Analysis of Machine Learning based Security Detection Systems in Cloud Computing Environment

[1]Ms.Samineni Nagamani, [2] Dr.S.Arivalagan, [3] Dr.M.Senthil, [4] Dr.P.Sudhakar

[1]Research Scholar, [2]Assistant professor, [3]Professor & HOD,, [4]Associate Professor,
[1,2,4]Department of CSE, [3]Department of AIML,
[1,2,4,]Annamalai University, [3] QIS College of Engineering and Technology

Abstract: **Cloud Computing (CC) offers on-demand network access to a group of configurable computing resources like network, storage, service, server, and application, that is released quickly with lesser service provider connections or management endeavours. The distributed and open infrastructure of CC and service develops an attractive target for potential cyber-attacks with intruders. The classical Intrusion Detection and Prevention Systems (IDPS) were considered mostly ineffective that utilized in CC platforms because of their openness, dynamicity, and virtualization in existing services. This article offers an Analysis of Machine Learning oriented Intrusion Detection Systems in CC Environment. This paper identifies the probable solutions for intrusion detection and prevention in the cloud platform. The major features of IDS along with its types are defined clearly. Besides, the study surveys the recently developed IDS models for cloud environment, with the help of advanced approaches to resolve the issues posed by the CC needs. The reviewed methods are elaborated with the intention, technique used, and experimental results. At last, a detailed result analysis of the reviewed approaches was provided.**

Keywords**: Security; Intrusion detection system; Cloud computing; Machine learning; Deep learning**

1. Introduction

Cloud computing (CC) presents virtualized on-demand, scalable services to the end user with lesser infrastructural investment and greater flexibility [1]. This can be provided through the Internet using known network standards, protocols, and formats under the control of various managements. CC's aim is to give on-demand, network, convenient access to the shared pool of configurable computing resources (services, servers, storage, applications and networks), that is released and provisioned rapidly with service provider interactions or minimum management efforts [2]. Cloud framework uses virtualization techniques, and integrated technology and runs through standard Internet protocol. This might attract intruders because of several vulnerabilities included in it.

A severe security problem for the IDS is facing malicious software variation that leads to serious faults and network security breaches [3]. Cyberattack is challenging and more complicated in unknown malware attack detection because of the advancement of evasion method to steal essential data and evade IDS from identification [4]. Furthermore, there exist cybersecurity threats during internetwork transmission. Thus, new techniques and solutions are crucial for timely intrusion detection and attack prevention methods. Deep learning (DL) and Machine learning (ML) approaches were newly designed and use IDS for prevention and the detection of abnormal behavior in the network [5]. IDS gives the solution for distinct security-based problems with distinct kinds of intruders or malicious attacks in the network. In the presented method, the distinct IDSs are deliberated. Furthermore, DL based techniques for IDS were broadly described. Fig. 1 illustrates the framework of IDS in cloud platform.

