Journal of Engineering Design and Analysis
Volume 6, Print Issue 1 - 2023, Pg. No. 16-22
Peer Reviewed Journal

Review Article

Intrusion Detection Scheme through multilevel ML Classifier

Hemant Kumar Saini[1], Ankesh Gupta[1], Gurleen Kaur[1]

[1]Department of CSE,UIE Chandidarg University Mohali, Punjab, India.

I N F O

Corresponding Author:
Hemant Kumar Saini, Department of CSE,UIE Chandidarg University Mohali, Punjab, India.
E-mail Id:
hemantrhce@rediffmail.com
Orcid Id:
https://orcid.org/0009-0009-7426-4977
How to cite this article:
Saini HK, Gupta A, Kaur G. Intrusion Detection Scheme through multilevel ML Classifier. *J Engr Desg Anal* 2023; 6(1): 16-22.

Date of Submission: 2023-03-05
Date of Acceptance: 2023-04-10

A B S T R A C T

In today's era with the emerging trends of Big Data and IoT the network traffic range of services derived for the users according to their needs. Mostly public users use the open channels for the transmission of the data which would be a lot o concern over its security. To sustain such security various researches developed many defensive approaches but those are no longer effective. Intrusion detection system (IDS) deployed to detect the various intrusion assaults but they are not up to mark. This paper explores the various classes of intrusions and methodologies to mitigate them. The overview gives the useful resource for naïve researchers, make them better learning of the emerging intrusions and invoke the potential measure involving the Machine learning techniques for future investigation. In particular various potential risks and rewards of intrusive activities are highlighted which will persuade researchers to implement the proactive approaches to address such challenges. Also trying to proposed an IDS where the detection paradigms has been improved by ensembled learner and advanced hyperprameter optimization which lessen the false alarms and identify accurately.

Keywords: Intrusions, Intrusion Detection System, Machine Learning, HIDS, NIDS

Introduction

T With the demands in big data network systems provide the services to its clients for sharing confidential data over the Internet which leads to the security concern. Various intruders tremendously attempt to However, malicious hackers progressively attempt to pilfer the data by assaulting any of the characteristics such as confidentiality, integrity, availability etc. with naive methods. Confidentiality guarantees the data will be accessible to only with the authorized labels, while integrity ensures there is no alteration in middle of traffic and availability ensures the data accessible for the given time frame. Intrusive activities can be from any form therefore data at open channels needs to be pro-protective. Intrusion Detection System (IDS) is deployed to identify the flow of any threat on machine. in general, assaults classified in two groups: primary assaults and secondary assaults. secondary assaults comprised of Denial of Service (DoS) and sensing assaults, while secondary assaults are client admin and terminal assaults, as seen in Figure 1.

Denial of Service (DoS) assaults is the way to overburden the traffic by excessive pulling requests to prevent legitimate users from accessing services. These assaults do not typically involve in stealing or damaging information, but this tried to crash the mailbox, database which causes the losses in time and money. This can't be easily detected so it needs to be more preventive.

Sensing Assaults are such malformed practices where the attackers robotically scans for the open ports and hinders the users from accessing materials and facilities of the host

or network. Diverse approaches used to detect the open ports such as Ipsweep, Portsweep, Nmap and the Satan. Such assaults are highly detrimental in nature for security of the host or network.

Client admin Assaults are common practices done by most of the attackers where they keep eye on the traffic and try to actively change the data over the internet through different techniques such as Warezclient, Phf, Ftpwrite, Imap, Warezmaster. By implementing such tools assaulter gain the access to machine without the need of authentication and changes the data. Such type can easily be probed.

Terminal assaults having the potential consequences where the root privileges is trying to tamper on host or network system by unauthorized actions. Mostly the attackers uses terminals of the root and try to access the confidential files and exfiltrate them through shared protocols such as mail and/or FTP which include Loadmodule, Perl, Buffer Overflow. Henceforth when the root terminal is in hands of attacker then it causes more so this is more significant to be aware of such types of assaults and their potential consequences.

