International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(Review Article)

Machine Learning approaches in IDS

Md Boktiar Hossain [1,*] and Khandoker Hoque [2]

[1] *Department of Information and Communication Engineering, University of Rajshahi, Rajshahi 6205, Bangladesh.*
[2] *Department of Electrical and Electronic Engineering, Brac University, Dhaka, Bangladesh.*

International Journal of Science and Research Archive, 2022, 07(02), 706-715

Publication history: Received on 18 November 2022; revised on 22 December 2022; accepted on 28 December 2022

Article DOI: https://doi.org/10.30574/ijsra.2022.7.2.0303

Abstract

With the rapid expansion of digital infrastructures, cybersecurity threats have become increasingly sophisticated, necessitating advanced protection mechanisms. Traditional security solutions, such as firewalls and rule-based intrusion detection systems (IDS), often fail to detect evolving attack patterns. Machine Learning (ML) has emerged as promising approaches for enhancing IDS capabilities by identifying anomalies and predicting cyber threats with higher accuracy. This paper provides a comprehensive review of ML methodologies applied to intrusion detection systems, focusing on their effectiveness, challenges, and future directions.

Despite their advancements, ML based IDS face several challenges, including data imbalance, high computational complexity, and adversarial attacks that manipulate detection mechanisms. The lack of interpretability in deep learning models hinders their deployment in critical security infrastructures. To address these limitations, future research should focus on explainable AI, federated learning for decentralized threat intelligence, and integration with blockchain technology for enhanced data integrity.

Keywords: Machine Learning (Ml); Intrusion Detection System (Ids); Cybersecurity; Anomaly Detection; Supervised Learning

1. Introduction

The internet is transforming people's jobs, learning, and lifestyles, and today, allowing to the integration of social life and the internet, which increases security threats in various ways. What counts now is learning how to identify network threats and cyberattacks, particularly those previously seen. Cybersecurity is defined as the process of implementing cyber protective measures and policies to protect data, programs, servers, and network infrastructures from unauthorized access or modification. The internet connects the majority of our computer systems and network infrastructure. As a result, cybersecurity emerged as the backbone for practically all types of corporations, governments, and even people to secure data, grow their businesses, and maintain privacy. People send and receive data across network infrastructure, such as a router, that can be hacked and manipulated by outsiders. The increased use of the internet has increased the amount and complexity of data, resulting in the emergence of big data. The constant rise of the internet and extensive data necessitated the creation of a reliable intrusion detection system. Network security is a subset of cybersecurity that safeguards systems connected to a network against malicious activity. The goal is to provide networked computers to ensure data security, integrity, and accessibility. Current cybersecurity research focuses on creating an effective intrusion detection system that can identify both known and new attacks and threats with high accuracy and a low false alarm rate [1, 2].

As organizations and individuals continue to digitize their operations, cyber threats have become more sophisticated, diverse, and frequent. One of the most critical areas of cybersecurity is Intrusion Detection Systems (IDS), which are

[*] Corresponding author: Md Boktiar Hossain

designed to identify and respond to unauthorized activities within a network. Traditional IDS solutions, such as signature-based and rule-based detection, often struggle to keep pace with the evolving nature of cyber threats. Consequently, Machine Learning (ML) has emerged as powerful tools to enhance the efficiency and adaptability of IDS in real-time threat detection and mitigation.

Cyber threats have evolved beyond simple virus attacks to more advanced techniques that exploit system vulnerabilities. Malware, including ransomware, trojans, and worms, continues to be a dominant threat, capable of infiltrating networks and causing significant harm. Phishing attacks, which manipulate users into revealing confidential information, have also grown in sophistication, leveraging social engineering tactics to deceive even the most vigilant users. DDoS attacks overwhelm network resources, rendering services unavailable, while zero-day exploits target undiscovered vulnerabilities, making them particularly difficult to counter. Insider threats pose significant risks, as trusted individuals with access to sensitive data can deliberately or inadvertently compromise security.

One of the major challenges in intrusion detection is the ability to detect and respond to evolving threats in real-time. Traditional IDS mechanisms, such as signature-based antivirus software and rule-based detection systems, often fail to identify new and sophisticated attack vectors. The rapid development of cyber threats outpaces the ability of conventional IDS solutions to adapt, making it imperative to explore advanced methodologies for threat detection and mitigation. The volume and variety of security data generated from networks, endpoints, and cloud environments make manual analysis impractical, necessitating automated solutions that can analyze vast amounts of information efficiently.

