks. Such assaults sometimes could be mechanical, device-generated, & development for obfuscation can be diverse and progressively leading to undetected data breaches. Because enormous quantity of information will processed everyday as well as the interdependency on the internet technology of the world [1], Every month, safety experts notice more and more security flaws in these infrastructures.

This technology flaws provide cyber criminals with ways to encroach through certain utilities and conduct disruptive operations[2]. Vulnerability management professionals & programmers also need to build multiple IDS to safeguard systems & software against perpetrators who can target the internet infrastructure & which could damage economical, pharmaceutical, or some valuable information in metadata. The wide- use of the IoT definition, large quantities of interconnected sensor devices, comprising including computers, automated devices, smart devices, detectors & so on.[3] Because of immense network scale and uncontrolled/anonymous internet structure, the security of the company's data and correspondence has emerged as a daunting problem for researchers[4].

Although firewalls are used for this prevention in most systems, Recognized as the second line of protection, intrusion detection systems (IDSs) acts as key towards increasing level for device stability. Intruders have been constantly attempting for discovering several evade circumvent networks risky prevalent. IDS are becoming an inevitable part of protection schemes[5].

The transformation in strategies for detecting intrusions is inspired for certain basic information, such as:

- Networked applications are so complex and very vulnerable to errors, and intruders/hackers can exploit these mistakes.
- There are some important security vulnerabilities in current network systems that Place them as a priority for hackers. While there are several instruments and works that aims to identify and repair these shortcomings, it is mostly not possible to close all of them. While there are several preventive systems for intrusion, full prevention is not feasible. As a result, IDS has emerged as an outstanding method for recording and trapping [6].
- Recognize the intrusions. A preventive mechanism can be modified automatically after this stage. Most prevention mechanisms protect the device from external criminals. Nowadays, many of those assaults were still performed along registered users for business, who are difficult to classify, This style of attack can be more harmful too.
- To overcome these prevention and detection mechanisms, new types of attack are developed. Security solutions can also be updated to a complex structure by using some learning or upgrading process. IDSs, as shown in Figure. 1, typically contain three key components[7].
- First the network flows can be tracked by a data collection system, then it is important to use these data to describe the functions, and a function vector.

- Eventually, a categorization model is carried out for the use of this vector also the tracked flows may be classified for natural either as an intervention as per existing experience[8]:

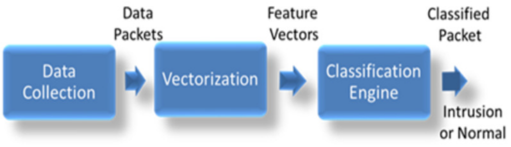

Figure. 1: Intrusion Detection System modules

Comparatively, automatic feature extraction is conducted by deep learning algorithms, allowing researchers to remove biased aspects for limited entity awareness & social effort. Such strategies provide an interface for layered statistical modeling where it is possible to extract elevated functionality over the final stage from the networks, whereas minimal data are obtained by the lowest stages. Artificial Intelligence (AI) originally inspired these types of architectures, which replicated the operation of main sensory modality throughout neural networks[9]. The description of information from disparate scenes can be extracted spontaneously by individual neurons.

The source is the information that the scenario receives through the eyelids, since this labeled entities are the outcome. It illustrates most strength of DL, that is what it functions in neural networks. DL is currently the leading avenues of study in the Machine Learning technologies, with tremendous progress in many fields. This research offers an summary of DL by diverse viewpoints, involving experience, complexities, possibilities, algorithms, mechanisms, implementations, concurrent and decentralized strategies techniques [10].

OBJECTIVE:

The objective of this paper includes by using deep learning approaches which has been used for intrusion detection system(IDS) by using different form of approaches like neural networks, big data and NSL-KDD technique.

RELATED WORK:

Although DL can be identified as major innovative area, whole report continuous to bring a big overall view and share technical skill with through intrusion detection system using various techniques.
There are various techniques related to the control of cyber criminals:

1. Work on safety concerns linked towards NIDS and HIDS has persisted from begging of electronic devices. Within current era, the implementation of ML deployment of NIDS and HIDS has been of primary concern between cyber-security experts and professionals. In order to improve NIDS and HIDS, ML and DL techniques often used[11].

- Network-Based Intrusion Detection Systems (NIDS): primarily use either statistical measures or computed thresholds on feature sets such as packet length, inter-arrival time, flow size and other network traffic parameters to effectively model them within a specific time window.
- NIDS is having some of the drawback such as They suffer from high rate of false positive and false negative alerts, A high rate of false negative alerts indicates that the NIDS could fail to detect attacks more frequently, and a high rate of false positive alerts means the NIDS could unnecessarily alert when no attack is actually taking place.
- Host-Based Intrusion Detection System(HIDS): The requisite framework for the study and collection of knowledge from target device vulnerabilities is given by various development packages such as Metasploit, Sqlmap, Nmap, Browser Exploitation[13].