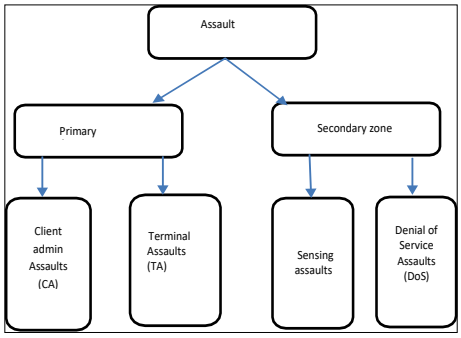

Figure 1.Evolution of types of Assaults

Despite of numerous research is effervescence to detect the DoS assault many existing models still can't predict the accurately especially in the primary zoneone such model is proposed by P.Amudha et. al.[7] which works as an IDS as depicted in Figure 2, which identifies the primary assaults. To get more accuracy in prediction it is being proposed the model can be improved with the use of effective Machine Learning techniques to get better detection and respond insuch threats.

Authors of[7] described the architectural framework as follows: The Data Gathering Unit- it gathers the statistics from the events sensorized, with the Detector-ID which progressively collect the data to identify the intrusive

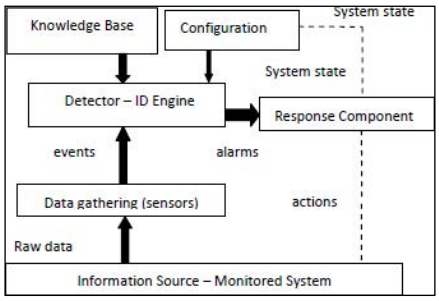

Figure 2.Identification of Assaults[7]

activity and alerts for malformed data activity. Here the Knowledge Base engine provides the pre-processed information to identify the misleading behavior in traffic and the configuration unit helps to provide the solution for it after checking with both constraints the ID engine send the alarm to the response unit Consequently, these devices work together to ensure an effective and efficient Intrusion Detection System.

The rest of the paper is organized as follows. Section 2 introduces last works done in IDS. Section III discusses the categories of IDS. In Section IV, opens the door for various research changes. Finally, Section V concludes the paper

Related Work

To investigate the IDS various researches explored the techniques which are entailed in Table 1.

Table I.Survey of Various Ids Techniques In Recent Years

S. No	Reference	Year	Overview
1.	M. Mehmood et. al.,[9]	2022	In addition, a method of study with three separate phases of activity has been proposed for the detection of invasive acts in systems. The strategy uses transformation and the min-max method to preprocess data at the early step. The random forest feature selection phase comes next, then a hybrid mode of SVM and Adaptive Neuro-Fuzzy System is used to increase the rate of attack detection. The simulation results show that the suggested method significantly improves the ability to spot irregular patterns.
2.	Cao B et. al.,[10]	2022	This work suggests an unique hybrid method that combines CNN and GRU techniques to identify invasive activities. The original dataset's positive and negative sample imbalance concerns were addressed using the ADASYN and RENN sample processing techniques. The best features were then chosen using an ensemble technique that included Random Forest and PCA. This strategy showed great effectiveness through a thorough review procedure, making it a potential solution.