2. Intrusion Detection Systems

Intrusion Detection refers to the process of analyzing network traffic and monitoring computer events to identify unusual activities. When this process is implemented through a software application, it is referred to as an Intrusion Detection System (IDS) [3]. IDS plays a crucial role in network security by detecting potential threats before they lead to service disruptions, unauthorized access, or data breaches [4, 5]. IDS can incorporate a graphical user interface that allows users to interact with the system and access various features for testing and training purposes [4]. Network-Based Intrusion Detection System (NIDS) analyzes network packets captured by devices such as routers, whereas Host-Based Intrusion Detection System (HIDS) monitors events occurring within a host system. A hybrid approach integrates both methodologies to enhance detection capabilities [6-10].

2.1. Anomaly Detection

This technique operates on the assumption that anomalous network traffic exhibits statistically low probability and can therefore be distinguished from normal traffic with high confidence. Algorithms based on unsupervised learning and statistical models enable anomaly detection systems to identify novel and previously unknown attacks.

2.2. Misuse Detection

Misuse detection, also known as signature-based detection, identifies threats by comparing network activities against a database of known attack patterns [10, 11]. This method utilizes supervised learning techniques to recognize and prevent malicious or suspicious activities that resemble previously recorded attack behaviors.

2.3. Attack Classification

As networks continue to evolve, cyber threats have become increasingly sophisticated and diverse. Various attack types have been classified, including Denial of Service (DoS), Probe, Remote to User (R2U), Worm, Backdoor, User to Root (U2R), and Trojan attacks [11, 12].

DoS attacks are among the most prevalent threats, aiming to overload network resources and render services inaccessible to legitimate users. Attackers employ multiple techniques to deplete network bandwidth and processing capabilities. In Probe attacks, adversaries scan all devices within a network to identify open ports, which can then be exploited for unauthorized access. The Remote to User (R2U) attack involves an attacker transmitting malicious packets across a network to gain local user privileges. Worms, on the other hand, are self-replicating malicious programs capable of spreading across multiple devices without user intervention [12]. Lastly, the User to Root (U2R) attack occurs when an intruder repeatedly attempts to escalate privileges, ultimately gaining control over critical network resources [11].

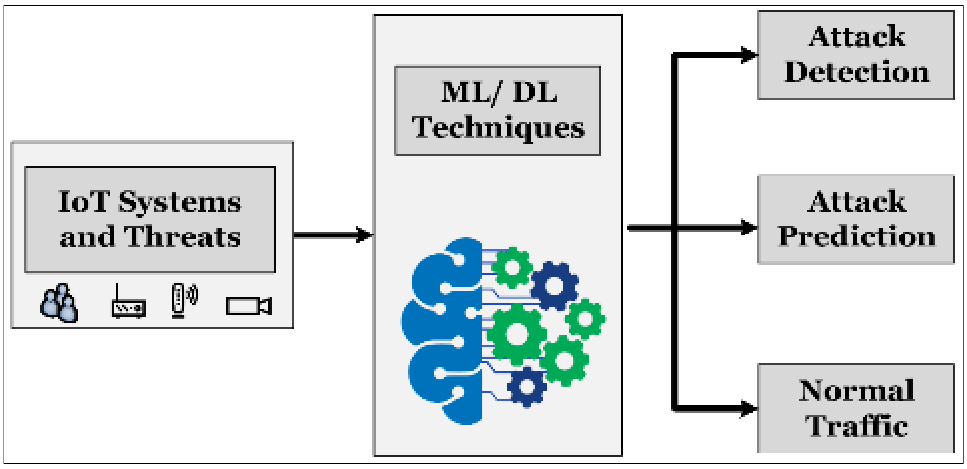

Figure 1 Role of ML Based IDS for IoT system [10].

3. Datasets

When it comes to intrusion detection systems, one should consider the dataset employed to ensure the system's accuracy. Nowadays, applications and networks are growing exponentially, necessitating resilient network security. It can be accomplished by selecting the proper datasets for training and testing. Following that, a summary of the most often used dataset in intrusion detection systems will be discussed.

