RESEARCH ARTICLE OPEN ACCESS

Distributed Intrusion Detection System Using Blockchain and Cloud Computing Infrastructure

Mr D Purushothaman MCA., M.E.,[1], K Ashok Kumar [2]

[1] Asst. Professor, Department of Computer Applications

[2] Student, Department of Computer Applications

[1], [2] Chadalawada Ramanamma Engineering College (Autonomous)

ABSTRACT

Nearly two decades after its emergence, the Cloud Computing remains gaining traction among organizations and individual users. Many security issues arise with the transition to this computing paradigm including intrusions detection. Intrusion and attack tools have become more sophisticated defeating traditional Intrusion Detection Systems (IDS) by large amount of network traffic data and dynamic behaviors. The existing Cloud IDSs suffer form low detection accuracy, high false positive rate and high running time. In this paper we present a distributed Machine Learning based intrusion detection system for Cloud environments. The proposed system is designed to be inserted in the Cloud side by side with the edge network components of the Cloud provider. This allows to intercept incoming network traffic to the edge network routers of the physical layer. A time-based sliding window algorithm is used to preprocess the captured network traffic on each Cloud router and pass it to an anomaly detection module using Naive Bayes classifier. A set of commodity server nodes based on Hadoop and MapReduce are available for each anomaly detection module to use when the network congestion increases. For each time window, the anomaly network traffic data on each router side are synchronized to a central storage server. Next, an ensemble learning classifiers based on the Random Forest is used to perform a final multi-class classification step in order to detect the type of each attack. Various experiment are performed in the Google Cloud Platform in order to assess the proposed system using the CIDDS-001 public dataset. The obtained results are satisfactory when compared to a standard Random Forest classifier. The system achieved an average accuracy of 97%, an average false positive rate of 0.21% and an average running time of 6.23s.

Keywords: - Nearly two decades after its emergence, the Cloud Computing remains gaining traction among organizations and individual users.

I. INTRODUCTION

Cloud computing is web based computing, wherever virtual shared servers offer software, infrastructure, stage, devices and different resources and facilitating to customers as a service to pay-as you-use premise and is outlined in Fig. 1 [1]. The IDS mainly submits to the monitoring of the computer network system, collecting and analyzing the information by using its main key points, and then discovering some behaviors that are contrary to the security strategy, or finding some signs of the attack, and making the timely and automatic response [2]. Intrusion detection is equivalent to a category issue, i.e., identifying whether it is typical or affected by any of the other four attack types: Denial of Service (DOS), User to Root (U2R), Probe (Probing) and Root to Local (R2L) [3]. These days, various associations have started to transfer their monstrous measure of noticeable information into public cloud [4].

In any case, transferring delicate information to open and distributed public cloud storage services acts security dangers such like accessibility, confidentiality and trustworthiness to associations. In addition, relentless cloud services have caused high levels of interruption and abuse [5].

Network attacks are a severe concern that confronts both cloud providers and massive number of mobile users who access distance clouds in our routine activities [6], providing security for user authentication with passwords or digital certificates and privacy information transmission [1]. To overcome these issues, a few approaches were proposed up until now. A structure that uses a profound learning approach to discover cyber attacks in mobile cloud condition is proposed [7].

An approach utilizing an Improved Genetic Algorithm (IGA) to construct a Deep Neural Network (DNN) based inconsistency NIDS is proposed. Here Genetic Algorithm (GA) is improved through optimization techniques, to be

specific Parallel Processing and Fitness Value Hashing [8]. A profound learning technique for interruption detection, called non-symmetric profound auto-encoder (NDAE) for unaided feature learning is proposed. Moreover, a profound learning classification model constructed utilizing stacked NDAEs is also proposed [9].

