IJCSE International Journal of Computer Sciences and Engineering Open Access

Review Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

Anomaly Detection and Categorization in Cloud Environment using Deep Learning Techniques

Nidhi Thakkar[1], Miren Karamta [2], Seema Joshi [3] and M. B. Potdar [4]

[1,3]GTU Cyber Security, Graduate School of Engineering & Technology, Gandhinagar 382028, Gujarat, India
[2,4]Bhaskaracharya Institute for Space Applications and Geo-Informatics, Gandhinagar 382007, India

Corresponding Author: nidhithakkar.129@gmail.com

DOI: https://doi.org/10.26438/ijcse/v7i5.211214 | Available online at: www.ijcseonline.org

Accepted: 08/May/2019, Published: 31/May/2019

Abstract— Cloud computing is a paradigm that allows on-demand network access to a shared pool of configurable and reliable computing resources to cloud customers in pay-per-use, fashion. Despite the existence of such merits, there are Security issues such as data integrity, users' confidentiality, and service availability because of its open and distributed architecture that place restrictions on the use of cloud computing. A preventive approach is to identify such issues and eliminate before it can cause the serious impact to the cloud users. Nowadays, Intrusion Detection Systems (IDSs) are the most widely used method to detect attacks on cloud. Recently, learning-based techniques for security applications are gaining popularity in the literature with the emergence in machine learning. A deep learning is a novel approach to detect cloud threats. The existing Cloud IDSs suffer from low detection accuracy and a high false positive rate. In this research, proposed solution will use deep learning algorithm to improve the effectiveness of our proposed solution. Furthermore, the comparisons with other deep learning algorithm to demonstrate the effectiveness of our proposed solution are given.

***Keywords* —** Cloud Security, Network Intrusion Detection System, Deep Learning

I. INTRODUCTION

Cloud computing is a paradigm that provides on-demand, user convenient, shared resources (network, storage, Applications, services) which could be easily accessible from cloud services provider. The on-demand and pay-as-you-go cloud characteristics are gaining more attraction of organizations and forcing them to become on-premises infrastructures to off premises data centers, and it can be accessed over the Internet and managed by cloud service providers [1].

Despite the existence of such merits, there are security issues such as data integrity, users' confidentiality, and service availability because of its distributed and open architecture it is more vulnerable to attackers. CSPs are also more susceptible to attackers who might exploit to take advantages of the vulnerabilities. Encryption is the widely used key to resolve the confidentiality and integrity however, to adopt cloud technology widely it is much more require to establish a relationship between CSP and cloud users[1][2].

To resolve the cyber attack in cloud environment, it is hard to detect the threats and countermeasures to mitigate the risk. Nowadays, some prevention and detection techniques are proposed for cloud environment. For Example In [2], [3] and [5] author has suggested some solution to detect the cyber threats in cloud environment. Alternatively, there are another ways to detect and prevent attacks game theory and supervised learning. However, the main disadvantage of these methods is low accuracy and they are not as efficient as the existing real life detection system.

In this paper, we proposed advanced deep learning technique to detect different cyber threats with high accuracy. Deep learning is a part of machine learning having structure of algorithm with neural networks. In many areas Deep learning has been implemented effectively over the past few years. For example, Deep learning is used in speech recognition, automatic translation, to identify in what customers are interested to buy and also in image identification [7][13].

In this paper we have introduced detection and classification deep learning technique. The main objective of this technique is to build the model as per defined requirement using training dataset and adjusting the weights of neural network. After that the neurons will be used to detect and classify the different cyber attacks with high accuracy. In this

paper we proposed combination of deep neural network and deep belief network to achieve high accuracy.
The remaining of the paper is organized as follows: Segment II of this paper describes Literature Review. Segment III discusses about deep learning. Segment IV explain basics about proposed system of this paper. Segment V discusses about Performance Measures and Experimental Results. Conclusion and future work are drawn in segment VI.