3.1. KDD CUP 1999

This dataset is the most widely used dataset for intrusion detection, based on the DARPA dataset. This dataset includes basic and high-level TCP connection information such as the connection window but no IP addresses. In addition, this dataset contains over 20 different types of attacks and a record for the test subset [12, 13].

3.2. UNSW-IDS15

Founded in 2015 by Australian Centre for Cyber Security (ACCS). Samples in this dataset contain normal and malicious traffic [14], and it has been collected from three real-world websites; BID (Symantec Corporation), CVE (Common Vulnerabilities and Exposures), and MSD (Microsoft Security Bulletin) and then to generate the dataset, it emulated in a laboratory environment. This dataset has nine attack families, such as worms, DoS, and fuzzers [12].

3.3. CIC-IDS2017

The Canadian Institute generated the dataset in 2017 for Cybersecurity. This dataset contains normal and attack scenarios and includes an abstract behavior for 25 users based on SSH, HTTPS, HTTP, FTP, and email protocols [15].

3.4. NSL-KDD

It is the improved KDD dataset, where a large amount of redundancy has been removed, and an advanced sub-dataset has been created. This dataset utilizes the same KDD99 attributes and belongs to four attack categories: DoS, U2R, R2L, and Probe [15].

3.5. PU-IDS

A derivative dataset from NSL-KDD is generated to extract a statistic from an input data and then utilized to create new synthetic instances. The traffic generator of this dataset obtained the same format and attributes as the NSL-KDD dataset [15].

3.6. DARPA1998

The DARPA1998 dataset [16] was built by the Lincoln laboratory of MIT and is a widely used benchmark dataset in IDS studies. To compile it, the researchers collected Internet traffic over nine weeks; the first seven weeks form the training set, and the last two weeks form the test set. The dataset contains both raw packets and labels. There are five types of labels: normal, denial of service (DOS), Probe, User to Root (U2R) and Remote to Local (R2L). Because raw packets cannot be directly applied to traditional machine learning models, the KDD99 dataset was constructed to overcome this drawback.

3.7. LBNL

The LBNL dataset contains anonymized traffic, which is comprised of only header data. The dataset was generated at the Lawrence Berkley National Laboratory, by gathering real outbound, inbound and routing traffic from two edge routers [38]. It lacked the labeling process and also no extra features were created [38].

3.8. UNSW-NB15

This is a dataset developed at UNSW Canberra by the researchers of [40] for the evaluation of IDS. The researchers used the IXIA PerfectStorm tool to generate a mixture of attack and benign traffic, at the Australian Center of Cyber Security (ACCS) over two days, in sessions of 16 and 15 h. They generated a dataset of size 100 GB in the form of pcap files with a substantial number of novel features. NB15 was planned as a step-up from the KDD99 dataset discussed above. It covers 10 targets: one benign, and nine anomalous, namely: DoS, Exploits, Analysis, Fuzzers, Worms, Reconnaissance, Generic, Shell Code and Backdoors [40]. However, the dataset was designed based on a synthetic environment for producing attack activities.

3.9. ISCX datasets [39]

The Canadian Institute for Cybersecurity has been working on the generation of numerous datasets that are used by independent researchers, universities and private industry around the world. A few datasets relevant to our work are IPS/IDS dataset on AWS (CSE-CIC-IDS2018), IPS/IDS dataset (CICIDS2017), CIC DoS dataset (application-layer), ISCX Botnet dataset, ISCX IDS 2012 dataset, ISCX Android Botnet dataset, and ISCX NSL-KDD dataset. Their latest dataset related to our work is CICIDS2017. This dataset covers benign and the most up-to-date common attacks, which is comparable to the real-world data [41].

4. Intrusion Detection System Landscape

Intrusion Detection Systems (IDS) play a critical role in identifying and mitigating cybersecurity threats in modern networks. IDS solutions are designed to detect unauthorized access, malicious activities, and potential security breaches within an organization's infrastructure. To understand the importance of IDS, it is essential to examine different types of cybersecurity threats that these systems aim to counter.

4.1. Malware

Malware, or malicious software, is a broad category of cyber threats that includes viruses, worms, trojans, ransomware, and spyware. These programs are designed to disrupt, damage, or gain unauthorized access to computer systems. Traditional IDS solutions use signature-based detection to identify known malware, while modern ML and DL-based IDS can detect previously unseen malware variants by analyzing behavioral patterns and anomalies.