3.	Fu Y. et. al. ,[11]	2022	Investigators have developed a unique technique called DLNID for identifying network abnormalities using deep learning models. They used a Bi-LSTM learning strategy, a CNN approach for extracting sequence characteristics from data traffic, adaptive synthetic sampling (ADASYN) to deal with the problems of data imbalance. They used the NSL-KDD dataset to illustrate the efficacy of their strategy, they reported a 90.73% accuracy and 89.65% F1 score, demonstrating the success of their suggested strategy.
4.	M. Ashfaq Khan and Y. Kim[12]	2021	A hybrid intelligent intrusion detection system (HIIDS) is used in this study to analyse crucial aspects from huge amounts of unlabeled raw network traffic data. Long Short-Term Memory (LSTM) for temporal feature recognition and Autoencoders (AE) for global feature identification have been merged. The proposed approach attained a remarkable accuracy rate of 97.52% during simulated trials. Thus, it can be said that HIIDS is a useful method for spotting intrusions in network traffic data.
5.	Li Y. et. al.,[13]	2021	In this study, a hybrid strategy based on the ID3 (Decision tree scheme) and ADASYN (Adaptive Synthetic) algorithms has been explored. The invasive data were initially processed via coding, the oversampling procedure was then carried out using the ADASYN algorithm. The accuracy percentage of the decision tree model that was created later using the ID3 method was 93.18%.
6.	J. Dong Lee et. al.,[14]	2021	This paper introduces the M-IDM intrusion classification architecture (Multi-class Classification based Intrusion Detection Model). The authors suggest that this method can utilise actual data from medical equipment like monitors (electrocardiogram and thermometers). According to simulation findings, the model's accuracy level was 96.7%. As a result, this strategy may be employed as an efficient intrusion detection technique.
7.	Y. S. Sydney and M. Kasongo,[15]	2020	The UNSW-NB15 and AWID IDS datasets were used to develop a novel deep learning-based Intrusion Detection System (IDS) dubbed WFEU-FFDNN. WFEU-FFDNN is a unique intrusion detection system (IDS) that detects malicious activity using a Feed-Forward Deep Neural Network (FFDNN). WFEU-FFDNN surpassed existing IDS systems in terms of accuracy and speed, according to the results. As a result, the suggested approach offers a viable and dependable alternative for identifying harmful activity.
8.	K.E.S.Hadeel Alazzam et.al.,[16]	2020	This work developed a unique feature selection technique for accurately identifying invasive behaviours. The authors used a pigeon-inspired optimizer to find the best feature set, which resulted in a high rate of accuracy. Furthermore, the algorithm's selection method proved to be both efficient and successful.
9.	U. Ahmad et. al.,[17]	2019	In a recent study, the authors analyzed the efficacy of classification technology for identifying intrusions. Through their simulations, they identified the advantages of the MLP scheme in effectively detecting intrusive activity under irregular traffic. Furthermore, they concluded that this scheme provides better accuracy when compared to other existing methods. In summary, their research highlights the effectiveness of MLP in recognizing intrusions, even under irregular traffic conditions.
10.	A. Hajimirzaei et. al.,[18]	2019	This method detects invasive activities by combining Neural Network (NN) with Bee Colony Scheme. The authors used the NSL-KDD dataset to test the approach's efficiency and got an excellent 98.41% detection accuracy. As a result, this method appears to be a potential alternative for identifying and stopping harmful activity.

Architecture of IDS

In[2] autonomous agents collect the intrusion activities based on their interests. Here the authors specify the novel interests on the new alerts. This way the architecture safe itself from unnecessary traffic among the agents and the network will not compromised even if the agent fails. But this degraded the performance of the system.

An Intrusion Detection System (IDS) is a vital tool to guarantee the security and service stability in networks by searching malicious activities and malicious actors. In these machines emotive actions is sensitized during the implementation process of individual functions, then categorize the activities into various categories, such as system component discovery, screening activities and response techniques, as illustrated

in Figure 3. Additional it also provides the data safety and service steadiness of a network system, which gives the complete security coverage. some of the other IDS are being discussed as follows:

Host based Intrusive Discovery System (HIDS) is another tool which can scan and inspect every file, log, kernel entry in order to check the intrusive activities. Another similar is Network Intrusion Detection System (NIDS) but HIDS uses the different techniques get system calls information and tracking which yields more accurate results than NIDS.

NIDS (Network Intrusion Detection System) observe the entire network traffic in place of individual system, making it a more reliable. Sometimes the internal intrusion discovery fails to detect due to the large data traffic, but NIDS sustain with the good hardware which scans the every network packet and detect malicious or abnormal activities. In this way, NIDS provides a more comprehensive and secure means of monitoring a network.