- Such information is used by malicious attackers to launch attacks on different applications, such as FTP servers, web servers, SSH servers, etc. Existing techniques such as firewalls, methods of cryptography and authentication are aimed at defending against certain threats by service providers. This methods include limitations, since unauthorized users can be acquire inappropriate approach for computer[14].

II. LITRATURE SURVEY

By using the above methods of security there are different types of issues in it and it can also be implemented by the latest algorithms and technique of deep learning methods are:

a) NIDS implementation method:

For their performance benchmarking, the NSL-KDD dataset. Accordingly, any dataset alluded to from this point forward should be considered as NSL-KDD. This approach makes it possible to more reliably compare work with those found in literature. Another downside is that for both preparation and study, most employees use training results. Finally, we address a few methods focused on deep learning that have been tried so far with related job styles [15].

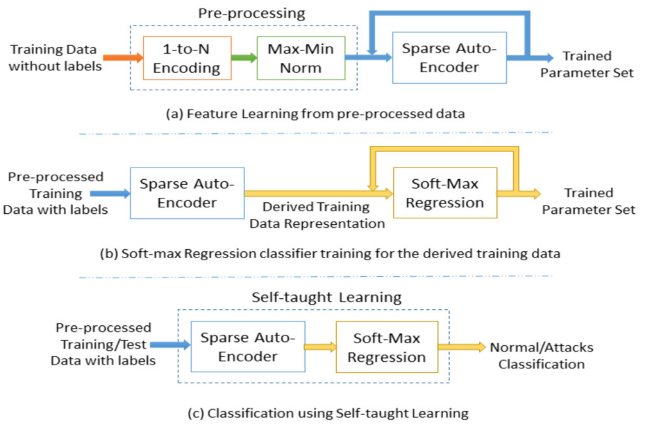
Pre-processing
Training Data without labels
1-to-N Encoding
Max-Min Norm
Sparse Auto-Encoder
Trained Parameter Set
(a) Feature Learning from pre-processed data
Pre-processed Training Data with labels
Sparse Auto-Encoder
Derived Training Data Representation
Soft-Max Regression
Trained Parameter Set
(b) Soft-max Regression classifier training for the derived training data
Self-taught Learning
Pre-processed Training/Test Data with labels
Sparse Auto-Encoder
Soft-Max Regression
Normal/Attacks Classification
(c) Classification using Self-taught Learning

Figure. 2: Various stages involved in our NIDS implementation.

Figure. 2 shows different stages involved in NIDS implementation with the new attributes, We use the NSL-KDD training data without labels for the first stage of self-taught learning for feature learning using coarse auto encoders. In the second step, we add the newly learned feature representation to the training data itself for the classification using soft-max. Retrospection. In our implementation, all the unlabeled and classified data for feature learning and classifier training come from the same source, i.e. training data from NSL-KDD [16].

III. BDHDLS phase

The first approach is to use the BDHDLS: It is a method that uses topographical map design to coordinate several DL models. Where the design of cluster is built separate observations through widely varying categories where every DL method is trained for comparison on each category to research special information dissemination with this category on similar samples[17].
Figure. 3 shows BDHDLS construction consists of five stages[18]:

- In the first step, the behavioral function and content function are retrieved & choose by using metadata process.
- In the 2nd stage, the entire data set segmented into one-level clusters using a concurrent to K-means cluster process, where the samples in each tree cluster typically exhibit identical traffic patterns.
- A hierarchical clustering procedure is conducted through the third stage to create the sub-tree of the cluster in every one-level cluster with lesser constraint conjunction.
- At fourth step, in each cluster, a DL model are design in the topographical map to learn distinctive data distribution patterns in category.
- Deep learning models' decision values are combined in the final step to allow the final judgment whether or not the sample is intrusive.

IV. DNN layers

The other approach uses the concept of the deep learning algorithm to network activity; data preparation, model definition and compilation, model fitting and model validation are the steps to create an effective deep learning approach (prediction) [19].

Figure. 4 describes Neurons : the building blocks of artificial neurons in neural networks. These are simple computing units that use an activation mechanism to produce an output signal and have weighted input signals[20]. Neuron has a bias and can be viewed as a feedback that also has a 1.0 value and must often get measured. Example, neural still had binary values, but in either three values are required[21].