4.2. Phishing

Phishing attacks involve fraudulent attempts to obtain sensitive information such as usernames, passwords, and financial data by disguising malicious entities as legitimate sources. These attacks often occur through email, social media, or fake websites. ML-powered IDS solutions can analyze communication patterns, identify suspicious links, and flag potential phishing attempts, reducing the risk of credential theft and financial fraud.

4.3. Distributed Denial-of-Service (DDoS) Attacks

DDoS attacks aim to overwhelm a target system or network by flooding it with excessive traffic, rendering it unavailable to legitimate users. Attackers use botnets to generate massive amounts of requests, exhausting server resources. Traditional IDS struggle with detecting sophisticated DDoS attacks, whereas ML and DL-based approaches can analyze network traffic patterns and detect early signs of an attack, allowing for proactive mitigation measures.

4.4. Zero-Day Exploits

Zero-day exploits target unknown or unpatched vulnerabilities in software and hardware before developers can release security updates. These attacks are particularly dangerous because traditional security measures cannot detect them without predefined signatures. Advanced IDS solutions leverage anomaly detection techniques in ML and DL to identify deviations from normal system behavior, providing early warnings of potential zero-day exploits.

4.5. Insider Threats

Unlike external cyber threats, insider threats originate from within an organization, making them challenging to detect. Malicious insiders may misuse their access privileges to steal data, sabotage systems, or leak confidential information. ML-based IDS can analyze user behavior, detect abnormal access patterns, and flag potential insider threats before significant damage occurs.

4.6. Ransomware

Ransomware is a type of malware that encrypts a victim's data and demands payment for decryption. Attackers typically use phishing emails or exploit vulnerabilities to deploy ransomware. DL-based IDS can detect ransomware activity by monitoring file access patterns and identifying encryption behaviors that deviate from normal usage.

Intrusion Detection Systems serve as the frontline defense against these evolving cyber threats. As threat actors continue to develop more sophisticated attack techniques, the integration of ML and DL in IDS is essential for improving detection accuracy, reducing false positives, and enabling proactive threat response. The next section will explore various ML and DL methodologies used in modern IDS and their effectiveness in combating cyber threats.

5. Machine Learning in IDS

Machine Learning (ML) has emerged as powerful tools in enhancing IDS by enabling intelligent threat detection, anomaly detection, and automated response mechanisms. ML techniques utilize algorithms that can learn from data, recognize patterns, and make informed decisions without explicit programming. This capability makes ML-based IDS solutions more adaptive to new and evolving threats. Supervised learning models, such as Support Vector Machines (SVMs) and Random Forests, can classify network traffic as benign or malicious based on labeled datasets, while unsupervised learning models, such as clustering techniques, can identify unknown attack patterns through anomaly detection.

5.1. Supervised Learning Techniques

Supervised learning methods use labeled datasets to train models in distinguishing between normal and malicious activities. The key techniques include:

- **Naïve Bayes (NB)**: A probabilistic classifier based on Bayes' theorem, often used for spam detection and IDSs due to its simplicity.
- **k-Nearest Neighbors (kNN)**: A distance-based classifier that assigns labels based on the closest data points in feature space.
- **Decision Tree (DT)**: A rule-based model that creates a tree-like structure for decision-making in intrusion detection.
- **Support Vector Machine (SVM)**: A powerful classification algorithm that separates data using hyperplanes, effective for binary classification tasks in IDSs.
- **Random Forest (RF)**: An ensemble of decision trees that improves accuracy and reduces overfitting in intrusion detection.
- **Ensemble Learning (EL)**: A combination of multiple models (e.g., boosting or bagging) to improve the IDS detection rate and reduce false positives.

5.2. Unsupervised Learning Techniques

Unsupervised learning methods do not require labeled data but instead detect anomalies or clusters based on patterns and statistical properties. Key techniques include:

- **K-Means**: A clustering algorithm that groups similar data points, useful for identifying anomalous network behaviors.
- **Principal Component Analysis (PCA)**: A dimensionality reduction technique that helps in detecting anomalies by reducing feature complexity and identifying deviations from normal patterns.