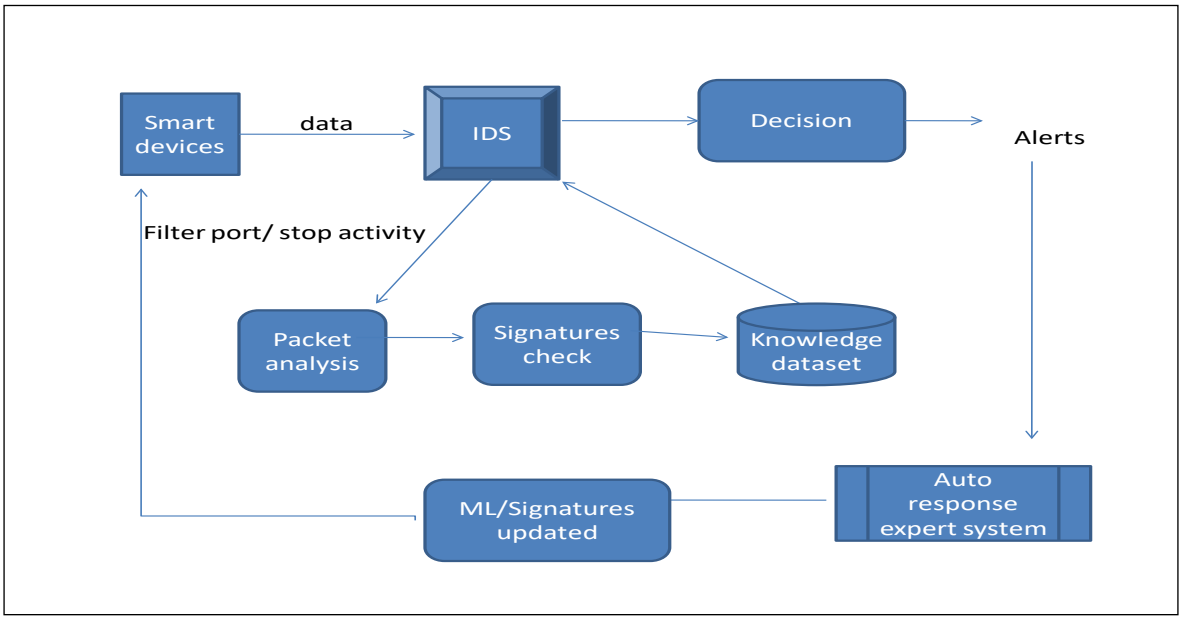

Figure 3.Architectural overview of IDS

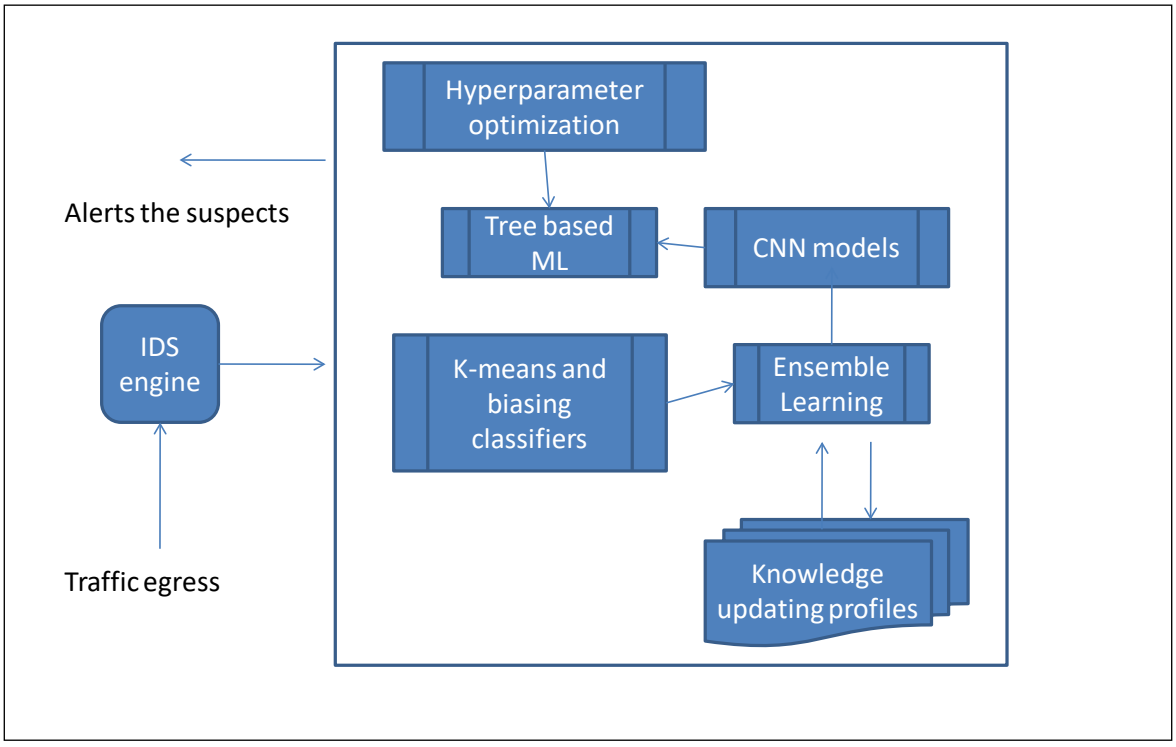

Figure 4.Proposed multilevel IDS idea with multi ML techniques filteration

The signature base Intrusion Detection System (SIDS) is only detect the known signature attacks which are in datasets therefore requiring regular training procedures with updated labeled data in order to detect with updated intrusive activities. Henceforth the signatures need to be regularly updated to get efficient results. Consequence this approach is imperative to ensure the latest signature detection to remain effective.

Anomaly-based Intrusion Detection Systems (AIDS) use regular statistics typically entails mining of knowledge based data that stored the past monitored activities. Consequently, AIDS throw the new events into past knowledge base data and detects he deviation which suspect the new detection this highlighting its significance.

The Active IDS approach is an automatic robotically set up to offer real-time security. However, while the approach can block assaulting acts in real time, it must be placed within the boundaries of the network for it to be effective.

The working methodology of passive IDS is different than the active IDS .it only alerts the admin, warning them to take the immediate actions at the risks. This is just an warning based system where alerting the admin but the complete responsibility is still in manual process where the expertise of admin is necessary.

Challenges Faced by IDS