II. LITERATURE REVIEW

Based on data Mining strategies Distributed IDS for Cloud architecture is given in the paper. Different Machine Learning Techniques for Intrusion Detection and its Comparative Analysis is given in [3]. In [5] A Study on NSL-KDD Dataset for Intrusion Detection System Based on Classification Algorithms was evaluated. The proposed model in [6] is used to identify the dispensed intrusions with the use of mobile dealers however; this explanation isn't always focused for cloud computing architecture. Performance Evaluation of Supervised Machine Learning Algorithms for Intrusion Detection is done in [7]. Adaptive Hybrid method for Network Intrusion Detection and Comparison among Machine Learning Algorithms is mentioned in [14]. Deep Learning based attack detection for Data Security and a deep learning approach to network intrusion detection explained in [9], [13], [24].

III. DEEP LEARNING

Deep Learning is an aspect of machine learning that is concern with the learning approach. However machine learning algorithms are linear, deep learning algorithms are based on hierarchy of abstraction. Basically, it teaches computer to act like human [22]. Most of deep learning algorithm uses neural network, so that it is also referred as "Deep Neural Network". Deep Neural Network (DNN), Deep Boltzmann machine (DBM), Restricted Boltzmann Machine (RBM), Deep Belief Network (DBN), Generalized Denoising AutoEncoders, Recurrent Neural Network (RNN), etc. are the different Deep Learning algorithms [8][9].

Deep Neural Networks:
A deep neural network is a neural network with a certain degree of complexity, a neural network with multiple hidden layers; it has more than two layers.

Deep neural networks use mathematical equations to solve any complex problem, it use mathematical modelling approach to solve data complexity. A neural network, widespread, is a technology constructed exactly replica of our brain simulation – mainly, pattern identification and the transit of input via numerous layers of neural connections. Professionals define deep neural networks as networks which consist of an input layer, an output layer and as a minimum one hidden layer in between. Every layer plays important role in a process that a few discuss with as “feature hierarchy”.

Deep Belief Networks:
Deep Belief Networks are base on probability algorithms. This probability generator algorithm works on different latent variables and stochastic. The latent variables normally hold binary values and also called hidden units / feature detectors. First two layers have consistent connections among them and they are use to memorize the feature and characteristics [7]. Lower layer connected with upper layers they gain knowledge from upper layer and lower layers are generating data vectors.

Generally there are different Deep Learning algorithms, i.e. Deep Neural Network (DNN), Deep Belief Network, Restricted Boltzmann machine, Convolutional Neural Network, Stacked Autoencoder, But the Deep Neural Network (DNN) is most widely used for general purpose. Here, for the proposed work also DNN and DBN is used and based on that some combination of different parameters are applied and One new proposed algorithm is developed which gives highest accuracy. In proposed Algorithm Dense Layer of DNN and RBM layer of DBN is combined and also some other combinations are also applied.

IV. PROPOSED MODEL

The proposed system to detect intrusion is composed of the following components:

Network traffic collector:
The First module of our proposed system will be network traffic data collector. This module is implemented in cloud environment near to every edge of network router. This will capture the incoming network traffic of that particular cloud infrastructure. The captured data will be stored in server and it will be passed to the next module [1].

Data Pre-processing:
The data collected by the first module will be pre-processed by this module. Different methods are used to normalize and formation of the collected network traffic data. Generally this module is placed at side by each network traffic collector module. This module is used by Hadoop and MapReduce to pre-process when the network traffic increases. This whole task of Pre-processing is implemented by MapReduce[8] [19].

Anomaly detection:
After Pre-processing the data will be passed to the detection module. To detect the anomaly we have implemented a hybrid deep learning method to achieve higher accuracy using NSL-KDD dataset. As per [15] this module is designed

to be inserted side by side data Pre-processing module to increase the detection speed.

Attacks traffic classification:

Anomaly detection module will decide whether it is anomaly or normal, after that this module will classify that which type of attack is ongoing. Using this administrator can take appropriate measures to mitigate the damage. This module will classify the server storage attacks based on FIFO algorithm.

The overall architecture of proposed system is as described below.