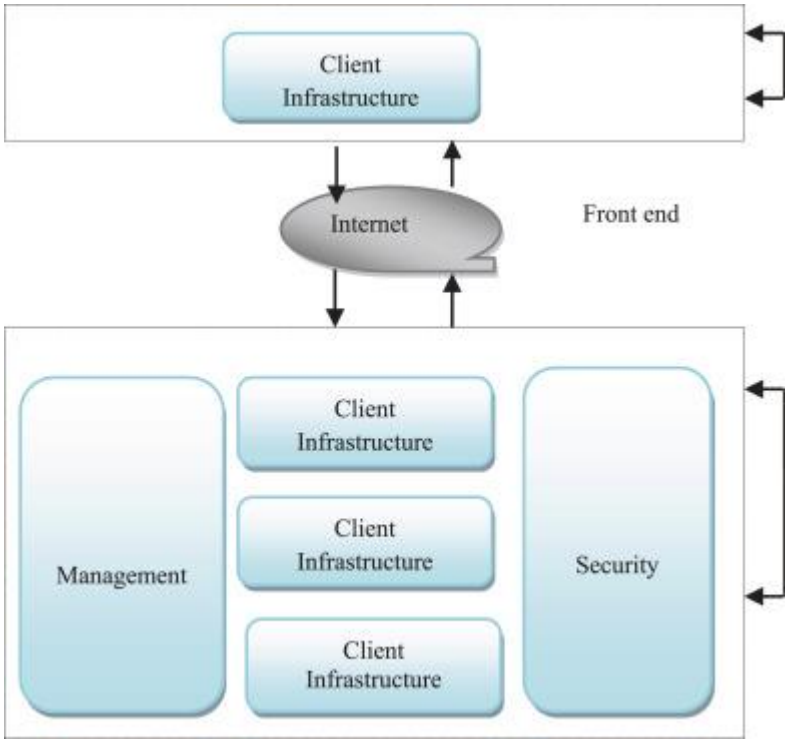

Fig. 1. Cloud computing architecture.

To a conclusion since deep learning techniques have the potential to remove better representations from the data to generate better models, and motivated by recurrent neural networks (RNN), a deep learning approach for an IDS using optimized custom RC-NN is proposed. Additionally to reduce the error rate a meta-heuristic Ant Lion optimization algorithm is proposed. The main contributions of this paper are reviewed as follows.

- The design and implementation of the detection system based on recurrent neural networks. An attack detection method known as the optimized custom RC-NN network is proposed which is hybrid from LSTM, a type of RNN and CNN. This detects the attacks in the network layer of cloud.
- Additionally Ant Lion optimization (ALO) algorithm is utilized within the proposed network layers to minimize the error rate thus improving accuracy.
- Estimating the effectiveness of the proposed solution using the DARPA dataset and CSE-CIC-IDS2018 dataset, and comparing the results with those generated with other existing approaches. The amalgamation of the proposed optimized custom RC-NN hybrid deep learning technique and meta-heuristic optimization provides better accuracy with less error rate.

The respite of this paper is sorted out as follows. In Section 2, the related research is presented that discusses some existing deep learning methods that facilitate the development of intrusion detection. An explanation of the proposed IDS architecture and the performance evaluation measures are presented in Section 3. Section 4 highlights the proposed model by a discussion about the experimental results and a comparison with a few previous studies. Finally, the conclusions and future work are discussed in Section 5 pursued by the references.

II. LITERATURE SURVEY

Some of the existing related topics for IDS with deep learning and optimization techniques in the cloud are listed and is detailed below as follows:

Iqbal et al. [10] described about cloud computing and its services via Software as a Service, Platform as a Service and Infrastructure as a Service. Nonetheless, these service delivery models are helpless against a scope of security attacks, abusing both cloud specific and existing web service vulnerabilities. Scientific classifications are a valuable apparatus for system designers as they give a systematic method for comprehension, identifying and tending to security dangers. Right now, Cloud based attacks and vulnerabilities were collected and classified with respect to their cloud models. Moreover scientific classification of cloud security attacks and potential relief methodologies was given the point of giving an inside and out comprehension of security requirements in the cloud condition.