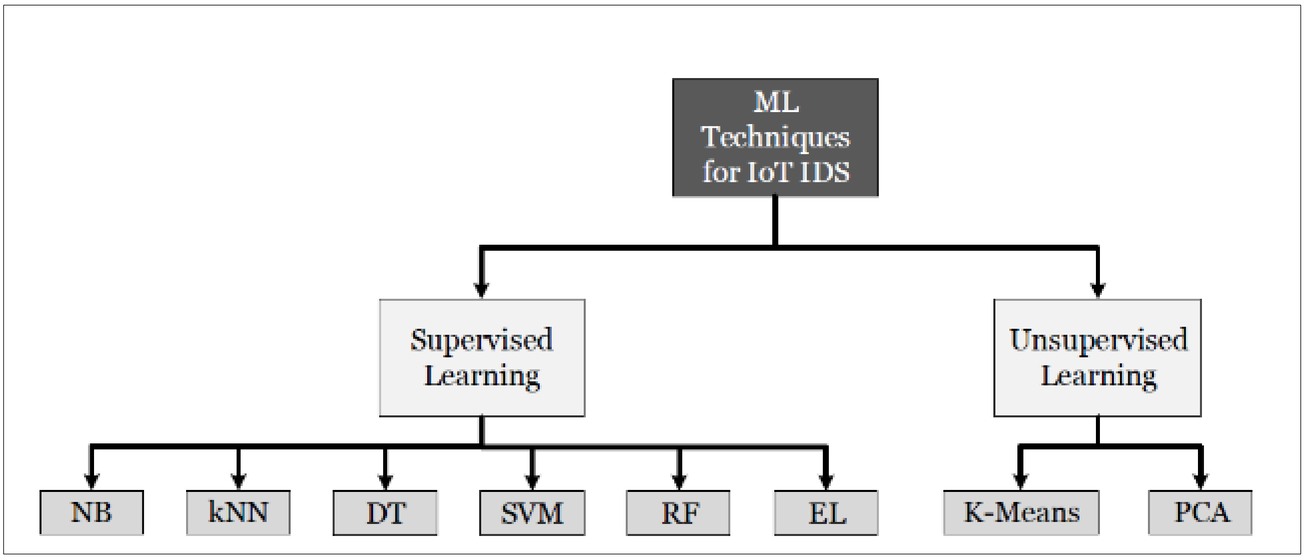

Figure 2 A taxonomy of ML Techniques for IoT-based IDSs [10].

Table 1 gives a brief overview of ML methods, their advantages and limitations along with reference to related research work conducted. In the end, Table 2 summarizes research works conducted to propose IDSs using various ML methods, as detailed below.

Table 1 Taxonomy of ML based methods for IoT systems security.

ML	Attack Types Handled	Pros	Cons
KB [17, 18, 19, 21]	HTTP attacks (Buffer overflow, Shell attacks) [19], DoS, Probe, R2L [18]	It requires very few samples for training [20]. It can classify in both binary and multi-label classification. It shows robustness to irrelevant features.	It fails to take into account interdependencies between features for classification purposes, which affect its accuracy [21].
KNN [22-26]	U2R, R2L, Flooding attacks, DoS, DDoS	Simple to use.	Determining optimal value of K and identifying missing nodes are challenging.
DT [27-29]	DDoS [29], U2R, R2L [27]	Easy and simple to use method.	It requires bigger storage. It is computationally complex It is easy to use only if few DTs are used.
SVM [30-32]	Scan, DDoS (TCP, UDP flood), smurf, portsweep	SVMs are highly scalable due to simplicity and are capable of performing tasks like anomaly-based intrusion detection in real-time including online learning. SVMs are considered suitable for data containing a large number of feature attributes. SVMs use lesser storage and memory.	The use of optimal kernel function in SVM, which is used to separate the data when it is not linearly separable, remains a challenge to achieve desired classification speed. It is difficult to understand and interpreting SVM-based models.
EL [33-35]	DoS, Probe, R2L, U2R attacks	It is robust to overfitting. Performs better than a single classifier.	Increased time complexity, due to the use of multiple classifiers in parallel.