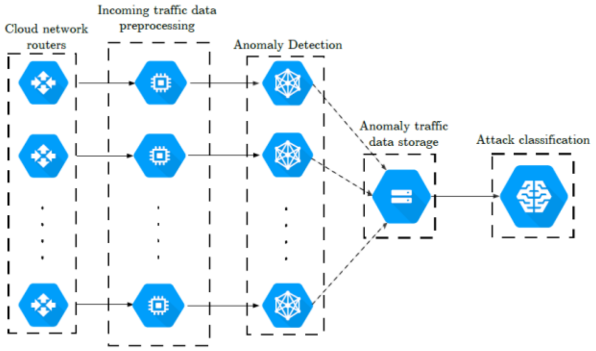
Cloud network routers
Incoming traffic data preprocessing
Anomaly Detection
Anomaly traffic data storage
Attack classification

Fig.1: Flowchart of the proposed IDS

V. Performance Measures and Experimental Results

To compare performance of different techniques or datasets, we can use multiple evaluation parameters like Precision, Recall and Accuracy. Among them Accuracy is the important parameter, in the network intrusion detection system performance evaluation [13][18]. We have considered Precision, Recall and Accuracy parameters which are described below.

where, TP = True Positive
FP = False Positive
TN = True Negative

FN = False Negative
Accuracy defines the percentage of the total quantity of accurate classifications.

Accuracy = (TP+TN)/ (TP +TN+FP+FN)

Precision is defined as the total quantity of data appropriately predicted positive(anomaly) data over the number of instance predicted as positive(anomaly) [4].

Precision = TP/ (TP+FP)

Recall defines the percentage of appropriately predicted positive (anomaly) data out of the number of actual positive (anomaly) data.

Recall = TP/ (TP+FN)

In this paper, first we have implemented Deep Neural Network (DNN), Deep Belief Network (DBN) and after that we have combined the dense layer of Deep Neural Network and RBM layer of Deep Belief Network. Also we have apply some combination of activation function, loss function, learning rate and other attributes. Based on that we have got the highest accuracy which is shown below. And based on that it will also classify the type of attack. The result of the analysis shows that proposed algorithm achieves higher accuracy as compared to existing algorithms. This will help the administrator to take appropriate actions against the attacks.

Table 1–Comparison of existing neural networks with proposed neural network

	Precision (%)	Recall (%)	Accuracy (%)
Deep Neural Network	89.362	80.769	86.381
Deep Belief Network	93.772	72.39	84.695
Proposed Algorithm	94.228	96.429	95.525

VII. CONCLUSION AND FUTURE WORK

In this paper we proposed a Network Intrusion Detection System (NIDS) to protect cloud network. The properties of network traffic that makes it to feasible to exploit the attack and gathering information from the user network. This paper proposed system architecture that makes sure Confidentiality, integrity and availability. In this paper we have implemented Different deep learning techniques and we have combined the dense layer of Deep Neural Network and RBM layer of Deep Belief Network and also applied some combination of activation function, loss function, learning rate and other attributes. And based on that it will also classify the type of attack. The result of the analysis shows that proposed algorithm achieves higher accuracy as compared to existing algorithms. This will help the administrator to take appropriate actions against the attacks.

ACKNOWLEDGMENTS

We are thankful to Shri T. P. Singh, Director, and BISAG for providing us infrastructure, motivation, and permitting to carry out this project at BISAG.

REFERENCES

[1] Mehmood, Yasir, et al. "Intrusion detection system in cloud computing: challenges and opportunities." *2013*, IEEE.

[2] Idhammad, Mohamed, Karim Afdel, and Mustapha Belouch. "Distributed intrusion detection system for cloud environments based on data mining techniques." *Procedia Computer Science* 127 (2018): 35-41.

[3] Hamid, Yasir, M. Sugumaran, and LudovicJournaux. "Machine learning techniques for intrusion detection: a comparative analysis." *Proceedings of the International Conference on Informatics and Analytics*. ACM, 2016.

[4] Haq, Nutan Farah, et al. "Application of machine learning approaches in intrusion detection system: a survey." *IJARAI-International Journal of Advanced Research in Artificial Intelligence* 4.3 (2015): 9-18.

[5] Dhanabal, L., and S. P. Shantharajah. "A study on NSL-KDD dataset for intrusion detection system based on classification algorithms." *International Journal of Advanced Research in Computer and Communication Engineering* 4.6 (2015): 446-452.

[6] Buczak, Anna L., and ErhanGuven. "A survey of data mining and machine learning methods for cyber security intrusion detection." *IEEE Communications Surveys & Tutorials* 18.2 (2016): 1153-1176.

[7] Nguyen, Khoi Khac, et al. "Cyberattack detection in mobile cloud computing: A deep learning approach." *2018 IEEE Wireless Communications and Networking Conference (WCNC).* IEEE, 2018.