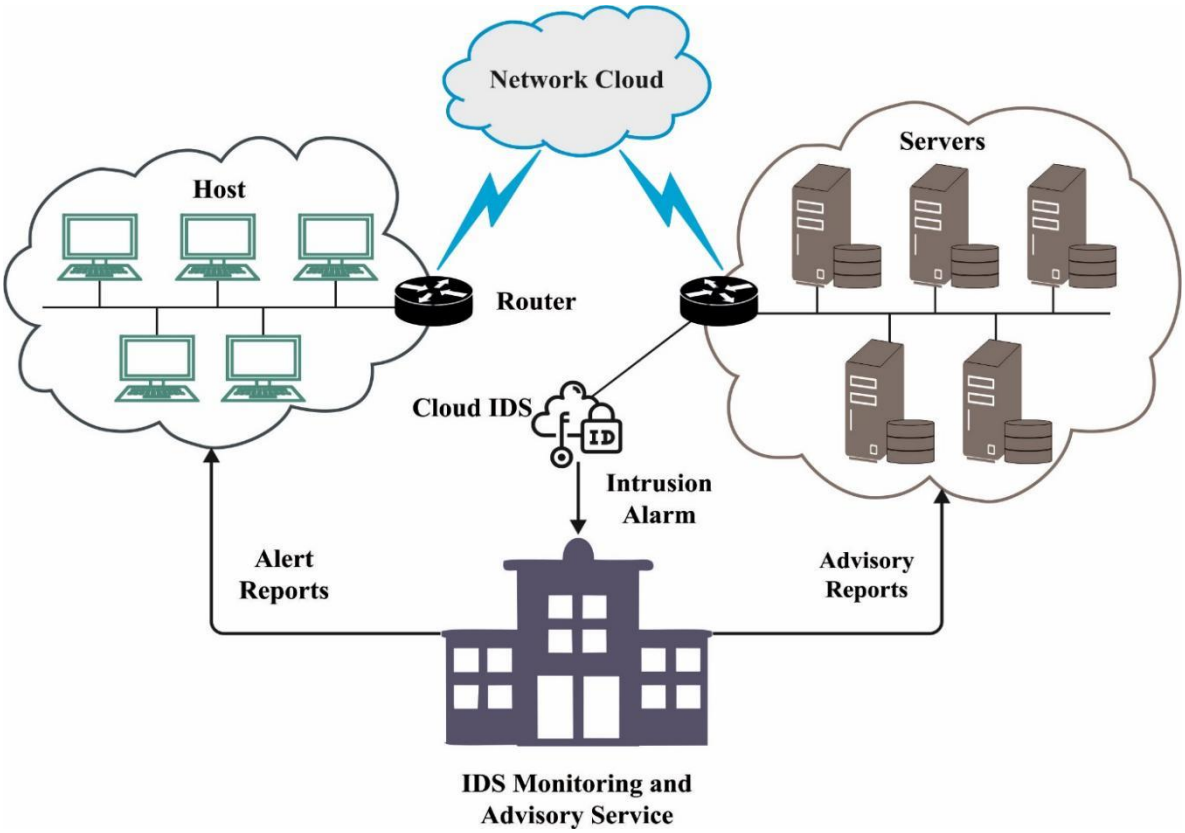

Fig. 1. Structure of IDS in cloud environment

Several studies were carried out in misuse and anomaly detection contexts with different ML methods [6]. Traditional ML technique suffers from insufficient labelled training dataset and relies on the extracted feature by human, making them hard to deploy on massive platform [7]. DL is an alternative model in the ML fields established mainly through ANNs and has high accuracy than the other traditional ML approaches [8]. In recent times, DL technique was effectively employed in different fields, namely visual processing, text, and audio along with contexts like natural language processing (NLP), sentiment analysis, social network analysis, wireless networking, recommender systems, etc [9]. In addition, DL has accomplished a considerable interest in the IDS contexts, and several DL based anomaly and misuse detection techniques are presented to handle different kinds of security attacks and intrusions [10].

This article offers an Analysis of Machine Learning oriented Intrusion Detection Systems in CC Environment. This paper identifies the probable solutions for intrusion detection and prevention in the cloud platform. The major features of IDS along with its types are defined clearly. Besides, the study surveys the recently developed IDS models for cloud environment, with the help of advanced approaches to resolve the issues posed by the CC needs. The reviewed methods are elaborated with the intention, technique used, and experimental results. At last, a detailed result analysis of the reviewed approaches was provided.

2. Background Information

The IDS primarily encompasses three sections. Firstly, information on cyberattack evidence is gathered from input dataset and later processed to detect and analyze the next segment cyberattack. Lastly, the attack was reported. DL and ML-based approaches are now exploited for predicting abnormal and normal behaviors and novel unidentified assaults in the network by analysing input dataset. (IDS technique is categorized into different kinds, for instance,

- specification-Based Detection,
- signature-oriented intrusion detection system (SIDS),
- anomaly-related IDS (AIDS),
- host-related IDS (HIDS),
- hybrid-related detection,

- distributed-based IDS (DIDS), and
- network-oriented IDS (NIDS).