ANumerous investigations are being explored for sake of technical challenges. In,[14] classified the IDS into five subcategories such as pattern-based, rule-based, statistics-based, state-based, heuristic-based IDS as seen in Figure 4. But this is much confusing due to the number of similar strengths and undistinguished techniques which do not gives clear idea. The signature based IDS usually work with uncertain degree of the false alarm and not able to identify the new attacks.[15] IDS exhibit the high false positive detection based on the stateful protocol analysis with the profile definition. The major challenge is the updating the profiles evolved over the time

Mostly seen in earlier studies that IDS cant able to detect the minute difference between the normal and malicious behavior.[17] So this was improved by pietraszek by aiding the signatures which can detect more accurately but this also not work in certain instances. Henceforth Adaptive Learner for Alert Classification (ALAC) approach proposed by Pietraszek[18] gives the new Machine Learning techniques where the IDS learn the new implicit classification rules with the human activity. But unfortunately this involvement of human there is more chances to get false alarms coupled with lesser detection rate[20] where this challenge raise the demand of balanced dataset which highly impact the model accuracy.

Approaches to solve the Challenges

Network and host-based intrusion detection approaches are great impact on security. To realize the IDS with full functionality the organizations or the government has to pay attention to mitigate such challenges in number of ways.some of the approaches are :

Ensuring an effective deployment

Since due to the low budget tendency it is not possible to deploy the NIDS and HIDS with optimized sensors throughout the organizations. So it is necessary to decide and tricky deployed the IDS where the critical assets can be secured with the threat visibility.