Deshpande et al. [11] reported a host based IDS model for Cloud computing condition. This model alarms the Cloud client against the malicious activities within the system by breaking down the system call traces. The method analyzed just discriminating system call traces, the bombed system call trace, as opposed to all. An untimely recognition of interruptions with reduced computational weight can be conceivable with this feature. The announced model gave security as a service (SecaaS) in the infrastructure layer of the Cloud condition. This work lacked an adaptive management module for initiating preventive actions after intrusion detection and the integration of Host stationed Intrusion Detection System (HIDS) and network based IDS (NIDS).

Loukas et al. [12] discussed about detection of cyber attacks against vehicles. As vehicles characteristically manage the cost of constrained processing resources, proposed resolution is rule-based or lightweight machine learning practices. A profound learning approach is shown that achieved high accuracy much more consistently than with typical machine learning practices and is not constrained to a solitary attack or the in-vehicle CAN bus as past work. As info, it utilizes information captured progressively that identify with both cyber and physical processes, which it takes care of as time arrangement information to neural network architecture. Both a profound multilayer perceptron and recurrent neural network architecture were used, with the last profiting by a long-transient memory concealed layer, which demonstrates extremely helpful for learning the fleeting context of diverse attacks. Disavowal of service, command injection and malware are utilized as instances of cyber attacks that are important for a robotic vehicle.

Abusitta et al. [13] discussed the capacity of cooperative cloud-based IDS in noticing complicated and obscure attacks connected with the composite architecture of the Cloud. In a cooperative setting, IDS consults different IDSs about suspicious interruptions and settle on a decision utilizing a conglomeration algorithm. In any case, undesired delays emerge from applying total algorithms and furthermore from standing by to receive feedback from consulting IDSs. These constraints render the decisions produced by existing cooperative IDS approaches ineffective continuously. To face these challenges, a machine learning-based cooperative IDS was suggested that misused the historical feedback information to give the capacity of proactive decision making. Specifically, the proposed model depended on a Denoising Autoencoder (DA), which was utilized as a structure block to construct a profound neural network. The intensity of DA lies in its capability to figure out how to reconstruct IDSs' feedback from halfway feedback. This permits to proactively settling on decisions about suspicious interruptions even in the absence of complete feedback from the IDSs.

Gai et al. [14] discussed about mobile cloud computing that is applied in numerous ventures to get cloud-based services by utilizing mobile technologies. With the improvement of the wireless networks, safeguarding dangers from wireless communications assumed an amazing job in the Web security area. IDS is an approach for defending wireless communications in the Fifth Generation (5G) context. Right now, fundamental techniques executed in IDS and mobile cloud computing are abridged. Tending to the security concern, a higher level structure of executing secure mobile cloud computing was proposed by receiving IDS practices for pertaining mobile cloud-based solutions in 5G networks.

Chiba et al. [15] illustrated about Cloud Computing applications in numerous sectors like industry, governments, education and entertainment. Nevertheless, transferring delicate information to public cloud storage services acts security dangers such like uprightness, accessibility and confidentiality to associations. Additionally, the open and

distributed structure of the cloud results in this class of computing, inclined to cyber attackers and gatecrashers. Right now, approach to automatically construct a Deep Neural Network (DNN) based peculiarity Network IDS utilizing a crossover optimization system (IGASAA) in light of IGA and Simulated Annealing Algorithm (SAA) was proposed. The IDS resulted was called "MLIDS" (Machine Learning based IDS). GA improves through optimization procedures, in particular Parallel Processing and Fitness Value Hashing, which reduces execution time, convergence time and spare processing power.

Hajimirzaei et al. [10] proposed an IDS in view of a combined multilayer perceptron (MLP) network, and artificial honey bee colony (ABC) and fuzzy clustering algorithms. Ordinary and anomalous network traffic packets were identified by the MLP, while the MLP preparing was finished by the ABC algorithm through streamlining the estimations of linkage loads and predispositions.