Signature-Based Intrusion Detection Systems (SIDS).
SIDS is known as knowledge-oriented recognition. It evaluates and analyses networks depending on corresponding signatures or known patterns to find assault signs in the signature dataset by comparing activities network and transmission. It saves the signature and behavior of all the attacks in the network. An alert will be generated if the attack sign is matched or found with saved signature dataset.
It shows that SIDS finds the attack whose signature was saved in databases. The novel attack was identified by means of SIDS, where it is not as precise in contradiction of attack variation. The alert scheme diminishes false alerts because of accurate and effective misbehaviour classification and identification to measure network administrators initiating defensive action.
Anomaly-Based Intrusion Detection System (AIDS).
AIDS otherwise called dynamic behavior or profile-based IDS and is the more commonly used method than SIDS because of its efficiency against innovative assaults. AIDS can be widely applied for resolving the shortcomings of SIDS. Unrecognized attack at distinct phases creates alert to identify the exposure and prevent them with promising methods. AIDS monitors the scheme reliably for collecting information for the identification of abnormal or normal. Zero-day attack detection was the basic objective of AIDS since novel anomalous action was concerned with pattern databases. It learns abnormal behaviors within the network. For instance, if there is stealing from an account or if any unauthorized activity occurs, the alarm is produced. Abnormal behaviors are novel typical actions, not affected intrusion, leading to a higher false-positive rate.
Customized Intrusion Detection Methods.
Customized and AIDS work similarly, whereas these techniques provide and develop rules and specifications manually to define normal network activity. The network can be observed based on the presented set of instructions and rules. It contains a minimal false positive rate because of resistance to novel assault variation. The personalized IDS has limitations because of restrictions and complexities in cost, advancement, and time consumption.
Hybrid Intrusion Detection Methods.
This method called compound recognition was established by integrating misuse, specification, and anomaly detection methods for overcoming the deficiency and improving the recognition of new and existing attack behaviors. For instance, SVELTE IDS approach has been introduced by means of hybrid technique (AIDS and SIDS) for 6LoWPAN network in IoT interconnected. This hybrid method was introduced for accomplishing steadiness of this technique's complexity, storage, processing, and cost.
Host-Based IDS (HIDS).
The HIDS is software mounted on network host computer which monitors, scrutinizes, analyses, and collects the information action reliably within the networks by examining servers, database logs, or firewalls. HIDS is constraint to detect an individual abnormal attack when identifying simple attack within the network.
Network-Based IDS (NIDS).
NIDS monitor network communication by collecting packet captures and other via NetFlow. The fundamental aim was to protect the network from exterior attack causes an alarm or alert once the malicious attack occurs. The IDS functions by more than one host through the network and external firewall by analyzing and monitoring network transmission by means of hardware or software. Software was installed over server for monitoring, whereas sensor was attached to the server for analyzing the network communication. Consequently, NIDS is secure and effective in identifying malicious attacks. NIDS has numerous restrictions; it could not analyze and process the large network dataset because of traffic flow and higher bandwidth. NIDS is unable to encrypt network packets.
Distributed IDS (DIDS).
DIDS encompasses various IDSs on wide-ranging networks for analyzing malicious incidents, attack information, and transmission monitoring management. Data incorporates several sensors (HIDS and NIDS based) and central analyzer for prevention and management of IDS.

3. Current IDS Models in Cloud Environment

In [11], devised a potential IDS called Chronological Salp Swarm Algorithm-related DBN to detect doubtful intrusion in cloud platforms. Hence, the presented Chronological Salp Swarm Algorithm-oriented DBN was formulated through integration of Salp Swarm Algorithm and Chronological concept. The best solution for IDS was exposed by making use of fitness function, which will accept reduced error value as best solution. Here, to generate an optimal and effective solution for IDS, the weights can be tuned optimally by the presented method.

Alkadi et al. [12] devise a deep blockchain framework (DBF) to provide privacy-oriented BC and security-oriented distributed ID with smart contracts in IoT platforms. The ID technique was used by a BiLSTM-DL technique to manage sequential network data and was evaluated through the datasets of BoT-IoT and UNSW-NB15. By utilizing the Ethereum library, the privacy-related BC and smart contract approaches were formulated to offer privacy to the dispersed ID engines.