Managing the High Volume of Alerts

HIDS and NIDS are being implemented with the combination of signature and anomaly-based detection techniques. So the sensors alerts as and when some known attack or the traffic outside the listing behaviors. This will impact the bandwidth consumption when the anomalous activity or the DNS traffic transits.

Sometime large false alarms also a significant overhead for the scans of every alerts which is not possible hence such slip of suspect can become a risk in security terms . although most of the IDS loaded with pre defined signatures but still the alerts should be managed well enough to detect without any slips.

Understanding and Investigating Alerts

Since IDS consist of the base line knowledge dataset which can observe a little. Therefore it is necessary to either implement an expert security system which is much capable enough to interpret the alerts and response in efficient time that invokes the suspicious activity sudden.

Knowing How to Respond to Threats

Many organizations that have IDS struggling with the identification of threats a half battle which they stoke but the major big issue is in appropriate response capability. There is need to deploy an effective and skilled responding robust heuristics which can manage the incoming traffic and proper control the threats which breach the security with relevant operations to swiftly remediate them

Risk Analysis for the IDS deployment

When implementing an IDS in any of the organization or the network body, it is necessary to first analyze the risks the internal environments suffers and also enlist the managed services . This could be done by hiring the dedicated security personnel but a managed service avoids such because it again depend on human behavior which could miss the suspect. Hence after listing the risks and finding the ongoing challenges which is more intricate and hard issue.

In present scenario developers from same organization theft the data internally thus the security developers need to give the more precise, faster, scalable techniques to invade the intrusive activities. No doubt the world link approx 26 billion devices by 2025 with the entrance of emerging "IoT" and Big Data age [18] where the huge traffic assigns so the risks of false alarming, unbalanced dataset, low detection rate and response time etc are being commissioning before deployment of any IDS.

Proposed Design

The various methodologies focuses on incorporating the convention to the expert system of IDS. The technique additionally saves time and response in fastest manner without human observation. One such design is thought to be making in future where ML is used with multi level IDS so that none suspect traffic data would be missing. This way the the future proposed design will predict the intrusive activities in best manner as given in Figure 5.

Conclusion

The various methodologies focuses on incorporating the convention to the expert system of IDS. The technique additionally saves time and response in fastest manner without human observation through automatic IDS. Here the different categories of IDS have been discussed which monitor and identify suspicious activity host as well as network. But all are differently work acceding to the individual needs ,numerous works have been explores and categories defined to roadmap the new investigations .it is also being proposed to add the machine learning techniques to aid the IDS which can help in better prediction. Several current researches briefly described to improve the IDS paradigm which opens the various challenges in this field. Thus is proposed a model on which working in future to get better detection by embedding new ML procedure into IDS