Alkadi et al. [11] introduced a Deep Blockchain Framework (DBF) designed to present security-rooted distributed IDS and privacy-rooted blockchain with smart contracts in IoT networks. The IDS was engaged by a Bidirectional LSTM algorithm to work with sequential network data. Hachimi et al. [18] focused on organizing a multi-stage machine learning-based IDS (ML-IDS) that identifies and classifies four sorts of jamming attacks.

III. IDS MODEL WITH DEEP LEARNING AND OPTIMIZATION

Cloud computing becomes a storage medium for several sectors leading to storage of large amount of packets. This tempts to various issues where security issues like integrity, availability and confidentiality, cyber attackers and intruders are considered as a major one. The cloud computing environment comprises of several security issues like virtualization level security issues, application level security issues, network level security issues, physical level security issues, etc. Among these the proposed works considered only the detection of network level security issues in the cloud environment. This happens since cloud services are delivered through the web browser. The network level attacks or intrusion include denial of service, User to root, Remote to User and probe. Thus, in this work IDS combining deep learning neural network technique with Ant lion optimization algorithm is proposed.

Initially the attack classification is done with the aid of neural network technique LSTM network, which is a type of RNN along with CNN and Ant lion optimization (ALO) algorithm. Since large amounts of packets give high error rate in the deep neural network, for training, the meta-heuristic ALO algorithm is utilized for optimizing this process thus providing less error rate and high classification rate. During new data arrival the IDS checks for intrusions or attacks and other policy infringement and finally provide a response in terms of attack classification. The proposed mechanism is demonstrated pictorially in Fig. 2.

The proposed IDS is proposed and evaluated with the aid of DARPA dataset and CSE-CIC-IDS2018 dataset. These datasets comprise of both mathematical and textual contents. To work with this text data initially preprocessing is done where the lowercase texts are converted and the punctuations are erased. Alternatively the document is padded and truncated to make them of the same length. Before providing the input to the optimized custom RC-NN network, feature analysis is done to extract features. There are 43 features in the DARPA dataset like protocol type, service, srcbytes, dstbytes, attack, etc. Specifically, these data packets enclose numerous attributes with diverse features. However, not all the 43 features are functional for IDS. Several features are unrelated and superfluous ensuing in an extended detection process and corrupting the performance. A few infrequent data in the dataset is removed as unused data and is sorted into dataset of reasonable size. Thus, choosing features which protect the most useful data of a data set is necessary to diminish the computation complexity and increase the accuracy of the learning process.

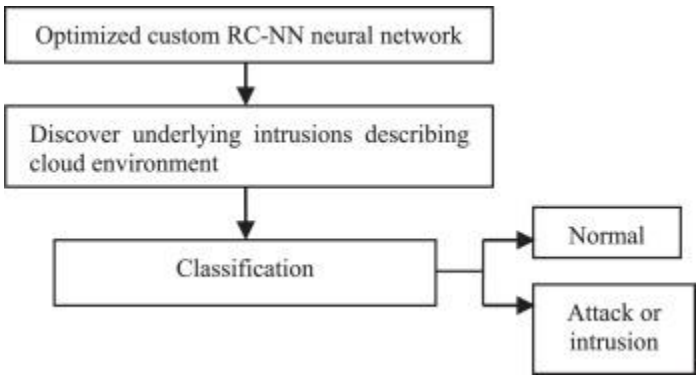

Fig. 2. Flow diagram for proposed methodology.

IV. PROPOSED SYSTEM:

In a large network, multiple Network Intrusion Detection Systems are deployed across the network.Theses distributed IDS share the logs and alert information with each other. Such an arrangement ofmultiple IDS is called a Distributed Intrusion Detection System (DIDS). The type and volume of information shared among the distributed IDS is configured by the administrator and need to be fine-tuned from time to time. It facilitates advanced persistent threat analysis, network monitoring, and instant attack analysis of the whole network.