In [13], devised a structure called IDSGT-DNN was for fostering security in cloud IDS mechanism. The developed game theory was enforced in DNN method including IWA for identifying the best solutions. Louati and Ktata [14] devise a DL-related multiagent mechanism for ID which integrates the favourable features of multi-agent mechanism with accuracy of DL techniques. Then, constituted various adaptive, autonomous, and intelligent agents that implanted 3 methods such as KNN, AE, and MLP. AE was employed as feature reduction mechanism, and KNN and MLP will be employed as classifiers.

Sreelatha et al. [15] devise a potential cloud IDS utilizing the sandpiper-oriented FS and extended equilibrium deep TL (EEDTL) classification for fostering overall security of a cloud-related computing architecture. Lastly, for classifying different attacks related to their chosen optimal features the EEDTL method was employed. For optimal tuning features in convolutional layers, TL employs a pre-trained network named AlexNet.

In [16], the authors developed a new structure for a deep LSTM-related IDS for detecting network traffic flow patterns as either normal or malevolent in a cloud. The presented IPS thwarts malicious attacks by diminishing computational period and rising the detection rate of malicious assaults. A network ID technique that combines BiLSTM and CNN networks were introduced in [17]. Initially, the KDD CUP 99 dataset was preprocessed through data extraction method. Lastly, CNN-BiLSTM and C5.0 DT utilizing the DL technique were integrated to skip design FS and directly employ DL method to study representational features of high dimensional datasets.

In [18], the authors developed an IDS was presented related to a new optimized custom RC-NN for ID with ALO method. In this work, CNN was made hybrid with LSTM. Wang et al. [19] intend to employ DL to derive indispensable feature representation automatically and realize high detecting performance efficiently. By utilizing the SCAE approach, robust and better low-dimensional features are learned from raw network traffic automatically. A new cloud IDS was proposed based on the SVM and SCAE classifier methods.

Almiani et al. [20] designed a full-automated IDS for Fog security against cyberattacks. This technique employs multi-layered RNN devised to be applied for Fog computing security that is near to the IoT devices and users.

Chkirbene et al. [21] developed two paradigms for classification and detection of intrusion. Trust-based ID and Classification System- Accelerated (TIDCS-A) and Trust-based ID and Classification System (TIDCS) for secured network. TIDCS decreases variety of characteristics in the input dataset based on a new model for selecting features. Originally, the feature was arbitrarily grouped to raise chance of making them take part in the generation of dissimilar groups and arranged according to the performance score. Abusitta et al. [22] introduced an ML-based cooperative IDS that effectively uses past feedback dataset to give the capability of pro-active decision making. Particularly, presented method relies upon Denoising Autoencoder (DA) that is utilized as a major component to create a DNN. The power of DA lies in its capability to learn how to recreate IDS feedback from partial feedback.

In [23], the authors proposed a host-based IDS (H-IDS) to protect VMs in the cloud platform. For this reason, firstly, significant features of every class are chosen by means of logistic regression, and then, this value was enhanced by means of the regularization method. Next, different attacks are categorized with the incorporation of three distinct classifiers: LDA, NN, and DT with the bagging algorithm for all the classes. Arjunan and Modi [24] designed robust security architecture for detecting intrusion at the virtual network layer of cloud. It integrates anomaly and signature-based methods for detecting potential attacks. It employs distinct

classifications that are; NB, DT, RF, extra trees, and LDA for effective and efficient IDS. To identify distributed attacks in whole Cloud and all the clusters, it gathers intrusion evidence from all the regions of the Cloud and uses Dempster-Shafer theory (DST) for making concluding decisions.

Ghanshala et al. [25] present an adaptable and light weighted IDS termed a Behavior-oriented Network ID (BNID) at network layer in cloud. The behavioural analysis of traffic is implemented at Cloud Network Node (CNN) to identify the intrusion. A security architecture was developed for deploying the BNID in cloud. The necessity of IDS deployment in all tenant virtual machines (TVM) was disregarded. BNID applies statistical learning methodologies with FS for the analysis of traffic behaviors and doesn't involve wide monitoring of memory writes.