		It reduces variance.	
RF [37, 38]	DoS, Probe, R2L, U2R	It produces a more robust and accurate output which is resistant to overfitting. It requires substantially fewer inputs and does not require the process of feature selection.	Since RF constructs several DTs, its use may be impractical in real-time applications requiring large dataset.

ML has become integral to optimizing resource management in cloud computing environments discussed in [36]. These techniques enable predictive analytics for workload forecasting, intelligent scheduling, and automated resource allocation, ensuring efficient utilization of computational resources while minimizing latency and energy consumption [36].

6. Challenges and Future Research Directions

The integration of Machine Learning (ML) in Intrusion Detection Systems (IDSs) for IoT networks faces multiple challenges, particularly in anomaly-based Network Intrusion Detection Systems (NIDSs) [39-42]. These challenges stem from data quality issues, real-time detection constraints, and the complexity of modeling IoT traffic. Below are the key technical challenges:

6.1. Data Availability and Quality Issues

- A high-quality dataset is crucial for training and evaluating IDS models.
- Most publicly available datasets lack critical attributes, such as:
- Missing labels (attack vs. normal traffic).
- Incomplete network flow features (e.g., missing headers, timestamps, or metadata).
- Absence of raw packet capture (pcap) files for detailed traffic analysis.
- Limited device coverage, where datasets focus on specific IoT devices, reducing generalizability.
- Creating a comprehensive, labeled dataset that includes diverse IoT traffic and attack scenarios remains an open research challenge.

6.2. Real-Time and Online Anomaly Detection

- Anomaly-based IDSs require continuous learning of normal behavior to detect intrusions.
- Training an IDS assumes a clean dataset without attack traffic, but real-world data often includes noisy or malicious traffic during the learning phase.
- This noise leads to false alarms, impacting IDS reliability in real-time environments.
- Efficiently implementing real-time anomaly detection while filtering out noise remains a significant technical hurdle.

6.3. Model Bias and High False Alarm Rates

- Anomaly detection models attempt to learn normal traffic patterns, but IoT networks are highly heterogeneous (e.g., different device types, communication protocols).

6.3.1. Class imbalance issue

- Normal traffic dominates the dataset, leading models to bias towards normal behavior.
- This results in high false-positive rates (normal traffic incorrectly flagged as malicious).
- Simultaneously, the inability to cover all variations of normal traffic increases false-negative rates (malicious activity being undetected).
- Designing ML models that balance normal vs. attack data while reducing false alarms remains a core challenge.

7. Conclusion

Machine Learning (ML) techniques significantly enhance Intrusion Detection Systems (IDS) by enabling automated threat detection and adaptive responses. Supervised learning models (SVM, RF, DT) effectively classify network traffic, but suffer from data imbalance and interpretability issues. Unsupervised learning methods (K-Means, PCA) help in detecting novel threats but face challenges due to high false positive rates. The integration of explainable AI, federated learning, and blockchain can improve the security and reliability of IDS in the future. Challenges such as data imbalance, computational complexity, and adversarial attacks need research to optimize ML-based IDS solutions

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] D. I. Edeh, ``Network intrusion detection system using deep learning technique,'' M.S. thesis, Dept. Comput., Univ. Turku, Turku, Finland, 2021.

[2] Halbouni, A., Gunawan, T. S., Habaebi, M. H., Halbouni, M., Kartiwi, M., & Ahmad, R. (2022). Machine learning and deep learning approaches for cybersecurity: A review. IEEE Access, 10, 19572-19585.

[3] Al-amri, R.; Murugesan, R.K.; Man, M.; Abdulateef, A.F.; Al-Sharafi, M.A.; Alkahtani, A.A. A Review of Machine Learning and Deep Learning Techniques for Anomaly Detection in IoT Data. Appl. Sci. 2021, 11, 5320. https://doi.org/10.3390/app11125320

[4] Y. Xin, L. Kong, Z. Liu, Y. Chen, Y. Li, H. Zhu, and M. Gao, ``Machine learning and deep learning methods for cybersecurity,'' IEEE Access, vol. 6, pp. 3536535381, 2018, doi: 10.1109/ACCESS.2018.2836950

[5] H. Dhillon, ``Building effective network security frameworks using deep transfer learning techniques,'' M.S. thesis, Dept. Comput. Sci., Western Univ., London, ON, Canada, 2021.