Advantages

With the growing size of network infrastructure, scalability and performance is always the major concern for single-mode IDS. It's difficult for the single-mode IDS to detect the attack pattern scattered across different geographical locations of an enterprise network. DIDS hasan advantage over single-mode IDS to collect and corroborate data among the peer IDS and detect the stealthy attack pattern. Many times in the case of Advanced Persistent Attack, it is observed that the attack may be initiated in some specific region of the network and then slowly spread to the entire network. In DIDS since all the IDS are connected, any attack detected in aspecific region or segment of the network can help the other IDS to learn and update their rule-base. The administrator can take preventive measures and protect the rest of the network from attack.

V. SIMULATION RESULTS AND DISCUSSIONS

In this research the hybrid deep learning network, i.e. optimized custom RC-NN network is utilized along with the meta-heuristic ALO algorithm. This algorithm efficiently classifies the intrusions recognized within the cloud network environment. The experimental analysis is performed in windows 7, 64 bit operating system with 4 GB memory. Additionally, the experimental results are executed in Matlab 2019a employing the DARPA dataset [4] and CSE-CIC-IDS2018 dataset [2].

Dataset description

The DARPA dataset uses reduced attributes with diverse threshold limits, 22 features preferred which give almost the same accuracy and sensitivity as the 41 attribute dataset. The different attack types identified in the DARPA dataset.

The CSE-CIC-IDS2018 dataset [6] is an extremely big dataset and comprises seven different attack eventualities: brute-force, heartbleed, botnet, dos, web assaults, ddos, and infiltration of the network from inside. The attacking infrastructure comprises of 50 machines. The victim agency has five departments with 420 machines and 30 servers. The dataset contains the captures community site visitors and gadget logs of all machine, together with eighty attributes mined from the confined traffic the usage of CICFlowmeter-v3.

Simulation results

To classify intrusion by the label, in the attack column from the dataset partition the data into classes and convert these labels to categorical. The distribution of the different attacks in DARPA dataset is pictorially given in Fig. 3 and its wordcloud illustration is depicted. A word cloud offers a visualization of intrusions characteristically connected with Cloud network security.

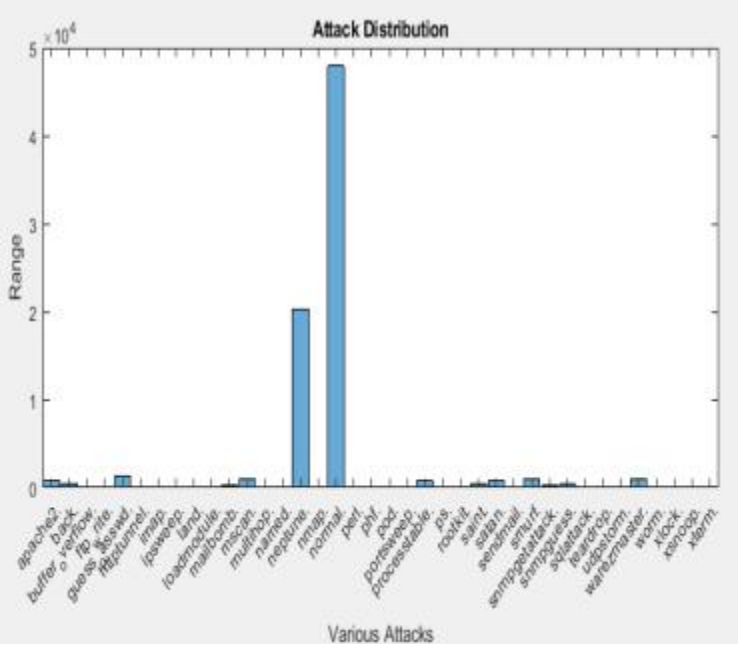

Fig. 3. Attack distribution.

VI. CONCLUSION

The intrusion detection technology is a significant way to preserve the cloud network security. Hence an innovative optimized custom RC-NN network model in detecting intrusion is proposed. This network is concluded effective when compared to other existing classifiers like LSTM, CNN and LSTM with CNN. These classifier outperforms show that the other classifiers with TPR as 0.8048 and TNR as 1 respectively indicating a better performance of the proposed IDS model. With this method, accuracy is increased by 0.9401 and 0.9428 with less error rate of 0.0012 and 0.0011. Hence, this approach can detect each attack better than other existing work. In future, the present work can be extended to model a management module for initiating preventive actions after intrusion detection with the help of classified results.