Gao et al. [26] introduced a fuzziness-related semi-supervised learning method based on ensemble learning for network ID on the cloud-related robotic systems that could resolve the abovementioned problems. Firstly, because of the better generalization capability of ensemble learning, the author constructs an ensemble mechanism trained by labeled dataset. Furthermore, for using the unlabeled dataset, a fuzziness-related technique is implemented for the analysis of dataset.

Li et al. [27] analyze the intrusion threat introduced by power information networks and conduct exhaustive investigation and research integrated with the IDS. It analyses the architecture of the CC and power knowledge network through DL-based method and gives a network interference recognition method.

4. Performance Validation

The intrusion detection results of the different ML and DL methods in cloud environment are depicted in Table 1. Fig. 2 reports an overall TNR and TPR examination of the diverse ML and DL models. The results indicated that the KNN and MLP models have shown improved results. For instance, based on TNR, the KNN and MLP have shown higher TNR of 99.97% and 99.83% whereas the BiLSTM, SVM, LSTM with Adam, CNN, and CNN+LSTM models have obtained lower TNR of 98.88%, 97%, 95.77%, 92.82%, and 95.75% respectively. Also, with respect to TPR, the KNN and MLP have revealed enhanced TPR of 99.88% and 99.54% whereas the BiLSTM, SVM, LSTM with adam, CNN+LSTM, and CNN approaches have attained reduced TPR of 99.47%, 99.32%, 98.97%, 97.6% and 90.96% correspondingly.

Table 1 Comparative analysis of distinct ML and DL approaches interms of distinct measures

Methods	TNR	TPR	Accuracy	DR	Precision
K-NN	99.97	99.88	99.95	99.88	99.88
MLP	99.83	99.32	99.73	99.32	99.32
BiLSTM	98.88	98.97	98.07	97.64	97.57
SVM	97.00	97.60	97.39	97.72	97.24
LSTM with adam	95.77	99.47	63.30	99.47	62.91
CNN	92.82	90.96	58.11	90.96	61.56
CNN+LSTM	95.75	99.54	62.97	99.54	62.95