[8] Van, Nguyen Thanh, Tran Ngoc Thinh, and Le Thanh Sach. "An anomaly-based network intrusion detection system using deep learning." *2017 International Conference on System Science and Engineering (ICSSE)*. IEEE, 2017.

[9] Feng, Fang, et al. "Anomaly detection in ad-hoc networks based on deep learning model: A plug and play device." *Ad Hoc Networks* 84 (2019): 82-89.

[10] Shone, Nathan, et al. "A deep learning approach to network intrusion detection." *IEEE Transactions on Emerging Topics in Computational Intelligence* 2.1 (2018): 41-50.

[11] Kwon, Donghwoon, et al. "A survey of deep learning-based network anomaly detection." *Cluster Computing* (2017): 1-13.

[12] Aldwairi, Tamer, Dilina Perera, and Mark A. Novotny. "An evaluation of the performance of Restricted Boltzmann Machines as a model for anomaly network intrusion detection." *Computer Networks* 144 (2018): 111-119.

[13] Almseidin, Mohammad, et al. "Evaluation of machine learning algorithms for intrusion detection system." *Intelligent Systems and Informatics (SISY), 2017 IEEE 15th International Symposium on*. IEEE, 2017.

[14] Haque, MdEnamul, and Talal M. Alkharobi. "Adaptive hybrid model for network intrusion detection and comparison among machine learning algorithms." *International Journal of Machine Learning and Computing* 5.1 (2015): 17.

[15] Ahmed, Mohiuddin, AbdunNaser Mahmood, and Jiankun Hu. "A survey of network anomaly detection techniques." *Journal of Network and Computer Applications* 60 (2016): 19-31.

[16] Liu, Qiang, et al. "A survey on security threats and defensive techniques of machine learning: a data driven view." *IEEE access* 6 (2018): 12103-12117.

[17] Kwon, Donghwoon, et al. "A survey of deep learning-based network anomaly detection." *Cluster Computing* (2017): 1-13.

[18] Belavagi, Manjula C., and BalachandraMuniyal. "Performance evaluation of supervised machine learning algorithms for intrusion detection." *Procedia Computer Science* 89 (2016): 117-123.

[19] Shanmugavadivu, R., and N. Nagarajan. "Network intrusion detection system using fuzzy logic." *Indian Journal of Computer Science and Engineering (IJCSE)* 2.1 (2011): 101-111.

[20] Shone, Nathan, et al. "A deep learning approach to network intrusion detection." *IEEE Transactions on Emerging Topics in Computational Intelligence* 2.1 (2018): 41-50.

[21] Park, Kinam, Youngrok Song, and Yun-Gyung Cheong. "Classification of Attack Types for Intrusion Detection Systems Using a Machine Learning Algorithm." *2018 IEEE Fourth International Conference on Big Data Computing Service and Applications (Big Data Service)*. IEEE, 2018.

[22] https://en.wikipedia.org/wiki/Deep_learning

Authors Profile

Ms. Nidhi Thakkar completed her bachelor degree in Computer Science and Engineering from Government Engineering College, Patan, Gujarat, India in the year of 2016. Now pursuing master’s degree in Computer Engineering(Cyber Security) from Gtu - Graduate School of Engineering and Technology.

Mr. Miren Karamta is working as a Project Scientist and IT Systems Manager at the Bhaskaracharya Institute for Space Applications and Geo-informatics (BISAG), Gandhinagar, Gujarat since January 2011. He has completed MTech in 2010 from Computer Engineering from DDIT.

Ms. Seema Joshi is working as Assistant Professor (Cyber Security) at GTU-Graduate School of Engineering and Technology. Gujarat Technological University, Chandkheda, Ahmedabad, Gujarat. She has worked in many Government Projects. She has published many papers in journals and conferences and two books with Pearson Education.

Dr. M. B. Potdar is a 1982 Ph. D. in Physics from Physical Research Laboratory of Dept. of Space, Govt. of India. Later for 28 years, he was associated with the Indian Space Research Organization in various capacities. He worked extensively in development of land and atmospheric applications of Remote Sensing data. Since March 2011, he is holding position as Project Director at BISAG and organizing and steering research in various areas of software development and applications of geo-spatial data.