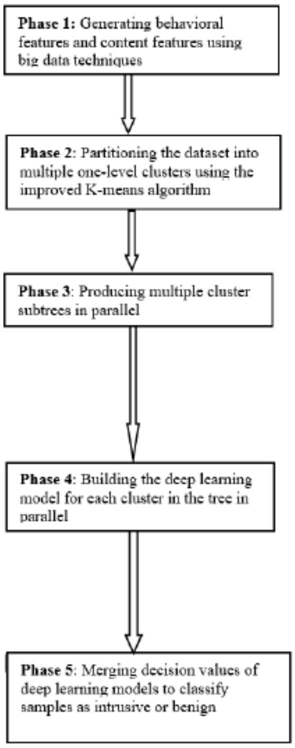
Phase 1: Generating behavioral features and content features using big data techniques
Phase 2: Partitioning the dataset into multiple one-level clusters using the improved K-means algorithm
Phase 3: Producing multiple cluster subtrees in parallel
Phase 4: Building the deep learning model for each cluster in the tree in parallel
Phase 5: Merging decision values of deep learning models to classify samples as intrusive or benign

Figure. 3: Different phase of hierarchical deep learning phases.

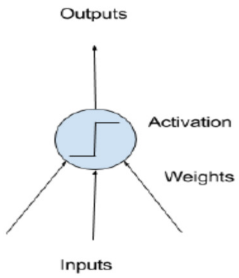

Figure. 4: Model of simple neuron

From every feedback include 1 with prejudices. Tiny unpredictable variables, along variables inside range of 0 to 0.3, are frequently initialized into weights, although more complicated initialization schemes can be used[22].

Figure. 5 shows different layers of DNN [23]:

a) Input Layer :

Since it's the revealed artificial neural portion, the input or observable layer is called the very first layer that receives data from any database. An input layer with one neuron per value of each input in the dataset also characterizes a neural network.

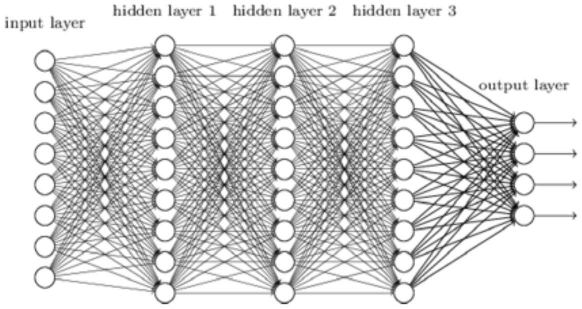

Figure. 5: Deep neural network with five layers

b) Hidden Layer:

The hidden layers are called hidden after the input layer, since they're never immediately open to feedback directly. The simple e.g. of a neural networks creating a neuron in this layer which gives value as output. The growth is the strength of computing & more powerful collection. You can create really DNN. There may several hidden layers in the neural network.

c) Output Layer :

Final layer is referred to as the output layer which is capable of transmitting values that meet the format needed for the issue to the variable or vector.

V. CONCLUSION

In this work, by introducing various types of methods such as big data, NSL-KDD and neural networks, the notion of deep learning in the field of intrusion detection system where it can be used for the help of cyber security which can be implemented using different types of implementation method it can be used for various field which can be implemented in industries and also for different places also with the suitable algorithm like and it can also be improved according to the requirements.

REFERENCES

[1] Homeland Security Council, National strategy for homeland security, https://www.dhs.gov/xlibrary/assets/nat strat home and security 2007.pdf, 2007.

[2] S. Dua and X Du, Data Mining and Machine Learning in Cybersecurity. Boston, MA, USA: Auerbach Publications,2011.

[3] B. Mukherjee, L. T. Heberlein, and K. N. Levitt, ``Network intrusion detection,'' *IEEE Netw.*, vol. 8, no. 3, pp. 26_41, May 1994.

[4] D. Larson, ``Distributed denial of service attacks_holding back the flood,'' *Netw. Secur.*, vol. 2016, no. 3, pp. 5_7, 2016.

[5] R. C. Staudemeyer, ``Applying long short-term memory recurrent neural networks to intrusion detection,'' *South Afr. Comput. J.*, vol. 56, no. 1, pp. 136_154, 2015.

[6] S. Venkatraman and M. Alazab, ``Use of data visualisation for zero-day Malware detection,'' *Secur. Commun. Netw.*, vol. 2018, Dec. 2018, Art. no. 1728303. [Online]. Available: https://doi.org/10.1155/2018/1728303.

[7] P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, ``A detailed investigation and analysis of using machine learning techniques for intrusion detection,'' *IEEE Commun. Surveys Tuts.*, to be published.

[8] deep learning approach for intelligent intrusion detection system r. vinayakumar, mamoun alazab, k. p. soman, prabaharan poornachandran, ameer al-nemrat, and sitalakshmi venkatraman, April 2019.