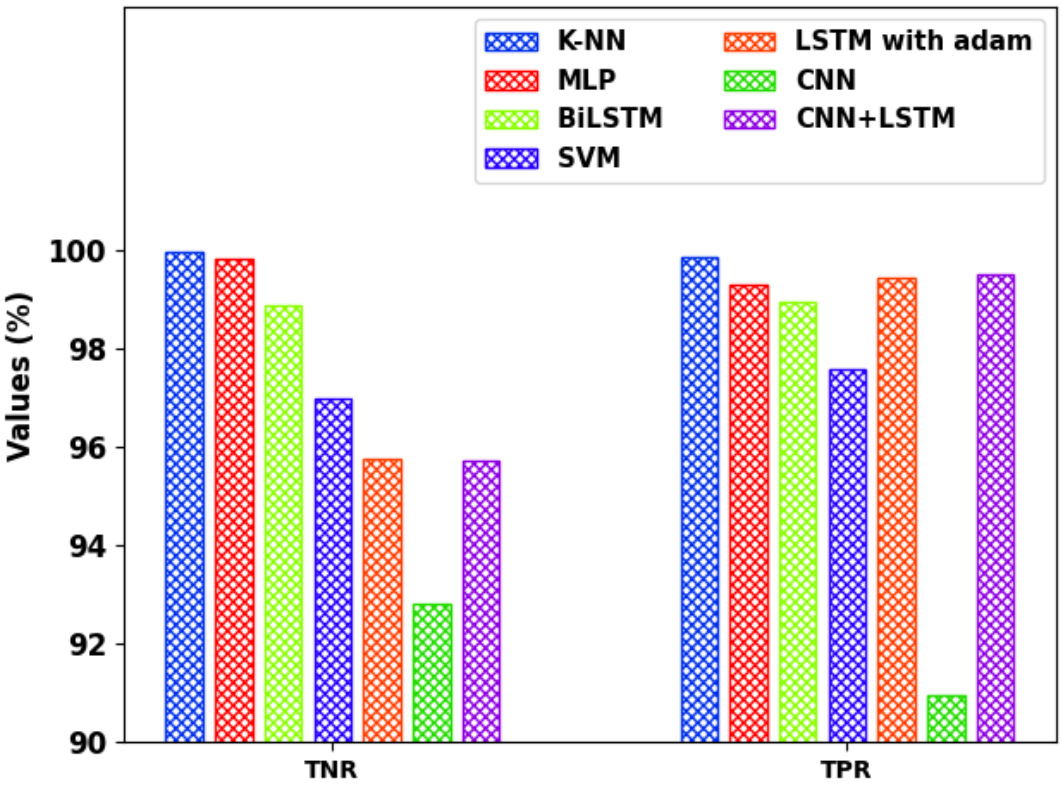

Fig. 2. TNR and TPR analysis of distinct ML and DL approaches

Fig. 3 demonstrates the overall $accu_y$, $prec_n$, and DR inspection of the varied ML and DL approaches. The outcomes stated that the KNN and MLP approaches have exhibited enhanced results. For instance, interms of $accu_y$, the KNN and MLP have revealed greater $accu_y$ of 99.95% and 99.73% whereas the BiLSTM, SVM, LSTM with adam, CNN+LSTM, and CNN models have obtained lower $accu_y$ of 98.07%, 97.39%, 63.3%, 62.97% and 58.11% correspondingly. Followed by, based on DR, the KNN and MLP have shown higher DR of 99.88% and 99.54% whereas the BiLSTM, SVM, LSTM with adam, CNN+LSTM, and CNN systems have obtained reduced DR of 99.47%, 99.32%, 97.72%, 97.64% and 90.96 respectively. Meanwhile, concerning $prec_n$, the KNN and MLP have shown maximal $prec_n$ of 99.88% and 99.32% whereas the BiLSTM, SVM, LSTM with adam, CNN+LSTM, and CNN models have obtained lower $prec_n$ of 97.57%, 97.24%, 62.95%, 62.91% and 61.56% correspondingly.