REFERENCES

[1] A. A. Titorenko and A. A. Frolov, "Analysis of modern intrusion detection system," 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), Moscow, 2018, pp. 142-143.

[2] Alexopoulos, Nikolaos & Vasilomanolakis, Emmanouil & Réka Iványkó, Natália & Mühlhäuser, Max. (2018). Towards BlockchainBased Collaborative Intrusion Detection Systems: 12th International Conference, CRITIS 2017, Lucca, Italy, October 8- 13, 2017.

[3] Axelsson, Stefan. Intrusion detection systems: A survey and taxonomy. Vol. 99. Technical report, 2000.

[4] H. M. Anwer, M. Farouk and A. Abdel-Hamid, "A framework for efficient network anomaly intrusion detection with features selection," 2018 9th International Conference on Information and Communication Systems (ICICS), Irbid, 2018, pp. 157-162.

[5] H.Debar, M. Dacier, and A. Wespi, “Towards a taxonomy of Intrusion Detection Systems”, Computer Networks, vol 31, n0. 8, pp. 805-822, 1999.

[6] Holtz, Marcelo D. ; Bernardo David ; Sousa Jr., R. T. . Building Scalable Distributed Intrusion Detection Systems Based on the MapReduce Framework. Telecomunicacoes (Santa Rita do Sapucai), v. 13, p. 22-31, 2011.

[7] J. Dean and S. Ghemawat, MapReduce: Simplified Data Processing on Large Cluster, USENIX OSDI,2004. Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020) IEEE Xplore Part Number:

[8] J. Yang, C. Shen, Y. Chi, P. Xu and W. Sun, "An extensible Hadoop framework for monitoring performance metrics and events of OpenStack cloud," 2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA), Shanghai, 2018, pp. 222-226.

[9] K. Kato and V. Klyuev, "Development of a network intrusion detection system using Apache Hadoop and Spark," 2017 IEEE Conference on Dependable and Secure Computing, Taipei, 2017, pp. 416-423.

[10] Konstantin Shvachko, Hairong Kuang, Sanjay Radia, and Robert Chansler, “The Hadoop Distributed File System,” IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST), pp.1-10, 2010.

[11] S. Ghribi, A. M. Makhlouf and F. Zarai, "C-DIDS: A Cooperative and Distributed Intrusion Detection System in Cloud environment," 2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC), Limassol, 2018, pp. 267-272.

[12] Suah Kim, Beomjoong Kim, and Hyoung Joong Kim. 2018. Intrusion Detection and Mitigation System Using Blockchain Analysis for Bitcoin Exchange. In Proceedings of the 2018 International Conference on Cloud Computing and Internet of Things (CCIOT 2018). ACM, New York, NY, USA, 40-44.

[13] W. Meng, E. W. Tischhauser, Q. Wang, Y. Wang and J. Han, "When Intrusion Detection Meets Blockchain Technology: A Review," in IEEE Access, vol. 6, pp. 10179-10188, 2018

. [14] Yeonhee Lee and Youngseok Lee. 2012. Toward scalable internet traffic measurement and analysis with Hadoop. SIGCOMM Comput. Commun. Rev. 43, 1 (January 2012), 5-13.

[15] Zohreh Abtahi Foroushani and Yue Li. 2018. Intrusion Detection System by Using Hybrid Algorithm of Data Mining Technique. In Proceedings of the 2018 7th International Conference on Software and Computer Applications (ICSCA 2018). ACM, New York, NY, USA, 119-123.

[16]Satish, Karuturi S R V, and M Swamy Das. "Multi-Tier Authentication Scheme to Enhance Security in Cloud Computing." IJRAR (International Journal of Research and Analytical Reviews) 6, no. 2 (2019): 1-8, 2019.