References

1. G Tavallaee, M.; Bagheri, E.; Lu,W.; Ghorbani, A.A. A detailed analysis of the KDD CUP 99 data set. In Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, Ottawa, ON, Canada, 8–10 July 2009; IEEE: Manhattan, NY, USA, 2009
2. Fu, Y.; Du, Y.; Cao, Z.; Li, Q.; Xiang, W. A Deep Learning Model for Network Intrusion Detection with Imbalanced Data. Electronics 2022, 11, 898.
3. Anita K. Jones and Robert S. Sielken "Computer System Intrusion Detection A Survey "International Journal of Computer Theory and Engineering, Vol.2, No.6, December, 2010.
4. Neelam Sharma, Saurabh Mukherjee, A Novel Multi-Classifier Layered Approach to Improve Minority Assault Detection in IDS, Procedia Technology, Volume 6, 2012, Pages 913-921.
5. P. Sharma, S. Saxena and Y. Mohan Sharma, "An Efficient Decision Support Model Based on Ensemble Framework of Data Mining Features Assortment & Classification Process," 2018 3rd International Conference on Communication and Electronics Systems (ICCES), 2018, pp. 487-491,
6. Jaiswal, O., Saini, P.K., Shalini, Sharma, Y.M. (2021). Analyze Classification Act of Data Mining Schemes. In: Goyal, D., Gupta, A.K., Piuri, V., Ganzha, M., Paprzycki, M. Second International Conference on Information Management and Machine Intelligence. Lecture Notes in Networks and Systems, vol 166. Springer.
7. Arul, Amudha & Subburathinam, Karthik & Sivakumari, S. "Classification Techniques for Intrusion Detection An Overview.International Journal of Computer Applications., 2013, 76. 33-40.
8. James P. Anderson. Computer Security Threat Monitoring and Surveillance, 1980. Last accessed: Novmeber 30,2008.
9. M. Mehmood, T. Javed, J. Nebhen, S. Abbas, R. Abid et al., "A hybrid approach for network intrusion detection," Computers, Materials & Continua, vol. 70, no.1, pp. 91–107, 2022.
10. [Cao B, Li C, Song Y, Fan X. Network Intrusion Detection Technology Based on Convolutional Neural Network and BiGRU.Comput Intell Neurosci. 2022 Apr 12;2022:1942847
11. Fu Y, Du Y, Cao Z, Li Q, Xiang W. A Deep Learning Model for Network Intrusion Detection with Imbalanced Data. Electronics. 2022; 11(6):898
12. M. Ashfaq Khan and Y. Kim, "Deep learning-based hybrid intelligent intrusion detection system," Computers, Materials & Continua, vol. 68, no.1, pp. 671–687, 2021.
13. Li Y, Xu W, Li W, Li A, Liu Z. Research on hybrid intrusion detection method based on the ADASYN and ID3 algorithms.Math Biosci Eng. 2021 Jan;19(2):2030-2042.
14. J. Dong Lee, H. Soung Cha, S. Rathore and J. Hyuk Park, "M-idm: a multi-classification based intrusion detection model in healthcare iot," Computers, Materials & Continua, vol. 67, no.2, pp. 1537–1553, 2021.
15. Y. S. Sydney and M. Kasongo, "A deep learning method with wrapper based feature extraction for wireless intrusion detection system," Computers & Security. Elsevier, vol. 92, pp. 15, 2020.
16. K. E. S. Hadeel Alazzam and Ahmad Sharieh, "A feature selection algorithm for intrusion detection system based on pigeon inspired optimizer," Expert Systems with Applications. Elsevier, vol. 148, pp. 113249, 2020.
17. U. Ahmad, H. Asim, M. T. Hassan and S. Naseer, "Analysis of classification techniques for intrusion detection," in 2019 Int.Conf. on Innovative Computing,

New Delhi, India, IEEE, pp. 1–6, 2019.
18. A. Hajimirzaei and N. J. Navimipour, "Intrusion detection for cloud computing using neural networks and artificial bee colony optimization algorithm," ICT Express, vol. 5, no. 1, pp. 56–59, 2019.
19. B. Ingre, A. Yadav and A. K. Soni, "Decision tree based intrusion detection system for NSL-KDD dataset," in Int. Conf. on Information and Communication Technology for Intelligent Systems, Springer, pp. 207–218, 2017.
20. S. Aljawarneh, M. Aldwairi and M. B. Yassein, "Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model," Journal of Computational Science, vol. 25, pp. 152–160, 2018.
21. C. Yin, Y. Zhu, J. Fei and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," IEEE Access, vol. 5, pp. 21954–21961, 2017.
22. A. M. Yogita Hande, "A survey on intrusion detection system for software defined networks (sdn)," Research Anthology on Artificial Intelligence Applications in Security. IGI Global, vol. 16, no. 1, pp. 20, 2021.
23. A. R. Javed, M. O. Beg, M. Asim, T. Baker and Al-Bayatti, "Alphalogger: Detecting motion-based sidechannel assault using smartphone keystrokes," Journal of Ambient Intelligence and Humanized Computing, pp. 1–14, 2020.