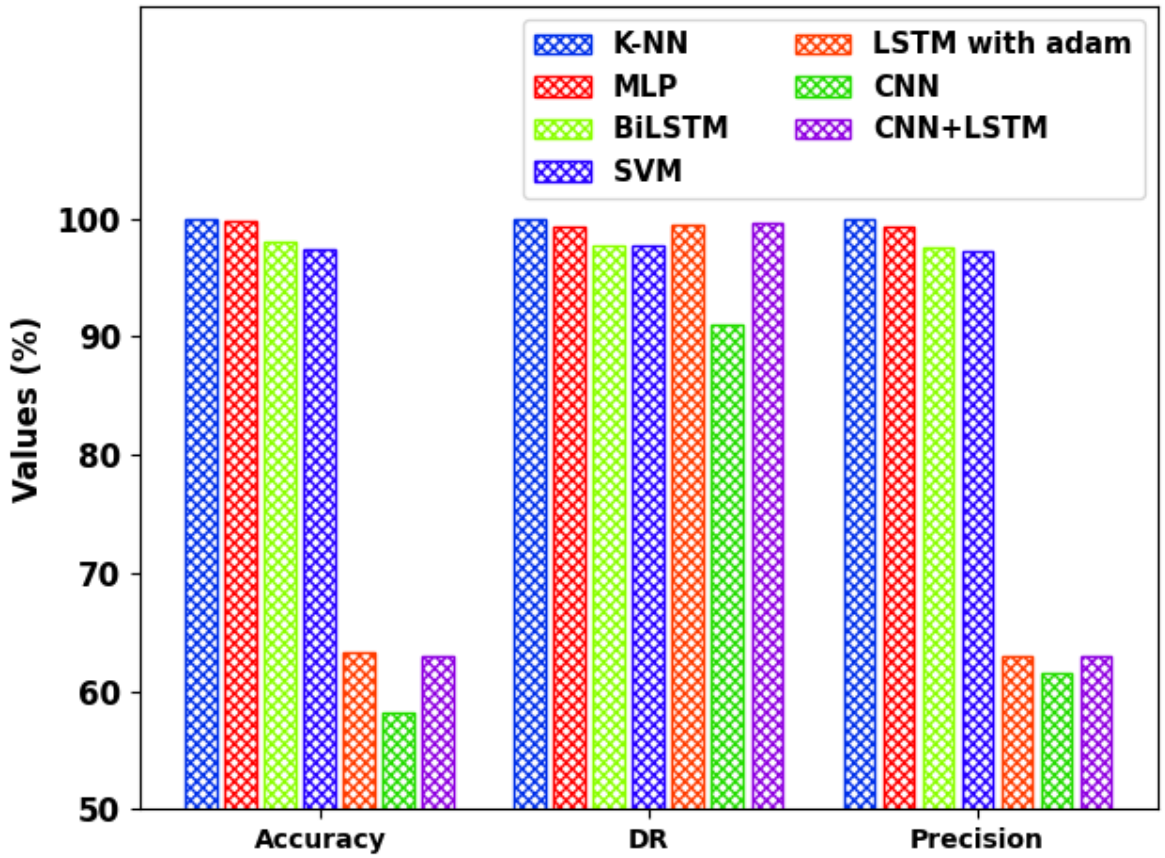

Fig. 3. $Accu_y$, DR, and $prec_n$ analysis of distinct ML and DL approaches

5. Conclusion

This article has offered a detailed investigation of several ML based IDS models for CC environment. This paper identified the probable solutions for intrusion detection and prevention in the cloud platform. The major features of IDS along with its types are defined clearly. Besides, the study surveys the recently developed IDS models for cloud environment, with the help of progressive approaches to resolve the issues posed by the CC needs. The reviewed methods are elaborated with the intention, technique used, and experimental results. At last, a detailed result analysis of the reviewed approaches was provided. In the future, we observe the performance of IDS models on real time datasets.

References:

1. Onyema, E.M., Dalal, S., Romero, C.A.T., Seth, B., Young, P. and Wajid, M.A., 2022. Design of Intrusion Detection System based on Cyborg intelligence for security of Cloud Network Traffic of Smart Cities. *Journal of Cloud Computing*, *11*(1), pp.1-20.
2. Singh, P. and Ranga, V., 2021. Attack and intrusion detection in cloud computing using an ensemble learning approach. *International Journal of Information Technology*, *13*(2), pp.565-571.
3. Samy, I.A.A. and Mary, M.S., 2022. An Improved Ecc Algorithm for Secure Cloud Storage System With the Help of Sha-256 Based User Authentication and Deep Learning Based Intrusion Detection System.
4. Mayuranathan, M., Murugan, M. and Dhanakoti, V., 2021. Best features based intrusion detection system by RBM model for detecting DDoS in cloud environment. *Journal of Ambient Intelligence and Humanized Computing*, *12*(3), pp.3609-3619.
5. Samriya, J.K., Tiwari, R., Cheng, X., Singh, R.K., Shankar, A. and Kumar, M., 2022. Network intrusion detection using ACO-DNN model with DVFS based energy optimization in cloud framework. *Sustainable Computing: Informatics and Systems*, *35*, p.100746.