[6] Bharati, S., Mondal, M. R. H., Podder, P., & Prasath, V. S. (2022). Federated learning: Applications, challenges and future directions. International Journal of Hybrid Intelligent Systems, 18(1-2), 19-35.

[7] Bharati, S., & Podder, P. (2022). Machine and deep learning for iot security and privacy: applications, challenges, and future directions. Security and communication networks, 2022(1), 8951961.

[8] Sirajul Islam, M., Rouf, M. A., Shahariar Parvez, A. H. M., & Podder, P. (2022). Machine Learning-Driven Algorithms for Network Anomaly Detection. In Inventive Computation and Information Technologies: Proceedings of ICICIT 2021 (pp. 493-507). Singapore: Springer Nature Singapore.

[9] Pinto Kumar Paul; Subrato Bharati; Prajoy Podder; M. Rubaiyat Hossain Mondal, "10 The role of IoMT during pandemics," in Computational Intelligence for Managing Pandemics , De Gruyter, pp.169-186.

[10] Asharf, J.; Moustafa, N.; Khurshid, H.; Debie, E.; Haider, W.; Wahab, A. A Review of Intrusion Detection Systems Using Machine and Deep Learning in Internet of Things: Challenges, Solutions and Future Directions. Electronics 2020, 9, 1177. https://doi.org/10.3390/electronics9071177

[11] M. Alkasassbeh and M. Almseidin, ``Machine learning methods for network intrusion detection,'' 2018, arXiv:1809.02610

[12] P.Wu, ``Deep learning for network intrusion detection: Attack recognition with computational intelligence,'' M.S. thesis, School Comput. Sci. Eng., Univ. New South Wales, Sydney NSW, Australia, 2020.

[13] M. Ring, S.Wunderlich, D. Scheuring, D. Landes, and A. Hotho, ``A survey of network-based intrusion detection data sets,'' Comput. Secur., vol. 86, pp. 147-167, Sep. 2019.

[14] T. Hamed, R. Dara, and S. C. Kremer, ``Network intrusion detection system based on recursive feature addition and bigram technique,'' Comput. Secur., vol. 73, pp. 137-155, Mar. 2018.

[15] A. Kim, M. Park, and D. H. Lee, ``AI-IDS: Application of deep learning to real-time web intrusion detection,'' IEEE Access, vol. 8, pp. 70245-70261, 2020.

[16] DARPA1998 Dataset. 1998. Available online: http://www.ll.mit.edu/r-d/datasets/1998-darpa-intrusion-detection-evaluation-dataset (accessed on 16 October 2019).

[17] Panda, M.; Patra, M.R. Network intrusion detection using naive bayes. Int. J. Comput. Sci. Netw. Secur. 2007, 7, 258–263.

[18] Mukherjee, S.; Sharma, N. Intrusion detection using naive Bayes classifier with feature reduction. Procedia Technol. 2012, 4, 119–128.

[19] Swarnkar, M.; Hubballi, N. OCPAD: One class Naive Bayes classifier for payload based anomaly detection. Expert Syst. Appl. 2016, 64, 330–339.

[20] Box, G.E.; Tiao, G.C. Bayesian Inference in Statistical Analysis; John Wiley & Sons: Hoboken, NJ, USA, 2011; Volume 40.

[21] Ng, A.Y.; Jordan, M.I. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. In Advances in Neural Information Processing Systems; 2002; pp. 841–848. Available online: https://ai.stanford.edu/~ang/papers/nips01-discriminativegenerative.pdf (accessed on 13 July 2020).

[22] Adetunmbi, A.O.; Falaki, S.O.; Adewale, O.S.; Alese, B.K. Network intrusion detection based on rough set and k-nearest neighbour. Int. J. Comput. ICT Res. 2008, 2, 60–66.

[23] Li, L.; Zhang, H.; Peng, H.; Yang, Y. Nearest neighbors based density peaks approach to intrusion detection. Chaos Solitons Fractals 2018, 110, 33–40.

[24] Su, M.Y. Real-time anomaly detection systems for Denial-of-Service attacks by weighted k-nearest-neighbor classifiers. Expert Syst. Appl. 2011, 38, 3492–3498.