[9] Deep Learning in Intrusion Detection Systems Gozde Karatas Mathematics and Computer Sciences Dept. Istanbul Kultur University Istanbul, Turkey g.karatas@iku.edu.tr, Onder Demir Computer Engineering Department Marmara University, Technology Faculty Istanbul, Turkey odemir@marmara.edu.tr, Ozgur Koray Sahingoz Computer Engineering Department Istanbul Kultur University Istanbul, Turkey sahingoz@gmail.com, December 2018.

[10] A. H. Ngu, M. Gutierrez, V. Metsis, S. Nepal, and Q. Z. Sheng, "Iot middleware: A survey on issues and enabling technologies," IEEE Internet of Things Journal, vol. 4, no. 1, pp. 1–20, Feb 2017.

[11] W. Fu, X. Xin, P. Guo, and Z. Zhou, "A practical intrusion detection system for internet of vehicles," China Communications, vol. 13, no. 10,pp. 263–275, Oct 2016.

[12] Applying Big Data Based Deep Learning System to Intrusion Detection Wei Zhong_, Ning Yu, and Chunyu Ai BIG DATA MINING AND ANALYTICS ISSN 2096-0654 03/06 pp181–195 Volume 3, Number 3, September 2020, DOI: 10.26599/BDMA.2020.902000.

[13] A Deep Learning Approach for Network Intrusion Detection System Quamar Niyaz, Weiqing Sun, Ahmad Y Javaid, and Mansoor Alam College Of Engineering The University of Toledo Toledo, OH-43606, USA {quamar.niyaz, weiqing.sun, ahmad.javaid,

mansoor.alam2}@utoledo.edu, BICT 2015, December 03-05, New York City, United States Copyright © 2016 ICST DOI 10.4108/eai.3-12-2015.2262516.

[14] A Deep Learning Approach for Intrusion Detection System in Industry Network Ahmad HIJAZI Univ.Grenoble Alpes, G-SCOP, F-38000 Grenoble, France ahd.hjz@gmail.com, EL Abed EL SAFADI Univ.Grenoble Alpes, G-SCOP, F-38000 Grenoble, France Abed.safadi@grenoble-inp.fr, Jean-Marie FLAUS Univ.Grenoble Alpes, G-SCOP, F-38000 Grenoble, France Jean-marie.Flaus@grenoble-inp.fr.

[15] J. Schmidhuber, Deep learning in neural networks: An overview, Neural Networks, vol. 61, pp. 85–117, 2015.

[16] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, Deep learning approach for intelligent intrusion detection system, IEEE Access, vol. 7, pp. 41525–41550, 2019.

[17] S. M. Kasongo and Y. X. Sun, A deep learning method with filter based feature engineering for wireless intrusion detection system, IEEE Access, vol. 7, pp. 38597–38607, 2019.

[18] P. Nagar, H. K. Menaria, and M. Tiwari, Novel approach of intrusion detection classification deep learning using SVM, presented at First International Conference on Sustainable Technologies for Computational Intelligence, Singapore, 2020, pp. 365–381.

[19] M. Akter, G. D. Dip, M. S. Mira, M. A. Hamid, and M. Mridha, Construing attacks of internet of things (IoT) and a prehensile intrusion detection system for anomaly detection using deep learning approach, presented at International Conference on Innovative Computing and Communications: Proceedings of ICICC 2019, Singapore, 2020, pp. 427–438.

[20] Z. Q. Liu, M. U. D. Ghulam, Y. Zhu, X. L. Yan, L. F. Wang, Z. J. Jiang, and J. C. Luo, Deep learning approach for ids, presented at Fourth International Congress on Information and Communication Technology: ICICT 2019, Singapore, 2020, pp. 471–479.

[21] C. Sekhar and K. V. Rao, A study: Machine learning and deep learning approaches for intrusion detection system, presented at Int. Conf. Computer Networks and Inventive Communication Technologies, Coimbatore, India, 2019, pp.845–849.

[22] G. Nguyen, S. Dlugolinsky, V. Tran, and A. L. Garc´ıa, Deep learning for proactive network monitoring and security protection, IEEE Access, vol. 8, pp. 19696–19716, 2020.

[23] A. Abusitta, M. Bellaiche, M. Dagenais, and T. Halabi, A deep learning approach for proactive multi-cloud cooperative intrusion detection system, Future Generation Comput. Syst., vol. 98, pp. 308–318, 2019.

[24] A. Liu and B. Sun, An intrusion detection system based on a quantitative model of interaction mode between ports, IEEE Access, vol. 7, pp. 161725–161740, 2019.

[25] T. Aldwairi, D. Perera, and M. A. Novotny, An evaluation of the performance of restricted boltzmann machines as a model for anomaly network intrusion detection, Comput. Networks, vol. 144, pp. 111–119, 2018.