6. Hizal, S., ÇAVUŞOĞLU, Ü. and AKGÜN, D., 2021, June. A New Deep Learning Based Intrusion Detection System for Cloud Security. In *2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)* (pp. 1-4). IEEE.
7. Sethi, K., Kumar, R., Prajapati, N. and Bera, P., 2020, January. Deep reinforcement learning based intrusion detection system for cloud infrastructure. In *2020 International Conference on COMmunication Systems & NETworkS (COMSNETS)* (pp. 1-6). IEEE.
8. Balamurugan, V. and Saravanan, R., 2019. Enhanced intrusion detection and prevention system on cloud environment using hybrid classification and OTS generation. *Cluster Computing*, *22*(6), pp.13027-13039.
9. Selvapandian, D. and Santhosh, R., 2021. Deep learning approach for intrusion detection in IoT-multi cloud environment. *Automated Software Engineering*, *28*(2), pp.1-17.
10. Bharati, M.P. and Tamane, S., 2020, October. NIDS-Network Intrusion Detection System Based on Deep and Machine Learning Frameworks with CICIDS2018 using Cloud Computing. In *2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC)* (pp. 27-30). IEEE.
11. Karuppusamy, L., Ravi, J., Dabbu, M. and Lakshmanan, S., 2022. Chronological salp swarm algorithm based deep belief network for intrusion detection in cloud using fuzzy entropy. *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, *35*(1), p.e2948.
12. Alkadi, O., Moustafa, N., Turnbull, B. and Choo, K.K.R., 2020. A deep blockchain framework-enabled collaborative intrusion detection for protecting IoT and cloud networks. *IEEE Internet of Things Journal*, *8*(12), pp.9463-9472.
13. Balamurugan, E., Mehbodniya, A., Kariri, E., Yadav, K., Kumar, A. and Haq, M.A., 2022. Network optimization using defender system in cloud computing security based intrusion detection system withgame theory deep neural network (IDSGT-DNN). *Pattern Recognition Letters*, *156*, pp.142-151.
14. Louati, F. and Ktata, F.B., 2020. A deep learning-based multi-agent system for intrusion detection. *SN Applied Sciences*, *2*(4), pp.1-13.
15. Sreelatha, G., Babu, A.V. and Midhunchakkaravarthy, D., 2022. Improved security in cloud using sandpiper and extended equilibrium deep transfer learning based intrusion detection. *Cluster Computing*, pp.1-16.
16. Mani, S., Sundan, B., Thangasamy, A. and Govindaraj, L., 2022. A New Intrusion Detection and Prevention System Using a Hybrid Deep Neural Network in Cloud Environment. In *Computer Networks, Big Data and IoT* (pp. 981-994). Springer, Singapore.
17. Gao, J., 2022. Network Intrusion Detection Method Combining CNN and BiLSTM in Cloud Computing Environment. *Computational Intelligence and Neuroscience*, *2022*.
18. Thilagam, T. and Aruna, R., 2021. Intrusion detection for network based cloud computing by custom RC-NN and optimization. *ICT Express*, *7*(4), pp.512-520.
19. Wang, W., Du, X., Shan, D., Qin, R. and Wang, N., 2020. Cloud intrusion detection method based on stacked contractive auto-encoder and support vector machine. *IEEE transactions on cloud computing*.
20. Almiani, M., AbuGhazleh, A., Al-Rahayfeh, A., Atiewi, S. and Razaque, A., 2020. Deep recurrent neural network for IoT intrusion detection system. *Simulation Modelling Practice and Theory*, *101*, p.102031.
21. Chkirbene, Z., Erbad, A., Hamila, R., Mohamed, A., Guizani, M. and Hamdi, M., 2020. TIDCS: A dynamic intrusion detection and classification system based feature selection. *IEEE Access*, *8*, pp.95864-95877.
22. Abusitta, A., Bellaiche, M., Dagenais, M. and Halabi, T., 2019. A deep learning approach for proactive multi-cloud cooperative intrusion detection system. *Future Generation Computer Systems*, *98*, pp.308-318.
23. Besharati, E., Naderan, M. and Namjoo, E., 2019. LR-HIDS: logistic regression host-based intrusion detection system for cloud environments. *Journal of Ambient Intelligence and Humanized Computing*, *10*(9), pp.3669-3692.
24. Arjunan, K. and Modi, C.N., 2017, January. An enhanced intrusion detection framework for securing network layer of cloud computing. In *2017 ISEA Asia Security and Privacy (ISEASP)* (pp. 1-10). IEEE.

25. Ghanshala, K.K., Mishra, P., Joshi, R.C. and Sharma, S., 2018, December. BNID: a behavior-based network intrusion detection at network-layer in cloud environment. In *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)* (pp. 100-105). IEEE.
26. Gao, Y., Liu, Y., Jin, Y., Chen, J. and Wu, H., 2018. A novel semi-supervised learning approach for network intrusion detection on cloud-based robotic system. *IEEE Access*, *6*, pp.50927-50938.
27. Li, T., Zhao, H., Tao, Y., Huang, D., Yang, C. and Xu, S., 2022. Power Intelligent Terminal Intrusion Detection Based on Deep Learning and Cloud Computing. *Computational Intelligence and Neuroscience*, *2022*.