[25] Pajouh, H.H.; Javidan, R.; Khayami, R.; Ali, D.; Choo, K.K.R. A two-layer dimension reduction and two-tier classification model for anomaly-based intrusion detection in IoT backbone networks. IEEE Trans. Emerg. Top. Comput. 2016.

[26] Li, W.; Yi, P.; Wu, Y.; Pan, L.; Li, J. A new intrusion detection system based on KNN classification algorithm in wireless sensor network. J. Electr. Comput. Eng. 2014, 2014.

[27] Goeschel, K. Reducing false positives in intrusion detection systems using data-mining techniques utilizing support vector machines, decision trees, and naive Bayes for off-line analysis. In Proceedings of the SoutheastCon 2016, Norfolk, VA, USA, 30 March–3 April 2016; pp. 1–6.

[28] Kim, G.; Lee, S.; Kim, S. A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Syst. Appl. 2014, 41, 1690–1700.

[29] Alharbi, S.; Rodriguez, P.; Maharaja, R.; Iyer, P.; Subaschandrabose, N.; Ye, Z. Secure the internet of things with challenge response authentication in fog computing. In Proceedings of the 2017 IEEE 36th International Performance Computing and Communications Conference (IPCCC), San Diego, CA, USA, 10–12 December 2017; pp. 1–2.

[30] Liu, Y.; Pi, D. A novel kernel svm algorithm with game theory for network intrusion detection. KSII Trans. Internet Inf. Syst. 2017, 11.

[31] Hu, W.; Liao, Y.; Vemuri, V.R. Robust Support Vector Machines for Anomaly Detection in Computer Security. ICMLA. 2003, pp. 168–174. Available online: https://web.cs.ucdavis.edu/~vemuri/papers/rvsm.pdf (accessed on 13 July 2020).

[32] Wagner, C.; François, J.; Engel, T. Machine learning approach for ip-flow record anomaly detection. In International Conference on Research in Networking; Springer: Berlin, Germany, 2011; pp. 28–39.

[33] Aburomman, A.A.; Reaz, M.B.I. A novel SVM-kNN-PSO ensemble method for intrusion detection system. Appl. Soft Comput. 2016, 38, 360–372.

[34] Gaikwad, D.; Thool, R.C. Intrusion detection system using bagging ensemble method of machine learning. In Proceedings of the 2015 International Conference on Computing Communication Control and Automation, Pune, India, 26–27 February 2015; pp. 291–295.

[35] Reddy, R.R.; Ramadevi, Y.; Sunitha, K. Enhanced anomaly detection using ensemble support vector machine. In Proceedings of the 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), Chirala, India, 23–25 March 2017; pp. 107–111.

[36] Tanvir Mahmud, "ML-driven resource management in cloud computing", World Journal of Advanced Research and Reviews, 2022, 16(03), 1230-1238 (Accepted).

[37] Chang, Y.; Li, W.; Yang, Z. Network intrusion detection based on random forest and support vector machine. In Proceedings of the 2017 IEEE International Conference on Computational Science and Engineering (CSE) and

IEEE International Conference on Embedded and Ubiquitous Computing (EUC), Guangzhou, China, 21–24 July 2017; Volume 1, pp. 635–638.

[38] Zhang, J.; Zulkernine, M. A hybrid network intrusion detection technique using random forests. In Proceedings of the First International Conference on Availability, Reliability and Security (ARES'06), Vienna, Austria, 20–22 April 2006; p. 8.

[39] Bhuyan, M.H.; Bhattacharyya, D.K.; Kalita, J.K. Towards Generating Real-life Datasets for Network Intrusion Detection. IJ Netw. Secur. 2015, 17, 683–701.

[40] Sharafaldin, I.; Lashkari, A.H.; Ghorbani, A.A. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. In Proceedings of the 4th International Conference on Information Systems Security and Privacy, ICISSP, Funchal, Portugal, 22–24 January 2018; pp. 108–116.

[41] Moustafa, N.; Slay, J. UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In Proceedings of the 2015 Military Communications and Information Systems Conference (MilCIS), Canberra, Australia, 10–12 November 2015; pp. 1–6.

[42] Koroniotis, N.; Moustafa, N.; Sitnikova, E.; Turnbull, B. Towards the development of realistic botnet dataset in the internet of things for network forensic analytics: Bot-IoT dataset. Future Gener. Comput. Syst. 2019, 100, 779–796.